CLASSICAL ELECTRICITY AND MAGNETISM

Wolfgang K. H. Panofsky

University of California

4th Edition, Revised 1951

DISCLAIMER

TABLE OF CONTENTS

Page

INTRODUCTION

The first part of this course deals with the derivation of Maxwell's equations from the fundamental experimental laws of Coulomb, Ampere and Faraday. Historically the derivation of the equations from these experiments was not possible without logical difficulties and for this reason some authors* prefer to postulate Maxwell's equations initially and then treat the subject as a set of examples in the solution of Maxwell's equations. Although this approach offers greater freedom from logical difficulties, it is not the one followed in this course, since it does not represent the manner in which physical theory evolves in practice. We shall therefore proceed from the experimental facts, and as a consequence will have to draw somewhat doubtful inferences, at various times, which will need further experimental verification. This approach is, however, the approach which is the most common one used in the development of any physical theory.

The second part of this course deals with problems associated with the theory of the electron. The problems are treated entirely classically with special emphasis however as to the features and particularly the difficulties which have a more general significance. The special theory of relativity is treated more thoroughly than is customary in a course in electricity and in particular its experimental basis is discussed.

The rationalized MKS system of units will be used in this course. This is a system which combines practical electromagnetic units with the meter kilogram mechanical units. In this system of units Maxwell's electromagnetic field equations will appear in a form which does not explicitly involve the velocity of light as a factor. The MKS units are rationalized in the sense that all equations describing phenomena having spherical symmetry will contain a factor of 4π, while all equations

---

* e.g. Stratton, Electromagnetic Theory, McGraw Hill, 1941

describing phenomena having cylindrical symmetry will contain a factor of 2π, and all equations that are expressed in general vector language will not contain any multiple of π. The reason for adopting this system of units, in this course, is not that its superiority or inferiority over the other systems is a matter of great significance, but simply that the majority of modern reference books on the subject are now written in this system of units.

These notes are the results of lectures given 1947/48 and 1948/49. The first edition was prepared with the aid of Roger Wallace and Howard Chang. The second edition incorporates valuable corrections contributed by various authors, particularly Richard Madey and Lee Aamodt.

CHAPTER 1 THE ELECTROSTATIC FIELD

The static interaction between material bodies can be described by two alternative methods: Either by formulating the action at a distance between the interacting bodies or by separating the interaction process into the production of a field by one system and the action of the field on another system. These approaches are physically indistinguishable in the static case if correctly formulated, but in the non-static case, i.e.,in the case of time varying sources one is forced to ascribe physical reality to the field owing to the finite velocity of propagation of the interaction.* We shall therefore even formulate the electrostatic interactions as a field theory and then study the extension of the theory to non-static cases. A field theory must satisfy the requirement that the action on a given volume V surrounded by a surface S can be fully described as a function of the field on S and thus be describable without reference to the source.

Field theories applicable to various types of interaction differ by the number of parameters necessary to define the field and the symmetry character of the field. The electric field is a three-dimensional vector field, i.e. a field definable by the specification of 3 components.

All vector fields are uniquely defined if their circulation density and source density are given functions of the coordinates of all points in space. Also, all electric fields are linear fields which obey the principle of superposition. The program for deriving Maxwell's equations consists of attempting to obtain a set of equations which will independently derive the source and circulation densities, that is the divergence and curl of the field vectors describing the electromagnetic fields.

We shall first consider the electrostatic field only in a vacuum. The electric field is defined by the force which is produced on a test charge q by

*Even here it is possible, in fact sometimes advantageous, to replace the field concept by the concept of "delayed direct interation".

the equation:

$$\lim_{q \to 0} \frac{\vec{F}}{q} = \vec{E} \qquad (1.1)$$

q = test charge (Coulombs)

$\vec{F}$ = force on test charge q (Newtons)

$\vec{E}$ = electric field defined by this equation (Volts per meter)

The definition expressed by equation (1.1) is a suitable one only if macroscopic phenomena are involved. The fundamental difficulty in equation (1.1) lies in the finite magnitude of the charge of an electron which does not permit the limit indicated as $q \to 0$ to be experimentally carried to an indefinitely small value of the charge. On the other hand, the necessity for introducing the limit is that it is assumed in the definition of an electric field that the test charge will not influence the behavior of the sources of the field and that for that reason the test charge must be vanishingly small compared to all the sources of the field. This restriction therefore limits the validity of this definition to cases where the sources producing the field are equivalent to a large number of unit electronic charges.

Coulomb's experimentally established law for the force between two point charges (formulated as an action at a distance law):

$$\vec{F}_2 = \frac{q_1 q_2 \vec{r}_{12}}{4\pi k_o r^3} = \frac{-q_1 q_2}{4\pi k_o} \vec{\nabla}\left(\frac{1}{r}\right) \qquad (1.2)$$

$\vec{r}_{12}$ = the radius vector position of charge q_2 measured from an origin located at charge q_1. (Meters)

$k_o = (10^{-7}/4\pi c^2$ Farads/Meter)

$\vec{F}_2$ = the force on charge q_2 due to the presence of charge q_1.

$\vec{\nabla}$ = the gradient operator acting on the coordinates of the charge q_2.

permits immediately by comparison with equation (1.1) the statement:

$$\vec{E} = \frac{q}{4\pi k_o} \frac{\vec{r}}{r^3} = -\frac{q}{4\pi k_o} \vec{\nabla}\left(\frac{1}{r}\right) \qquad (1.3)$$

giving the electric field $\vec{E}$ of a point charge q at a position given by $\vec{r}$ whose origin is at q. (q corresponds to q_1 en equation 1.2.) We might note that by the vector $\vec{r}$ we mean the vector extending from a source point to a field point; unless differently specified the vector operator $\vec{\nabla}$ operates on the coordinates of the field point. In accordance with this convention:

$$(1.4) \quad \vec{\nabla} r = + \frac{\vec{r}}{r} \; ; \; \vec{\nabla} \cdot \vec{r} = 3$$

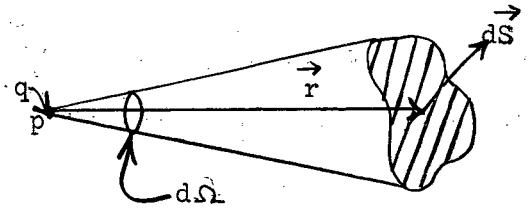
$$\vec{r} = \vec{i}(x_{Field} - x_{Source}) + \vec{j}(y_{Field} - y_{Source}) + \vec{k}(z_{Field} - z_{Source})$$

According to Gauss' flux theorem:*

$$(1.6) \quad \iint_S \vec{E} \cdot \vec{ds} = \frac{q}{k_o}$$

---

* Gauss' electric flux theorem is a direct consequence of Coulomb's law, since if we consider an element of surface $\vec{dS}$, as shown in Figure (1.1),



$\vec{dS}$ = element of surface expressed as a vector directed along the outward normal to the element.
$d\Omega$ = the solid angle subtended by $\vec{dS}$ at the point p.

Figure (1.1)

at a distance $\vec{r}$ from a charge q at a point p, we can, by taking the dot product of both sides of equation (1.3) with $\vec{dS}$, secure:

$$\vec{E} \cdot \vec{dS} = \frac{1}{4\pi k_o} \frac{q}{r^3} \vec{r} \cdot \vec{dS} = \frac{q}{4\pi k_o} \, d\Omega$$

The last equality comes from the expression (1.5) for the solid angle:

$$(1.5) \qquad d\Omega = \left(\frac{\vec{r}}{r} \cdot \vec{dS} \right) \frac{1}{r^2} = \frac{\vec{r} \cdot \vec{dS}}{r^3}$$

Now integrating equation (1.5) over a closed surface which includes the point p, Gauss' theorem results, since:

$$\iint_S d\Omega = 4\pi$$

and using Gauss' divergence theorem:**

$$(1.7)\quad \iint_S \mathbf{E}\cdot d\mathbf{s} = \iiint \nabla\cdot\mathbf{E}\, dv$$

together with the fact that:

$$\iiint_v \rho\, dv = q$$

ρ = the charge density per unit volume at the point where the electric field is $\vec{E}$.

equation (1.3) can be put into the form:

$$(1.9)\quad \vec{\nabla}\cdot\vec{E} = \frac{\rho}{k_o}$$

It further follows from equation (1.3) and the fact that the curl of the gradient of a scalar vanishes and from the principle of superposition, that:

$$(1.10)\quad \vec{\nabla}\times\vec{E} = 0$$

The electrostatic field is thus seen to be solely derived from sources and is therefore irrotational. It then follows from the theorem, that the curl and the divergence of a vector field completely define the vector field, that the electrostatic field is completely defined by a charge distribution.

Since the electric field is irrotational, we may define an electrostatic potential ϕ given by:

$$(1.11)\quad \vec{E} = -\vec{\nabla}\phi$$

ϕ = the potential at the point where the electric field is $\vec{E}$. (Volts)

In Cartesian coordinates we see from equation (1.11) that:

$$(1.12)\quad E_x = \frac{-\partial\phi}{\partial x}, E_y = \frac{-\partial\phi}{\partial y}, E_z = \frac{-\partial\phi}{\partial z}$$

(E_x, E_y, E_z) = the components of the vector field $\vec{E}$ parallel to the x,y,z, axes respectively.

and from the condition of equation (1.10) we have:

$$(1.13)\quad \frac{\partial E_y}{\partial z} = \frac{\partial E_z}{\partial y}, \quad \frac{\partial E_z}{\partial x} = \frac{\partial E_x}{\partial z}, \quad \frac{\partial E_x}{\partial y} = \frac{\partial E_y}{\partial x}$$

**Gauss' divergence theorem is a general vector analytic expression holding for for any vector field.

which assures that the definition of ϕ in equation (1.11) implies a unique potential within an additive constant.

(1.14) Stokes' Theorem:

$$\iint_S (\vec{\nabla} \times \vec{E}) \cdot d\vec{S} = \oint \vec{E} \cdot d\vec{\ell}$$

$d\vec{\ell}$ = the vector tangent to a closed path of integration.

when combined with equation (1.10) yields:

(1.15) $$\oint \vec{E} \cdot d\vec{\ell} = 0$$

which shows that the electrostatic field is a conservative field; that is no work is done on a test charge if it is moved around a closed path in the field. Therefore no continuous external sources of power are needed to maintain an electro-static field. Since the work done in moving a test charge from one point to another is independent of the path, we can therefore uniquely define the work necessary to carry a unit charge from an infinite distance up to a given point as the potential of that point. This definition will not lead to any difficulties if finite sources are considered. If one considers sources of less than three dimensions, that is sources infinite in one or more directions, then this definition will lead to difficulties and a point other than infinity has to be taken for a reference position. From equations (1.9) and (1.11) we can easily obtain Poisson's equation:

(1.16) $$\nabla^2 \phi = \frac{-\rho}{k_o}$$

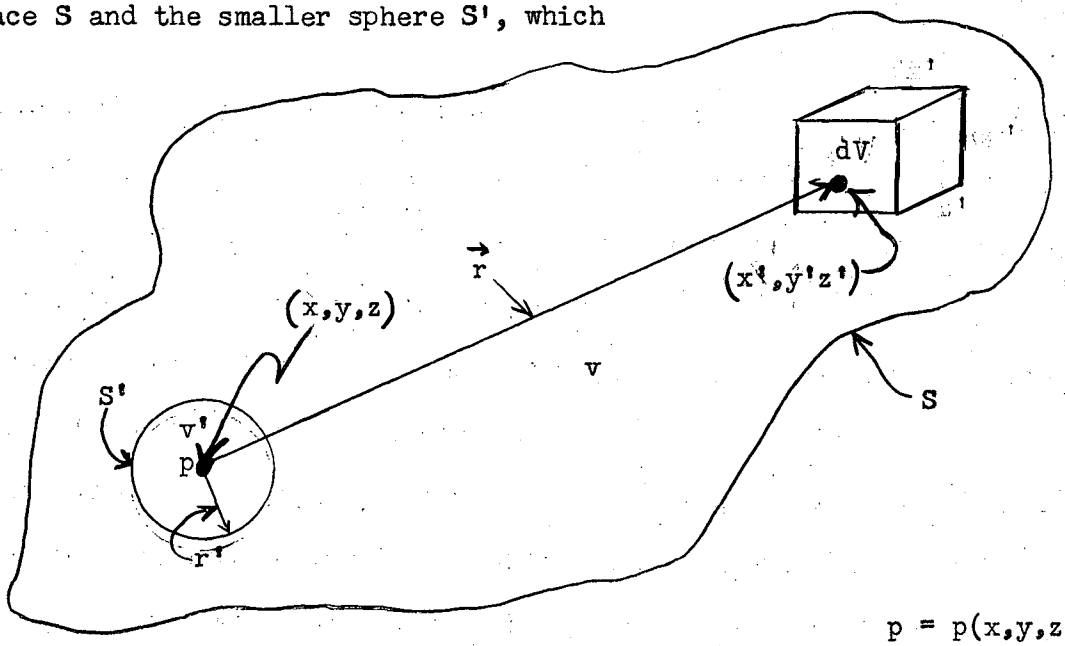
and in a region of zero charge density, Laplace's equation:

(1.17) $$\nabla^2 \phi = 0$$

The electrostatic potential at a given point was defined in equation (1.11) in terms of the electric field at that point. In order to obtain an explicit expression for the potential due to a known distribution of charge within a finite region of space, as well as in terms of the boundary values of the potential and its derivatives, we make use of Green's theorem:

(1.18) $$\iiint_V (\phi \nabla^2 \psi - \psi \nabla^2 \phi) dv = \iint_S (\phi \vec{\nabla} \psi - \psi \vec{\nabla} \phi) \cdot d\vec{S}$$

where ϕ and Ψ are scalar functions of position and are analytic with analytic derivatives in the region of integration and are also analytic on its boundary. Let us make ϕ the electrostatic potential defined in (1.11) and let $\Psi = \frac{1}{r}$ the unit point source solution of Laplace's equation. Since this point solution has a singularity at p (see Figure 1.2) we will integrate only over the region contained between the surface S and the smaller sphere S', which



$p = p(x,y,z)$

$\vec{r} = \vec{r}(x,y,z;x',y',z')$

Figure (1.2)

encloses the point p. Noting that:

(1.19) $\nabla^2\left(\frac{1}{r}\right) = 0$ and $\vec{\nabla}\left(\frac{1}{r}\right) = \frac{\vec{r}}{r^3}$

Thus:

(1.20) $\nabla^2\Psi = 0$ and $\vec{\nabla}\Psi = -\frac{\vec{r}}{r^3}$

We now have on substituting from (1.19) and (1.20) into (1.18):

(1.21) $$-\iiint_{V-V'} \frac{1}{r}\nabla^2\phi\, dv = \iint_S \left(\frac{-\phi\vec{r}}{r^3} - \frac{\vec{\nabla}\phi}{r}\right)\cdot d\vec{S} - \iint_{S'} \left(\frac{-\phi\vec{r}}{r^3} - \frac{\vec{\nabla}\phi}{r}\right)\cdot d\vec{S}'$$

The minus sign before the second term on the right side arises from the inward direction of the normal on the sphere S'. If we now shrink S' until it is small enough so that ϕ and $\vec{\nabla}\phi$ become essentially constant over S' the second term on the right side of (1.21) becomes:

$$(1.22)\quad \iint_{S'} \left(\phi \frac{\vec{r}}{r^3} + \frac{\vec{\nabla}\phi}{r} \right) \cdot dS' \equiv \phi_{s'} \iint_{S'} \frac{r'dS'}{r'^3} + \iint_{S'} \frac{\vec{\nabla}\phi \cdot d\vec{S'}}{r'}$$

If we now pass to the limit as $r' \longrightarrow 0$, we have:

$$(1.23)\quad \lim_{r' \longrightarrow 0} \iint_{S'} \left(\frac{\phi \vec{r}}{r^3} + \frac{\vec{\nabla}\phi}{r} \right) \cdot d\vec{S} = 4\pi\phi_p$$

Note that the term in $\vec{\nabla}\phi$ vanishes since it is of order r'. We thus obtain on substitution of (1.23) into (1.21) and solving for ϕ_p the potential at the point p: (The integration is over the primed coordinates.)

$$(1.24)^{*}\quad \phi_p = \frac{1}{4\pi} \left[\frac{1}{k_o} \iiint_v \frac{\rho}{r}\, dv + \iint_S \left(\phi \frac{\vec{r}}{r^3} + \frac{\vec{\nabla}\phi}{r} \right) \cdot dS \right]$$

The first term of (1.24) is simply the contribution to the potential by the volume charge distribution within v, since by Poisson's equation, the integrand is the charge density, evaluated at the point of integration, divided by the distance between the point of integration and the point of observation. Note that the distance r is a

---

* We can formally obviate the necessity of introducing the sub-volume V' by writing Poisson's equation for a point charge using the Dirac δ - function:

$$(1.25)\quad \nabla^2\left(\frac{1}{r}\right) = -4\pi\delta(r)$$

where:

$$\delta(r) = 0, \quad r \neq 0$$

$$\iiint_v \delta(r)dv = 1$$

Eq. (1.24) then follows directly by substitution of (1.25) into Green's theorem (1.18). We shall have occasion to use this simplified procedure although its justification rests in the above formal proof.

function of the coordinates both of the point of observation, p, and of the point of integration (x', y', z').

The second and third terms of equation (1.24) are surface terms; that is they summarize the effect of the charge distribution outside the region v which is not contained explicitly in the integral over the charge density of the first term. We therefore conclude that the potential at any point within S is uniquely determined by the charge distribution within S and by the values which ϕ and $\vec{\nabla}\phi$ have at all points on the surface S. In particular the potential within a charge-free volume is uniquely determined by the potential and its normal derivative over the surface enclosing the volume. We have only shown that knowledge of the potential and its normal derivative over the surface is sufficient to determine the potential uniquely inside, but we have not shown that these two pieces of information are necessary. As we shall see later, it is in fact sufficient in a charge-free region to have either the potential or its normal derivative given over a surface in order to determine the potential at every point within the surface to within an arbitrary additive constant. The reason for this sufficiency is the fact that ϕ and $\vec{\nabla}\phi$ may not be independently specified over the surface, since ϕ must be a solution of Laplace's equation. The second and third terms of equation (1.24) which define the potential in terms of the boundary values will be interpreted later by stating that the effect of any charge distribution outside of S is equivalent to a charge distribution on S and a dipole layer distribution on S.

If we extend the surface S to include all charges in space, and arbitrarily expand S away from such charges, then the second and third terms of equation (1.24) vanish. This follows since the integrands in ϕ and $\vec{\nabla}\phi$ vary at large distances at least to the inverse third power. Since the surface of integration increases as r^2 both terms vanish to at least to order $\frac{1}{r}$. We conclude therefore that the potential can be calculated by the direct superposition of the individual potentials of all

the volume charge distribution, but that we can, if we wish, replace any part of the distribution by an equivalent surface charge layer and dipole layer distribution.

The volume term of equation (1.24) can be looked on as being a particular integral of Poisson's equation, while the surface terms are complementary integrals of the differential equation in the sense that they are general solutions of the homogeneous equation; that is, Laplace's equation.

CHAPTER 2 FIELD SINGULARITIES

In the last chapter the solution of the potential problem was divided into boundary contributions and into a volume integral extending over the source-charges. These charge integrals will not lead to singular values of the potentials if the charge distributions are volume distributions. If the charges are considered to be surface, line, or point charges, then singularities will result as shown in Table 2.1.

Type of Charge Distribution	Behavior of Potential near distribution	Behavior of field Near distribution
Surface	r	Constant
Line	$\log r$	r^{-1}
Point	r^{-1}	r^{-2}
Point 2^n pole	r^{-n-1}	r^{-n-2}

TABLE (2.1)

Note that if either surface or line charges are infinite (i.e. the fields are considered 1 or 2 dimensional) the potential cannot be referred to infinity. Let us now discuss in more detail the nature of the potentials corresponding to such singular sources.

The potential ϕ_p ; at the point p:

$$(2.1) \qquad \phi^{(1)}_p = \frac{1}{4\pi k_o} \frac{q_{p'}}{r(p,p')} = \frac{1}{4\pi k_o} \frac{q_{p'}}{\sqrt{(x-x')^2 + (y-y')^2 + (z-z')^2}}$$

due to the charge $q_{p'}$ located at the point p', as in Figure (2.1) has a first order singularity at the point p' corresponding to $r = 0$. Singularities of higher order can be generated by superimposing onto this potential a potential corresponding to a similar charge but of opposite sign, displaced a distance $\Delta x'$ from it. This process is equivalent to differentiating eq. (2.1) with respect to x'. If we denote

differentiation with respect to x' by $\frac{\partial}{\partial x'}$ and call $\vec{\nabla}'$ the operator corresponding to $\vec{\nabla}$ but referred to the source co-ordinates we see that in general for any function of the relative co-ordinates only

$$\frac{\partial}{\partial x'}\left\{ f\left[(x-x'),(y-y'),(z-z')\right]\right\} = -\frac{\partial}{\partial x}\left\{ f\left[(x-x'),(y-y'),(z-z')\right]\right\}$$

and in general

(2.2) $$\vec{\nabla}' f = -\vec{\nabla} f$$

Since the derivative of a solution of Laplace's equation is also a solution, the process of differentiation with respect to the source point as physically described above will generate new solutions with successive higher order singularities near r = 0. Such potentials are called multipole potentials.

For a single differentiation we obtain for instance:

(2.3) $$\phi^{(2)} = \frac{\partial \phi^{(1)}}{\partial x'}\Delta x' = \frac{q\Delta x'\,(x-x')}{4\pi k_o r^3} = (q\Delta x')\frac{\cos\theta}{4\pi k_o r^2}$$ (See Figure 2.1)

and if we let:

(2.4) $$q\vec{\Delta}_{x'} = \vec{p}^{(2)}$$

be the dipole moment of the distribution (positive from - to + charge) we can write this as:

(2.5) $$\phi^{(2)} = \frac{1}{4\pi k_o}\vec{p}^{(2)}\cdot\vec{\nabla}'\left(\frac{1}{r}\right) = -\frac{1}{4\pi k_o}\vec{p}^{(2)}\cdot\vec{\nabla}\left(\frac{1}{r}\right) = \frac{1}{4\pi k_o}\frac{\vec{p}^{(2)}\cdot\vec{r}}{r^3}$$

Note the sign conventions regarding $\vec{r}$ and $\vec{\nabla}$ discussed in Chapter 1. This solution is the dipole potential.

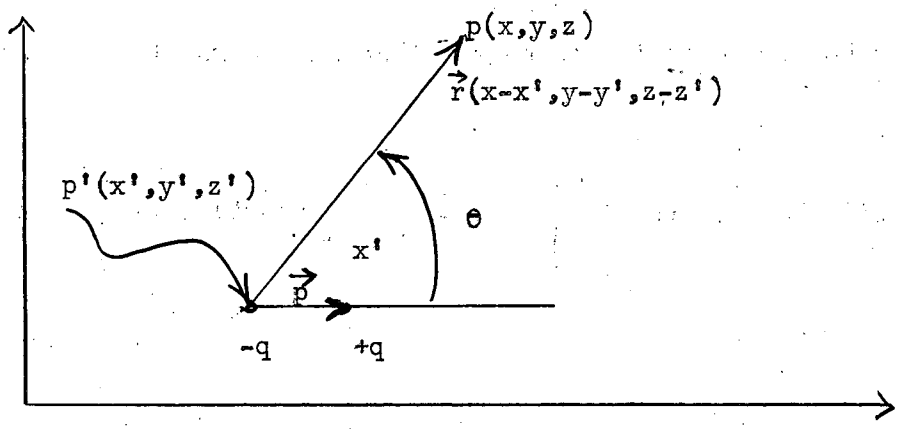


Figure (2.1)

The potential distribution and consequently the fields of higher multiples of the charge, or multipoles, can be generated by the same method of geometrical construction. For example the potential field of 2^{n+1} pole is generated by taking the potential field of a 2^n pole and subtracting from it the potential field of another 2^n pole that is displaced infinitesimally in an arbitrary direction (or superposing the potential of the displaced 2^n pole with opposite sign).

The potential of a 2^n pole will therefore have the general form:

$$(2.6) \qquad \phi^{(2^n)} = \frac{p^{(n)}}{4\pi k_o n!} \frac{\partial^n}{\partial x' \partial y' \cdots} \left(\frac{1}{r}\right) = \frac{(-)^n p^{(n)}}{4\pi k_o n!} \frac{\partial^n}{\partial x \partial y \cdots} \left(\frac{1}{r}\right)$$

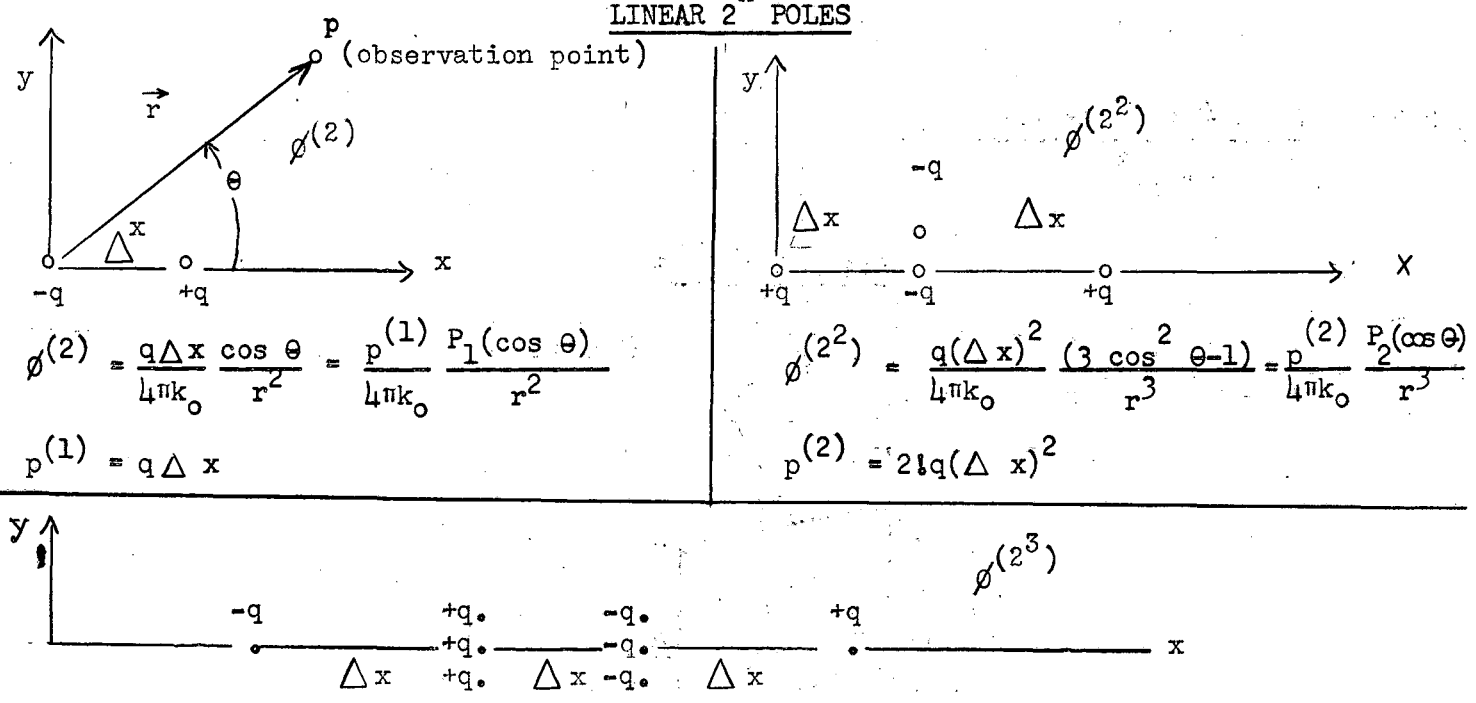
where $p^{(n)}$ is the multipole moment, defined by the recurrence relation $p^{(n)} = np^{(n-1)} \Delta x'_n$ where $\Delta x'_n$ is the displacement leading to the 2^n - pole. In the special case in which all the displacements are in one direction, we have a linear 2^n pole:

$$(2.7) \qquad \phi^{(2^n)}(x,y,z) = \frac{p^{(n)}}{4\pi k_o n!} \frac{\partial^n}{\partial x'^n} \left(\frac{1}{r}\right) = \frac{p^{(n)}}{4\pi k_o} \frac{P_n(\cos\theta)}{r^{n+1}}$$

where $P_n(\cos\theta)$ are the Legendre polynomials.

A few examples of such multipoles are shown in Figure (2.2). An arbitrary linear combination of higher multipoles constitutes a general Taylor-Laurent expansion of the potential. For that reason one can state the theorem that at a large distance from a charge distribution the potential of this charge distribution can be expressed as an infinite series of multipole potentials. By a large distance we mean a distance that is large compared to any dimension of the charge distribution. We shall see later that the multipole expansion is actually equivalent to an expansion in Legendre polynomials, and that the various orders of the Legendre polynomials actually correspond to multipoles of various orders.

LINEAR 2^n POLES



$$\phi^{(2)} = \frac{q\Delta x}{4\pi k_o}\frac{\cos\theta}{r^2} = \frac{p^{(1)}}{4\pi k_o}\frac{P_1(\cos\theta)}{r^2}$$

$$p^{(1)} = q\Delta x$$

$$\phi^{(2^2)} = \frac{q(\Delta x)^2}{4\pi k_o}\frac{(3\cos^2\theta - 1)}{r^3} = \frac{p^{(2)}}{4\pi k_o}\frac{P_2(\cos\theta)}{r^3}$$

$$p^{(2)} = 2!q(\Delta x)^2$$

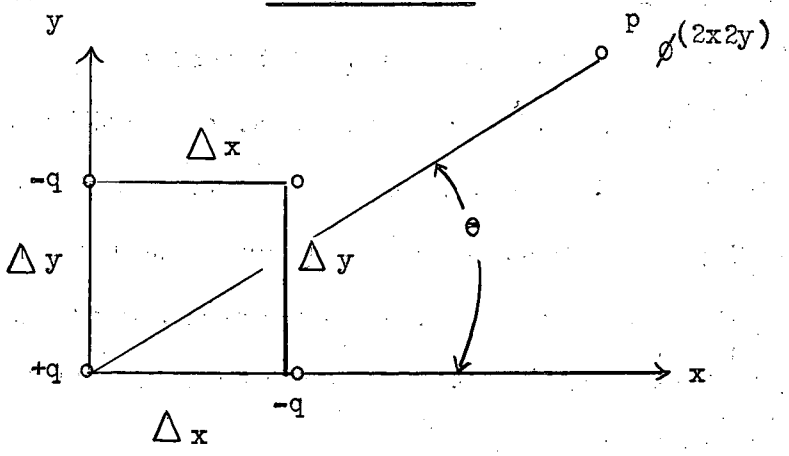
$$\phi^{(2^3)} = \frac{q(\Delta x)^3}{4\pi k_o}\frac{15\cos^3\theta - 9\cos\theta}{r^4} = \frac{p^{(3)}P_3(\cos\theta)}{4\pi k_o r^4}$$

$$p^{(3)} = 3!q(\Delta x)^3$$

FIGURE (2.2)

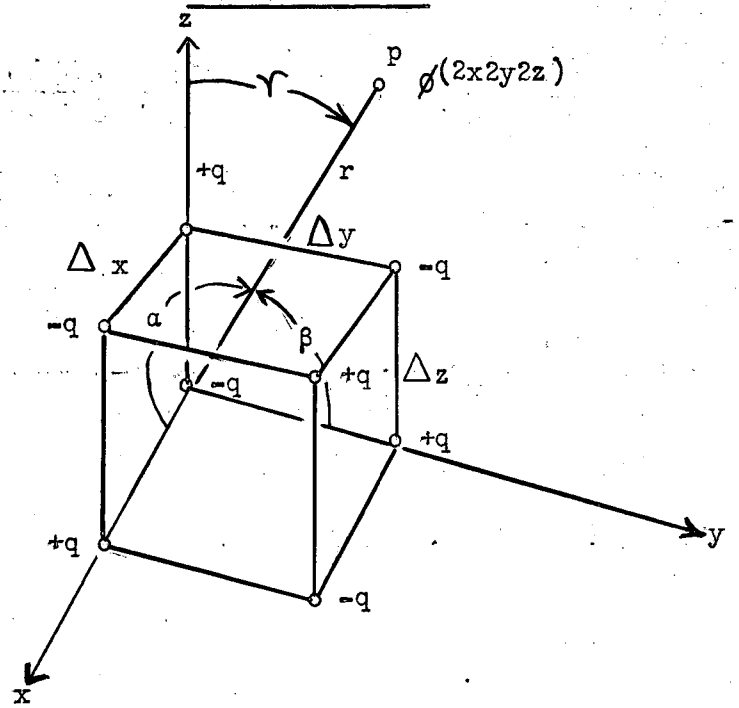
FIGURE (2.2) (CONTINUED)

PLANE 2^n POLE



$$\phi^{(2x2y)} = \frac{3q\Delta x\,\Delta y \cos\theta \sin\theta}{4\pi k_o r^3}$$

SPATIAL 2^n POLE



$$\phi^{(2x2y2z)} = \frac{15q(\Delta x)(\Delta y)(\Delta z)\cos\alpha\cos\beta\cos\gamma}{4\pi k_o r^4}$$

The energy of a dipole in an electrostatic field is given by:

(2.8) $U = -\vec{p} \cdot \vec{E}$ — U = the energy of position of an already created dipole in an electrostatic field. (Joules)

which follows directly from the torque relation:

(2.9) $\vec{L} = \vec{p} \times \vec{E}$ — $\vec{L}$ = the torque exerted in such a direction as to rotate the positive end of the dipole toward the field direction. (Newton Meters)

of a dipole. The force acting on a dipole, when the relative orientation of the dipole and the field is not free to change, in an inhomogenous electric field is given by:

(2.10)* $\vec{F} = -\vec{\nabla} U = -\vec{\nabla}(-\vec{p} \cdot \vec{E}) = (\vec{p} \cdot \vec{\nabla})\vec{E}$

Note that this force vanishes for constant fields, as would be expected from symmetry arguments.

Let us now consider the interaction force and energy between two dipoles, such as those shown in Figure (2.3), whose moment vectors are oriented at an arbitrary angle in space to each other. Combining the force equation (2.10),

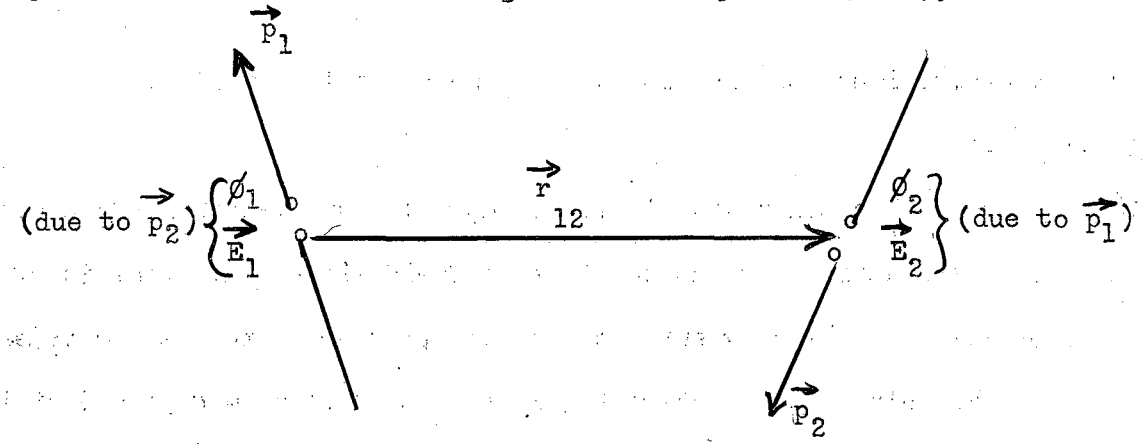


FIGURE (2.3)

the potential equation (2.5) and the field equation (1.11), we have for the force $\vec{F}_1$

---

*Since: $\vec{\nabla}(\vec{p} \cdot \vec{E}) = \vec{p}\times(\vec{\nabla} \times \vec{E}) + (\vec{p} \cdot \vec{\nabla})\vec{E} = (\vec{p} \cdot \vec{\nabla})\vec{E}$ if $\vec{\nabla} \times \vec{E} = 0$ If $\vec{\nabla} \times \vec{E} \neq 0$ an additional term is obtained.

on dipole 1 in the field of dipole 2, or conversely the force $\vec{F}_2$:

$$(2.11)\qquad \phi_1 = -\frac{1}{4\pi k_o}\cdot\vec{p}_2\cdot\vec{\nabla}_1\left(\frac{1}{r}\right),\quad \vec{E}_1 = -\vec{\nabla}_1\phi_1 = +\frac{1}{4\pi k_o}\vec{\nabla}_1\left\{\vec{p}_2\cdot\vec{\nabla}_1\left(\frac{1}{r}\right)\right\}$$

$$\vec{F}_1 = (\vec{p}_1\cdot\vec{\nabla}_1)\vec{E}_1 = \frac{1}{4\pi k_o}(\vec{p}_1\cdot\vec{\nabla}_1)\left[\vec{\nabla}_1\left\{\vec{p}_2\cdot\vec{\nabla}_1\left(\frac{1}{r}\right)\right\}\right]$$

$$\vec{F}_2 = \frac{1}{4\pi k_o}(\vec{p}_2\cdot\vec{\nabla}_2)\left[\vec{\nabla}_2\left\{\vec{p}_1\cdot\vec{\nabla}_2\left(\frac{1}{r}\right)\right\}\right]$$

We can get the interaction energy between the two dipoles of Figure (2.3) by combining Equations (2.5), (2.11) and (2.8). We have for the energy U_{12} of the dipole 1 in the field of dipole 2, and conversely for U_{21}:

$$(2.12)\qquad U_{12} = -\vec{p}_1\cdot\vec{E}_1 = -\frac{(\vec{p}_1\cdot\vec{\nabla}_1)}{4\pi k_o}\left[\vec{p}_2\cdot\vec{\nabla}_1\left(\frac{1}{r}\right)\right]$$

$$U_{12} = +\frac{1}{4\pi k_o}\left[\frac{\vec{p}_1\cdot\vec{p}_2}{r^3} - \frac{3}{r^5}(\vec{p}_1\cdot\vec{r})(\vec{p}_2\cdot\vec{r})\right]$$

$$U_{12} = U_{21}$$

These are general expressions for the interaction energy of two dipoles.

Surface Singularities

We have considered point singularities and will now discuss surface singularities. A surface singularity is another type of field singularity that is often of interest in electrostatics. Usually surface singularities up to the second order or dipole form are the only ones of interest. Surface charges or monopole distributions will be discussed later. Let $\vec{\tau}$ be the dipole moment per unit area of a double layer surface charge arrangement. The potential arising from such a distribution is given by:

$$(2.13)\qquad \phi_p = \frac{1}{4\pi k_o}\iint\frac{\vec{\tau}\cdot\vec{r}}{r^3}\,dS$$

This expression reduces, in the case when $\vec{\tau}$ is uniform and normal to the surface over the dipole sheet,* when simplified by (1.5), to the relation:

$$(2.14)\quad |\phi_p| = \frac{|\tau|}{4\pi k_o} \iint \frac{\vec{r}\cdot d\vec{S}}{r^3} = \frac{|\tau|}{4\pi k_o}\,\Omega$$

Here Ω is the solid angle subtended by the boundaries of the dipole sheet at the point of observation as in Figure (2.4). The solid angle subtended by a non-closed surface jumps discontinuously by 4π as the point of observation crosses the dipole sheet. This means that in the ideal case of an infinitely thin dipole charge layer, the potential function will have a discontinuity of magnitude $|\tau|/k_o$, but will have a continuous derivative at the dipole sheet.

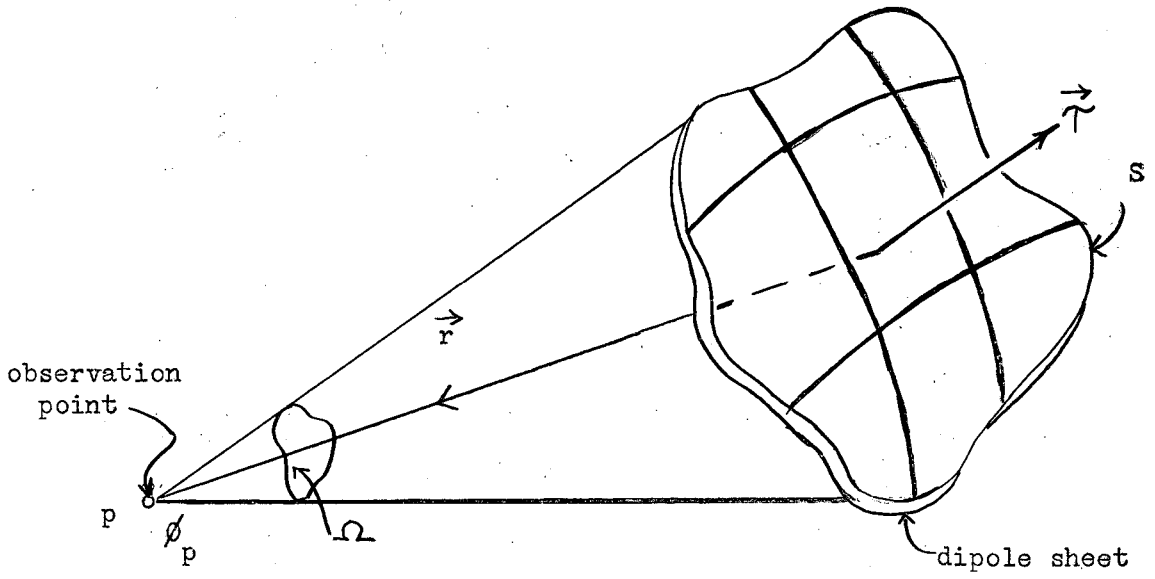


Figure (2.4)

On the other hand, a simple surface charge layer will not result in a discontinuity in potential, but will produce a discontinuity only in the normal derivative or the potential, the magnitude or discontinuity being σ/k_o where σ is the surface charge density of the charge layer. A comparison between the two cases is shown in Figure (2.5A and B). Since surface charge layers and dipole charge layers enable us to introduce arbitrary discontinuities in the potential and its derivatives at a particular surface, we can make

* No generality is lost since a dipole layer with $\vec{\tau}$ not parallel to $d\vec{S}$ is equivalent to a dipole layer with $\vec{\tau}$ parallel to $d\vec{S}$ plus a charge layer.

the potential vanish outside a given volume by surrounding the volume with a suitably chosen charge layer and dipole layer. This is a further explanation of the significance of the surface terms in Equation (1.24) which was derived from Green's theorem. We see that the surface terms in that expression, when ϕ and $\vec{\nabla}\phi$ are properly evaluated on the surface in terms of a $\vec{\tau}$ and a σ, are precisely those terms that are necessary to cancel the field of the charges inside the surface S in the region outside of S, provided that there are no other sources outside of S,

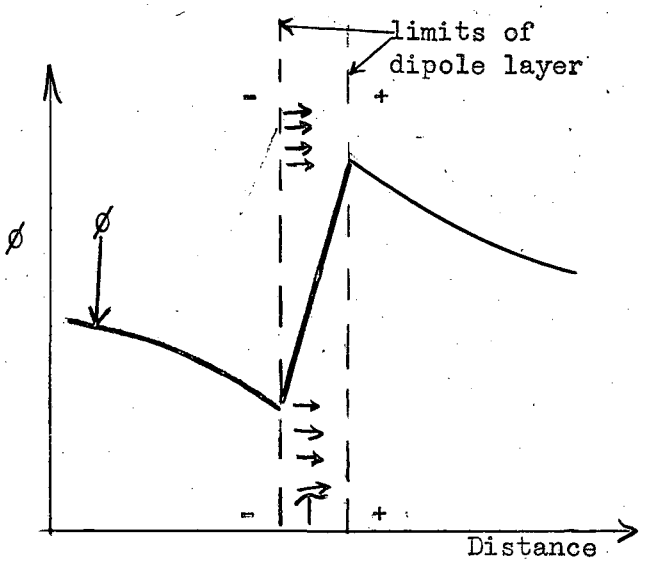


FIGURE (2.5A)

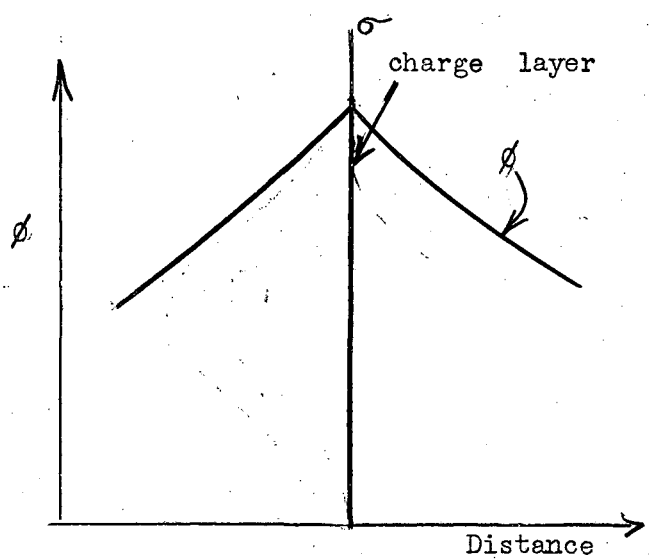


FIGURE (2.5B)

as can be seen by writing eq. (1.24)as:

$$(2.15)\qquad \phi_p = \frac{1}{4\pi k_o}\left\{ \iiint_V \frac{\rho\, dv}{r} + \iint_S \tau \frac{\vec{r}\cdot d\vec{S}}{r^3} + \iint_S \sigma \frac{dS}{r} \right\}$$

where $\tau = k_o\phi$

$$\sigma = k_o \frac{\partial \phi}{\partial n}$$

As an example of a combined surface charge and dipole layer distribution that will just cancel the field outside a given surface, yet leave the field inside the surface unchanged, consider a point charge q located at the point R = 0, and the surface R = a surrounding this charge. If we place a surface charge density $\sigma = -q/4\pi a^2$ per

unit area on the sphere R = a, it will give a potential:

$$\phi_\sigma = -\frac{q}{4\pi k_o a} \quad R < a$$

$$\phi_\sigma = -\frac{q}{4\pi k_o R} \quad R > a$$

If in addition we place a surface dipole layer of moment $\vec{\tau} = q\vec{R}/4\pi aR$ per unit area on the sphere R = a, it will give a potential:

$$\phi_\tau = \frac{+q}{4\pi k_o a} \quad R < a$$

$$\phi_\tau = 0 \quad R > a$$

The potential of the original charge q is:

$$\phi_o = \frac{+q}{4\pi k_o R}$$

Adding those potentials we obtain:

$$(2.16) \qquad \phi = \phi_o + \phi_\sigma + \phi_\tau = \frac{+q}{4\pi k_o R} \quad R < a, \qquad \phi = 0 \quad R > a$$

The potential due to a dipole layer is therefore double valued at the surface although this is, of course, only strictly true in the limiting case in which the dipole layer has an infinitesimal thickness; see Figure (2.5A). For this reason the case of a double valued potential at a surface does not have a physical reality. This method of generating a non-conservative potential is a useful one in the theory of magnetic fields to be considered later. The electric field produced by a dipole layer can be derived as follows:

The potential change corresponding to a displacement of the point of observation p by a distance $d\vec{x}$ as in Figure (2.6) is:

$$(2.17) \qquad d\phi_p = -\vec{E} \cdot d\vec{x}$$

The change in solid angle, $d\Omega$, subtended by the dipole layer at the field point is the same whether the field point moves $\vec{dx}$ or the layer moves $-\vec{dx}$. The latter case is shown:

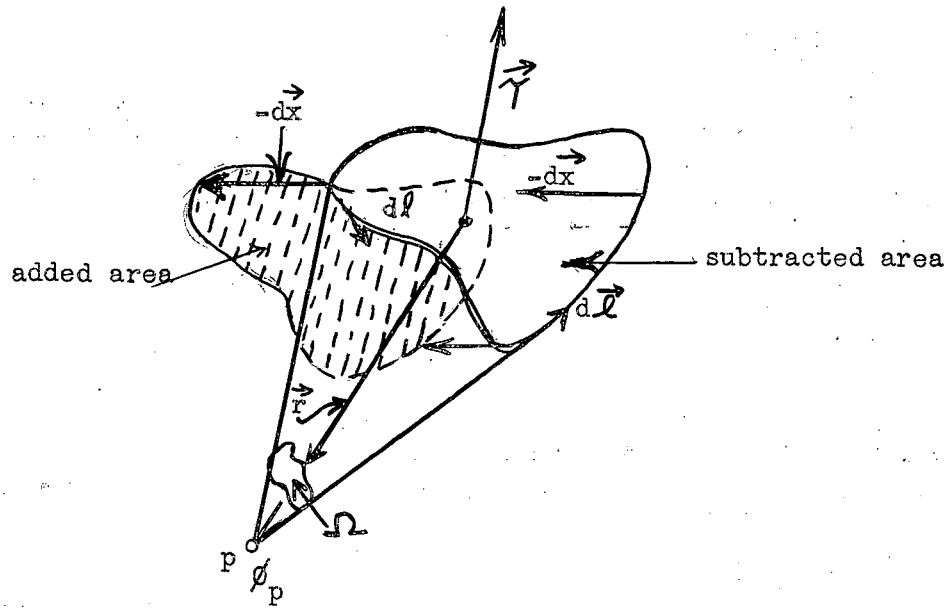


FIGURE (2.6)

The change in solid angle is seen to be:

$$(2.18) \qquad d\Omega = \oint \frac{(\vec{dx} \times d\vec{\ell}) \cdot \vec{r}}{r^3} = \oint \frac{(d\vec{\ell} \times \vec{r}) \cdot \vec{dx}}{r^3}$$

The change in potential corresponding to this solid angle change from (2.14) is:

$$(2.19) \qquad d\phi_p = -\frac{\tau}{4\pi k_o} d\Omega$$

Substituting (2.19) and (2.18) into (2.17) we have:

$$(2.20) \qquad -\vec{E} \cdot \vec{dx} = \frac{-\tau}{4\pi k_o} \oint \frac{(d\vec{\ell} \times \vec{r}) \cdot \vec{dx}}{r^3}$$

We may remove the dot product since $\vec{dx}$ is an arbitrary vector and (2.20) holds for all values of $\vec{dx}$. It would not be permissible to un-dot if this were not true.

$$(2.21) \qquad \vec{E} = \frac{\tau}{4\pi k_o} \oint \frac{d\vec{\ell} \times \vec{r}}{r^3} = \frac{-\tau}{4\pi k_o} \oint d\vec{\ell} \times \vec{\nabla}\left(\frac{1}{r}\right)$$

These expressions will again be useful in the derivation of the magnetic fields due to currents where the corresponding potential does have a multivalued behavior completely analogous to the properties of the surface dipole moment.

Volume Singularities

We shall now consider volume distributions of dipole moment. The potential due to the volume distribution is given by considering the point dipole moment of Equation (2.5) as a volume density and integrating over the volume:

(2.22) $$\phi_p = \frac{1}{4\pi k_o} \iiint \vec{P} \cdot \vec{\nabla}' \left(\frac{1}{r}\right) dv$$ $\vec{P}$ = the dipole moment per unit volume.

This can be transformed to a form which is physically more interesting by the relation:

(2.23) $$\vec{\nabla}' \cdot \left(\frac{\vec{P}}{r}\right) = \frac{1}{r} \vec{\nabla}' \cdot \vec{P} + \vec{P} \cdot \vec{\nabla}' \left(\frac{1}{r}\right)$$

Substituting the last term of (2.23) into (2.22) and using Gauss' divergence theorem (1.7) we have:

(2.24) $$\phi_p = \frac{1}{4\pi k_o} \left[\iiint \vec{\nabla}' \cdot \left(\frac{\vec{P}}{r}\right) dv - \iiint \frac{1}{r} \vec{\nabla}' \cdot \vec{P} \, dv \right] = \frac{1}{4\pi k_o} \left[\iint \frac{\vec{P} \cdot d\vec{S}}{r} - \iiint \frac{\vec{\nabla}' \cdot \vec{P}}{r} dv \right]$$

This expression can be interpreted as follows: The first term, which is a surface integral, is a potential equivalent to that of a surface charge density, while the second term is a potential equivalent to that of a volume charge density. The charge densities which have potentials equivalent to those caused by the volume polarization of a region of space are given by the expressions:

(2.25)* $$\sigma_p = P_n \qquad \rho_p = -\vec{\nabla} \cdot \vec{P}$$ σ_p = the polarization surface charge

ρ_p = the polarization volume charge

These relations can be derived from purely geometrical considerations. For example in case we have an inhomogeneous dipole moment per unit volume, ρ_p will represent the charge density which accumulates from incomplete cancellation of the ends of the individual dipoles distributed in the volume. σ_p, on the other hand, represents the polarization charge density on the surface produced by the lack

* Since $\rho_p = -\vec{\nabla} \cdot \vec{P}$ is a field equation the prime on $\vec{\nabla}$ can be dropped without ambiguity. The prime on the $\vec{\nabla}$ is only necessary in integral expressions, which relate a field quantity to an integral over a source quantity.

of neighbors for the dipoles which lie with their ends on the surface. It is evident that ρ_p will vanish in a homogeneous polarized medium. In fact, it will vanish provided only that the dipole moment per unit volume has a zero divergence. We will thus have for the potential due to the two forms of polarization charges, when we substitute (2.25) into (2.24):

$$(2.26) \qquad \phi_p = \frac{1}{4\pi k_o} \left[\iint \frac{\sigma_p dS}{r} + \iiint \frac{\rho_p dv}{r} \right]$$

CHAPTER 3 BOUNDARY CONDITIONS AND RELATION OF MICROSCOPIC TO MACROSCOPIC FIELDS

The dipole moments per unit volume considered in Chapter 2 are special examples of sources which give rise to electrostatic fields and can therefore be treated as special types of charge densities in Poisson's Equation (1.16). Since material media in an electric field give rise to such volume dipole distributions, the behavior of such media in an electric field can be described in terms of its polarization, that is its dipole moment per unit volume. It is customary, in order to clarify the understanding of polarization, to separate the total charge that produces an electrostatic field into two parts, namely a true, free, moveable, net charge ρ, and a bound, zero-net, polarization charge ρ_p. This division is to a certain extend arbitrary, in the sense that the polarization charge ρ_p simply represents separated charges which on the scale of observation being considered in a particular experiment are essentially inaccessible, but if a smaller scale were considered would be considered separately, as free charges. If, for example, we place a piece of metal between the plates of a condenser, we can describe the resultant field between the plates either in terms of the true charges produced on the metal, or in terms of an equivalent polarization of the piece of metal, depending on whether we consider the charges individually measurable or not, respectively. If, instead of the metal we introduce a piece of dielectric between the condenser plates, we are forced to describe the phenomena by a polarization charge, rather than by a true charge, since it is assumed in the theory that observation shall not be made on an atomic scale. An atomic scale observation would be necessary in order to "resolve" the volume polarization into individual charges.

It is thus seen that the distinction between ρ and ρ_p is an arbitrary one. This arbitrariness will in no way distrub the formalism used to describe the fields produced by such polarization charges. Since we have divided the sources of electric

fields into sources of the two types mentioned, the Poisson source Equation (1.16) becomes:

$$(3.1) \qquad \nabla^2\varphi = -\vec{\nabla}\cdot\vec{E} = \frac{-1}{k_o}(\rho + \rho_p)$$

Note that the symbol ρ denotes only the true free charge, at the point where the divergence is being taken. If we express ρ_p in terms of the divergence of the polarization $\vec{P}$, as given by Equation (2.25) we obtain from Equation (3.1):

$$(3.2) \qquad \vec{\nabla}\cdot\left(\vec{E} + \frac{\vec{P}}{k_o}\right) = \frac{\rho}{k_o}$$

If we define the displacement vector $\vec{D}$ by:

$$(3.3) \qquad \vec{D} = k_o\vec{E} + \vec{P}$$

$\vec{D}$ = the electric displacement (Coulomb·Meter^{-2})

the source equation becomes simply:

$$(3.4) \qquad \vec{\nabla}\cdot\vec{D} = \rho$$

and also:

$$(3.5) \qquad \vec{\nabla}\cdot\vec{E} = \frac{\rho_t}{k_o} \qquad \rho_t = \rho + \rho_p$$

with the corresponding integral relations secured from Gauss' divergence theorem (1.7) by integrating over the volume containing all the charges:

$$(3.6) \qquad \iint \vec{D}\cdot d\vec{S} = q$$

$$(3.7) \qquad \iint \vec{E}\cdot d\vec{S} = \frac{q_t}{k_o}$$

$q_t = q + q_p$ = the total charge

$$q_p = \iiint -\left(\vec{\nabla}'\cdot\vec{P}\right)dv$$

$\vec{D}$ thus represents a partial electric field, namely that electric field whose sources are only the true charges. Note that the relation (3.3) between $\vec{D}$ and $\vec{E}$ is basically an additive one, the difference between $\vec{D}$ and $k_o\vec{E}$ being the polarization $\vec{P}$. Note also that the polarization, although defined in a purely geometrical fashion as the dipole moment per unit volume, has the properties of an electric field. The polarization field $\vec{P}$ is that field whose flux arises only from the polarization charges ρ_p.

The solution of an actual field problem involving polarized bodies will depend on the manner in which the polarization depends on the external field. In most cases, the polarization will be proportional to the electric field. This can be expressed by an equation of the type:

(3.8) $\vec{P} = k_o \chi \vec{E}$ χ = the electric susceptibility

Such a description excludes the consideration of electrets.* But electrets do not have great practical importance. In case we do have a simple medium whose polarization depends linearly on the imposed electric field, as expressed by Equation (3.8), then all three vectors, $\vec{D}$, $\vec{E}$ and $\vec{P}$, will be related by constants of proportionality as given by:

*Electrets are materials possessing a permanent electric dipole moment.

(3.9) $\vec{D} = k_o\vec{E} + \vec{P} = k_o(1 + \chi)\vec{E}$

If we let:

(3.10) $k = 1 + \chi$ k = the specific inductive capacity

then:

(3.11) $\vec{D} = kk_o\vec{E}$

and:

(3.12) $\vec{P} = k_o(k-1)\vec{E}$

Equation (3.8) pre-supposes that the medium polarizes isotropically, or that the polarization properties of the medium do not depend on the direction of the polarization. In general this is not true, and in fact it holds only in liquids, gases, amorphous solids and cubic crystals. In crystals of lower symmetry than cubic the relation between each of the components of the polarization vector and of the electric field vector is still linear but the constants of proportionality in the various

directions may be different. This means that the relation between the components of the polarization vector and the components of the electric field vector are given by a tensor relation of the type:

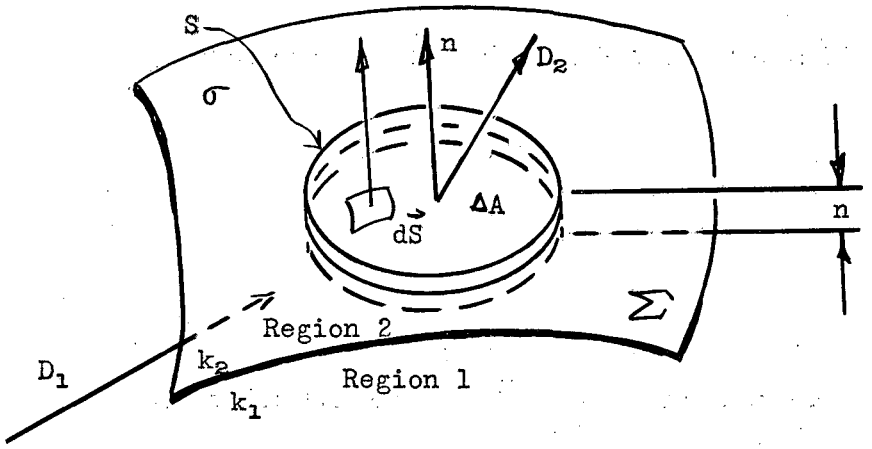
$$(3.13) \qquad P_i = k_o \sum_j \chi_{ij} \; E_j$$

The tensor χ_{ij} can be shown by energy considerations to be a symmetrical tensor, and therefore it can be expressed in terms of principal coordinates by a set of only three constants.

The case of non-linear media, that is media where $\vec{E}$ and $\vec{P}$ are not proportional and Equation (3.8) does not hold, will not be treated here, but the parallel case will be treated in the discussion of magnetic media where non-linearity is of more practical importance. It should be pointed out, however, that the relation that has been given between $\vec{P}$ and $\vec{E}$ is not a fundamental one, but is only a special simplifying assumption.

Maxwell's field equations, to be discussed later, are a set of equations whose sources are divided into accessible and inaccessible charges. In order to obtain a solution of Maxwell's equations the inaccessible charges must be related to the accessible charges or the fields produced by the accessible charges by additional relationships. The relations which evaluate the inaccessible charge sources in terms of the external fields which produce them are called the constitutive equations. Equation (3.8) is an example of such an equation. The constitutive equations are, of course, not basic, and depend on the properties of the material in which the inaccessible charges arise. None of the equations (3.8) to (3.13) are dependent upon the homogeneity of the medium, only upon its linearity. That is the susceptibility and specific inductive capacity may be arbitrary functions of the coordinates. A case of much interest is the one where the specific inductive capacity varies discontinuously, as at the boundary between two dielectrics. In this case, illustrated in figures (3.1 and 3.2), we imagine a small volume, whose dimension normal to the

interface is smaller than its dimensions parallel to the interface by an order of magnitude, to be so placed that one of its large surfaces lies in medium 1 and the other in medium 2 and both are parallel to the interface. In addition we imagine a small loop erected as in Figure (3.2) with its major extent lying parallel to the surface, one side in medium 1 and the other in medium 2. The little volume in Figure (3.1) will be used to derive the behavior of the normal components of the fields and the loop in Figure (3.2) will be used to derive the behavior of the tangential components of the fields as they cross the interface.



$\vec{n} = \frac{d\vec{S}}{dS}$ = the unit vector normal to ΔA, the top of surface S.

σ = surface charge on the surface Σ.

Figure (3.1)

We have from (3.6) for the surface integral of $\vec{D}$ over the little volume in Figure (3.1):

$$(3.14) \quad \iint \vec{D} \cdot d\vec{S} = q$$

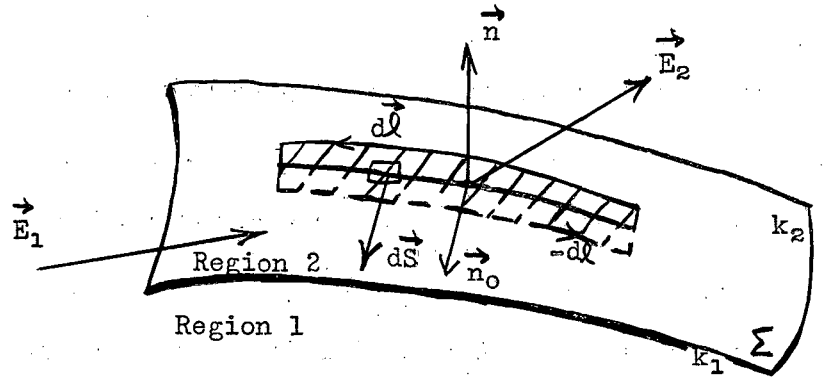
As we shrink the dimension h to zero, $q \to \sigma \Delta A$, and the contribution of the sides of the volume to the surface integral vanishes, so (3.14) becomes:

$$(3.15) \quad \vec{n} \cdot (\vec{D}_2 - \vec{D}_1) = \sigma$$

And if we may make the assumption of (3.11) we have:

$$(3.16) \quad \vec{n} \cdot (k_2\vec{E}_2 - k_1\vec{E}_1) = \frac{\sigma}{k_o}, \quad \vec{n} \cdot (k_2 \vec{\nabla}\phi_2 - k_1 \vec{\nabla}\phi_1) = -\frac{\sigma}{k_o}$$

We have made the assumption that ΔA is small enough so that the fields are essentially constant over it.



$\vec{n}$ = unit normal to the surface Σ.
$\vec{n}_o$ = unit normal to the loop of integration, lying in the surface Σ.
$d\vec{S}$ = differential of surface area of the integration loop.

Figure (3.2)

From (1.15) we have for the line integral of the electric field:

(1.15) $\oint \vec{E} \cdot d\vec{\ell} = 0$

If we apply this to the path shown in Figure (3.2) and let the ends of the path shrink to zero, we have:

(3.17) $\vec{E}_2 \cdot d\vec{\ell} - \vec{E}_1 \cdot d\vec{\ell} = 0$

Noting that $\frac{d\vec{\ell}}{d\ell} = \vec{n}_o \times \vec{n}$ we have:

(3.18) $\vec{n}_o \cdot \vec{n} \times (\vec{E}_2 - \vec{E}_1) = 0$

And since $\vec{n}_o$ can be oriented in any direction relative to $\vec{E}$, and (3.18) holds for all of them, we may remove the dot product:

(3.19) $\vec{n} \times (\vec{E}_2 - \vec{E}_1) = 0$

And hence from (1.11) we have:

(3.20) $\vec{n} \times (\vec{\nabla}\varphi_2 - \vec{\nabla}\varphi_1) = 0$

We have assumed that $\int d\vec{\ell}$ is short enough so that the fields are essentially constant over its length.

The relation (3.20) could have been secured from (1.10) and an application of Stokes' theorem, but (1.10) is actually not necessary in this proof. The relationship (3.20) will hold true even in the non-static case where $\vec{\nabla} \times \vec{E}$ is not zero but is given by the Faraday law of induction:

(3.21) $\vec{\nabla} \times \vec{E} = -\frac{\partial \vec{B}}{\partial t}$ $\vec{B}$ = the magnetic induction (Weber $\cdot$ Meter^{-2})

Applying Stokes' theorem (1.14) to (3.21) and carrying the integration over the path shown in Figure (3.2) and neglecting the contributions of the ends of the path

we have:

$$(3.22)\quad \oint \vec{E} \cdot d\vec{\ell} = \int \vec{E}_2 \cdot d\vec{\ell} - \int \vec{E}_1 \cdot d\vec{\ell} = -\iint \frac{\partial \vec{B}}{\partial t} \cdot d\vec{S}$$

As we shrink the integration path to zero the differential of surface area $d\vec{S}$ vanishes to a higher order than the differentials of path length $d\vec{\ell}$ and thus the term containing $\vec{B}$ drops out and we have (3.19) or (3.20) again. Thus these equations are correct expressions of the boundary conditions, even in the general non-static case.

If, as we saw in Chapter 2, there is no surface dipole layer, the potential is continuous across the boundary and we have:

$$(3.23)\quad \varphi_1 = \varphi_2$$

We have therefore boundary conditions on the potentials and the fields on the two sides of an interface as given by Equations (3.15), (3.16), (3.19) and (3.20). These boundary conditions provide for the continuity of the field across an interface.

We first defined the electrostatic field in a vacuum, caused by free charges and then we introduced material media containing charges that are inaccessible to measurement. The behavior of these media has been described in terms of their dipole moment per unit volume. Certain difficulties arise in the definition of the electric field $\vec{E}$ in material media if one attempts to use a strictly phenomenological point of view. A definition of the field might be made by one of the following three methods, which will not necessarily be in agreement.

(A) We may define the field $\vec{E}$ on an atomic electron scale where the question of the polarizability of material media would presumably not arise. Then for our macroscopic definition of the field we would take the space time average of these atomic fields.

(B) We might consider that a hole be cut into the dielectric material in the field and define the field as that measured in this hole in terms of the force on a unit charge as was used in the vacuum definition.

This cavity definition of the field will make the field strength depend on the geometry of the cavity and on its orientation relative to the direction of the field in the medium. This will lead to a unique definition only if the shape and orientation of the cavity are standardized in an arbitrary manner.

(C) We may define the field as that acting on an individual molecule of the dielectric.

(A) Space Time Average Definition.

Consider a function $f(x,y,z;t)$ defined in a certain region of space during a certain time interval as in Figure (3.3). The space-time average of $f(x,y,z;t)$ when the function is averaged over a time interval 2T and a region of space of radius a is then given by:

$$(3.24)\quad \overline{f(x,y,z;t)} = \frac{1}{2T \times \frac{4\pi a^3}{3}} \int_{-T}^{T} \underbrace{\iiint}_{(\xi^2+\eta^2+\zeta^2 \leq a^2)} f\left[(x+\xi),(y+\eta),(z+\zeta);(t+\theta)\right] d\xi\, d\eta\, d\zeta\, d\theta$$

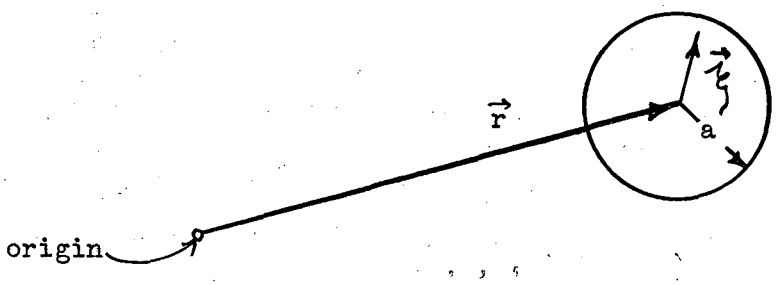


Figure (3.3)

This integral is a linear operation and may therefore be commuted with linear differential operators, as for example:

$$(3.25)\quad \overline{\vec{\nabla} f} = \vec{\nabla}\, \overline{f}$$

On an atomic scale an equation corresponding to (1.9) holds:

$$(3.26)\quad \vec{\nabla} \cdot \vec{\mathcal{E}} = \frac{\rho_t}{k_o}$$

$\vec{\mathcal{E}}$ = the atomic electric field.

ρ_t = the total charge density in the atomic distribution.

Taking the space time average of ρ_t in (3.26) we obtain:

$$(3.27)\quad \overrightarrow{\nabla \cdot \mathcal{E}} = \frac{\overline{\rho_t}}{k_o}$$

And from (3.25) and (3.5) we have:

$$(3.28)\quad \vec{\nabla} \cdot \overrightarrow{\mathcal{E}} = \vec{\nabla} \cdot \vec{E}$$

Hence the macroscopic field $\vec{E}$ is actually the space-time average of the atomic field $\overrightarrow{\mathcal{E}}$, even in the presence of dielectrics.

(B) Empirical Measurement of the Field in a Hole in the Medium Definition.

Consider the three shapes of cavities shown in Figure (3.4). From the boundary conditions of Equations (3.16) and (3.19), the field measured in the slot, whose boundaries are uncharged, in Figure (3.4A), whose major extent is oriented normal to the field is $k\vec{E}_m$, where $\vec{E}_m$ is the field in the medium. The field in the slot in Figure (3.4B) whose major extent is oriented parallel to the field is $\vec{n} \times \vec{E}_m = \vec{E}_m$. The field measured in the spherical cavity in Figure (3.4C), can be shown to be:

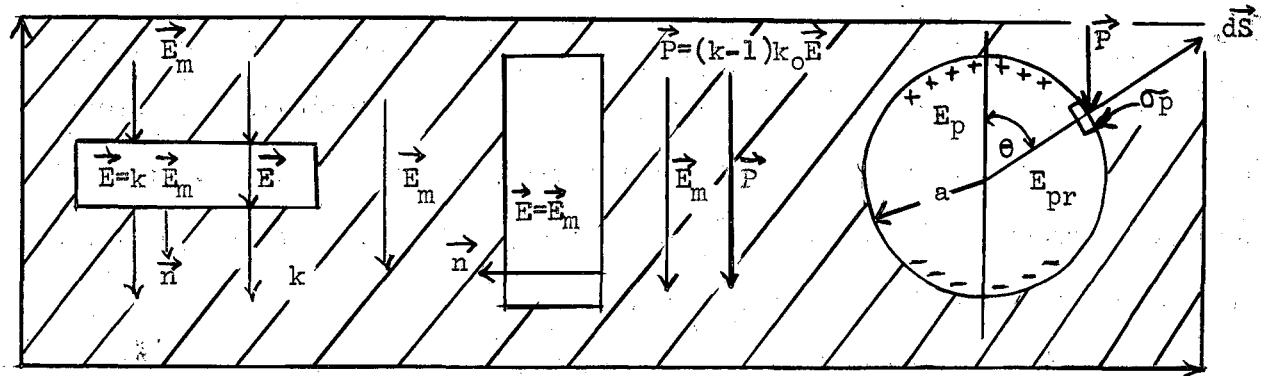


Figure (3.4A) Figure (3.4B) Figure (3.4C)

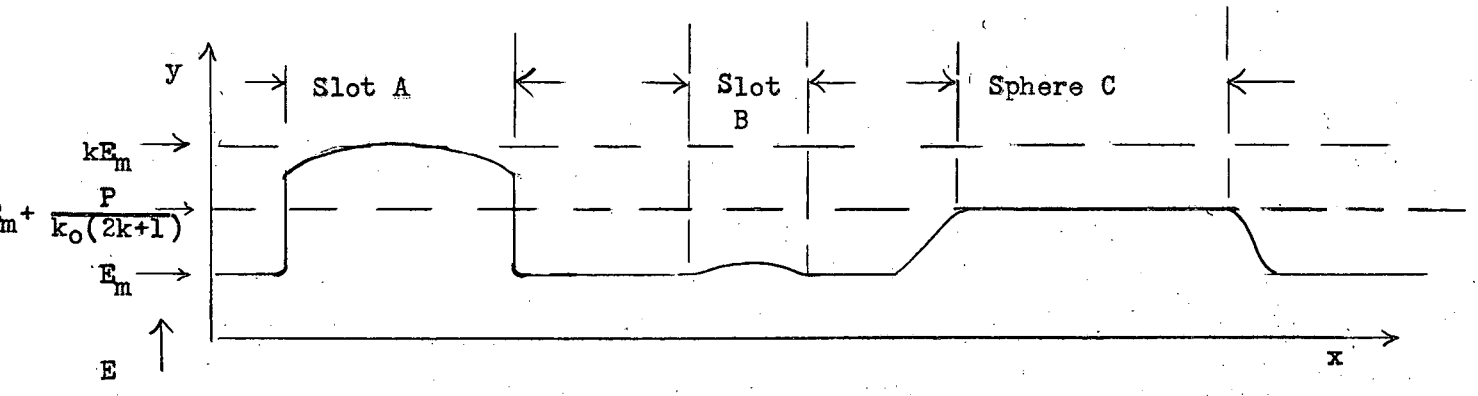
$$(3.29)\quad \vec{E} = \frac{3k\vec{E}_m}{2k+1} = \vec{E}_m + \frac{\vec{P}}{k_o}\,\frac{1}{(2k+1)}$$

by methods to be discussed in Chapter 6 for the solution of boundary value problems. For large values of k, (3.29) becomes:

$$(3.30)\quad \vec{E} = \frac{3}{2}\,\vec{E}_m$$

The three types of fields existing in the cavities are shown graphically in Figure (3.5). The cavity definitions will therefore each give a definite

value of the field provided that the geometry is standardized.



Field profile measured by the cavity technique on a horizontal line passing through the centers of Figures (3.4A, B and C).

Figure (3.5A) Figure (3.5B) Figure (3.5C)

(C) Molecular Fields

Consider a dielectric placed between the plates of a parallel plate condenser, as shown in Figure (3.6), the dielectric and condenser being sufficiently large in the directions parallel to the plates so that end effects may be neglected. Consider a molecule within this dielectric. Let us draw a sphere of radius a about this particular molecule, where this radius is intended to represent schematically the boundary between the microscopic and macroscopic range of phenomena concerning this molecule. The molecule is thus influenced by the fields arising from the following charges:

(1) The charge on the surfaces of the condenser plates.

(2) The surface charge on the dielectric facing the condenser plates.

(3) The surface charge on the interior of the spherical boundary of radius a.

(4) The charges of the individual molecules, other than the molecule under consideration, contained within the sphere of radius a.

Considering these cases:

(1) The charge on the condenser plates produces a field at the molecule in question equal to:

$$(3.31)\quad \frac{\vec{D}}{k_o} = \vec{E} + \frac{\vec{P}}{k_o}$$

(2) The polarization charge on the surface of the dielectric facing the condenser plates, $\sigma_p = \vec{P} \cdot \vec{n}_1$, produces a field at the molecule:

$$(3.32) \quad -\frac{\vec{P}}{k_o}$$

(3) The polarization charge present on the inside of the sphere produces a field that may be calculated as follows: The electric field at the center of the cavity, $\vec{E}_p$, due to the polarization on the surface of the cavity, is given by:

$$(3.33) \quad E_p = \frac{1}{4\pi k_o} \iint \frac{\sigma_p \cos\theta \, dS}{a^2}$$

The differential element of surface charge is:

$$(3.34) \quad \sigma_p dS = 2\pi a^2 \sin\theta \cos\theta \, d\theta \, |\vec{P}|$$

Substituting (3.34) into (3.33) we have:

$$(3.35) \quad E_p = \frac{|\vec{P}| 2\pi a^2}{4\pi k_o a^2} \int_o^{\pi} \sin\theta \cos^2\theta \, d\theta$$

and integrating we get:

$$(3.36) \quad \vec{E}_p = +\frac{\vec{P}}{3k_o}$$

Note that $\vec{E}_p$ is not the solution of the boundary value problem of a spherical cavity within a dielectric, as (3.29) was, but is the solution of the problem of a spherical cavity within a dielectric if the polarization is considered to be unaffected by the presence of the spherical cavity.

(4) The field due to the individual molecules within the sphere must be obtained by summing over the fields due to the dipoles within the sphere. We can secure the potential of an individual dipole from (2.5):

$$(3.37) \quad \phi = \frac{1}{4\pi k_o} \frac{\vec{p} \cdot \vec{r}}{r^3}$$

The field at a distance r from this dipole is:

$$(3.38)\quad \vec{E} = \frac{-1}{4\pi k_o}\vec{\nabla}\left(\frac{\vec{p}\cdot\vec{r}}{r^3}\right) = \frac{-1}{4\pi k_o}\left(\frac{\vec{p}}{r^3} - \frac{3(\vec{p}\cdot\vec{r})\vec{r}}{r^5}\right)$$

Summing over all the dipoles within the sphere we obtain the spatial average of the x component of the field:

$$(3.39)\quad \overline{E_x} = \frac{-1}{4\pi k_o}\sum\left(\frac{p_x}{r^3} - \frac{3(p_x x^2 + p_y xy + p_z xz)}{r^5}\right)$$

Since we have assumed that the dielectric is isotropic, the x, y and z directions are equivalent and we have:

$$(3.40)\quad \overline{x^2} = \overline{y^2} = \overline{z^2} = \frac{\overline{r^2}}{3}\ ,\ \overline{xy} = \overline{yz} = \overline{zx} = 0$$

Hence the field due to the dipoles within the sphere vanishes.

Thus adding the partial fields of (3.31), (3.32), (3.36) and the zero field of part (4) we have for the field acting on one molecule:

$$(3.41)\quad \vec{E}_{eff} = \overbrace{\vec{E} + \frac{\vec{P}}{k_o}}^{(1)} - \frac{\vec{P}}{k_o} + \frac{\vec{P}}{3k_o} = \vec{E} + \frac{\vec{P}}{3k_o}$$

The magnitude of the field of (3.41), derived for isotropic substances is also valid for cubic lattice crystals, but is not valid for crystals of lower symmetry.*

In order to describe the large scale behavior of a dielectric in terms of the constants of the molecules constituting the dielectric, the specific inductive capacity k must be associated with the polarizability of a single molecule α. This connection may be made by the use of Eqn. (3.41) which gives the field and thus the force acting on a single molecule within the body of a dielectric in terms of the external field. Therefore, if α is defined by the equation:

$$(3.42)\quad \vec{p} = \alpha \vec{E}_{eff}$$

α = the dipole moment induced in a molecule by a field of unit strength.

* Note that these expressions consider only dipole-dipole interactions between neighbors. This will clearly be inaccurate for substances having large oriented molecular groups.

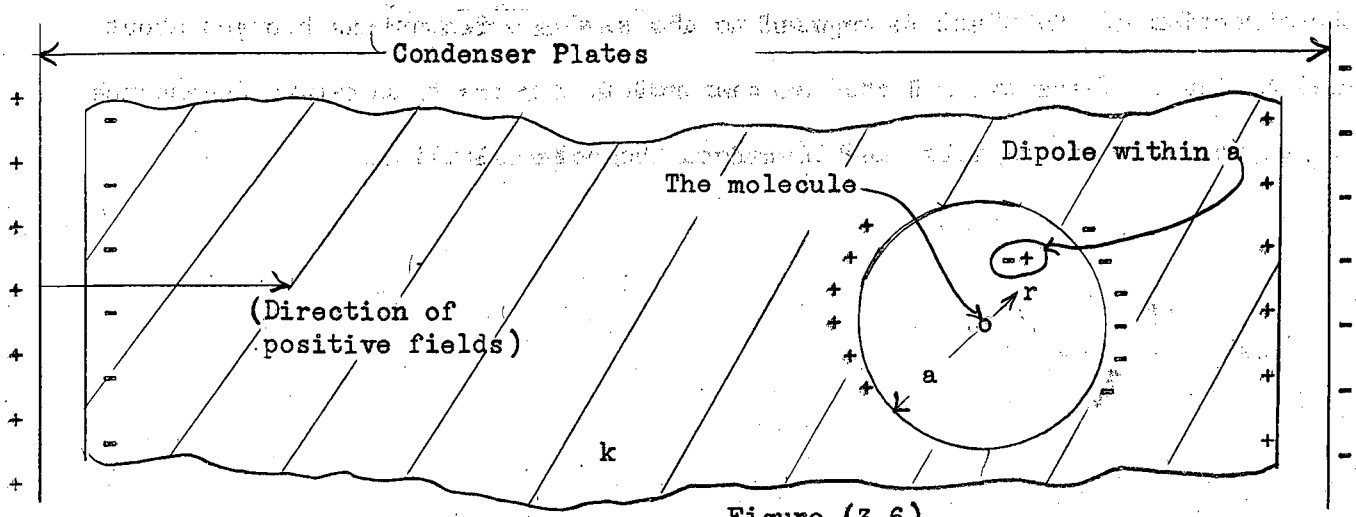


Figure (3.6)

Therefore if α is given for a particular material we may get $\vec{P}$.

(3.43) $\vec{P} = n\vec{p} = n\alpha\vec{E}_{eff}$

$n = \frac{Ng}{M}$

$\vec{P} = \frac{Ng\alpha}{M}\vec{E}_{eff}$

n = the number of molecules per unit volume.
N = Avagadro's Number.
g = the density of the material.
M = the molecular weight of the material.

Combining this with (3.12) and (3.41) we have, after the fields are eliminated:

(3.44) $\frac{k-1}{k+2} = \frac{Ng\alpha}{3Mk_o}$

This formula, known as the Clausius-Mosotti formula, gives the correct dependence of the specific inductive capacity on density for a wide class of solids and liquids. For dilute gases, where $k \cong 1$, Equation (3.44) becomes:

(3.45) $k - 1 \cong \frac{Ng\alpha}{Mk_o}$

which is, of course, to be expected since this relation corresponds to neglecting the interaction between each molecule and its neighbors. The molecular polarizability will in general arise from two basic physical causes: (1) The lengthening of the bonds between atoms, and (2) the preferred orientation of molecules within a fluid

along the direction of the field as opposed to the random orientations brought about by thermal motions. These two effects are responsible for the temperature dependence of the specific inductive capacity and therefore the polarizibility.

CHAPTER 4 GENERAL METHODS FOR THE SOLUTION OF POTENTIAL PROBLEMS

We shall now consider the solution of several types of potential problems. Unfortunately, no general methods of solution are available which will apply to all cases and therefore each individual case demands, to some extent, methods of its own. However, we shall discuss certain methods which apply to general classes of problems and can be discussed as individually characteristic of these classes.

A theorem of great importance for the treatment of potential problems is the uniqueness theorem. This theorem states that if a solution of a potential problem is found within a given physical boundary which gives either, a given potential distribution on that boundary, or which corresponds to a given charge distribution on that boundary, then this solution within this boundary is the only correct solution of the potential equations. This theorem is the justification for attempting any method of solution if the resulting solution can be shown to fulfill Laplace's equation in a charge-free region. Thus no matter how the solution is obtained, if it satisfies these conditions, the problem is considered solved.

The proof of this theorem can easily be given by means of Green's theorem. If we put $\phi\vec{\nabla}\phi$ into Gauss' divergence theorem (1.7) as the vector field, we obtain:

$$(4.1)\quad \iint_S \phi\vec{\nabla}\phi \cdot d\vec{S} = \iiint_v \vec{\nabla}\cdot(\phi\vec{\nabla}\phi)dv = \iiint_v \left[(\vec{\nabla}\phi)^2 + \phi\nabla^2\phi \right] dv$$

The last term vanishes, from Laplace's equation, if we choose our surface of integration in such a way as to exclude all charged regions from the region of integration. It may be necessary to employ surfaces internal to S in order to entirely exclude the charges from S and from v, as was done in Figure (1.2).

Let us suppose that two different potentials, ϕ_1 and ϕ_2 are each a solution of a given potential problem. Both ϕ_1 and ϕ_2 are to satisfy the boundary conditions; hence on the boundary $\phi_1 = \phi_2$ or $\frac{\partial\phi_1}{\partial n} = \frac{\partial\phi_2}{\partial n}$. If we substitute the difference

$\phi_1 = \phi_2$ for ϕ in equation (4.1) we have:

$$(4.2) \quad \iint (\phi_1 - \phi_2) \vec{\nabla} (\phi_1 - \phi_2) \cdot d\vec{S} = \iiint \left[\vec{\nabla}(\phi_1 - \phi_2)\right]^2 dv$$

Either the boundary condition $\phi_1 = \phi_2$ or $\frac{\partial \phi_1}{\partial n} = \frac{\partial \phi_2}{\partial n}$ assures the vanishing of the left side of Equation (4.2), and hence the vanishing of the right side. Since the integrand of the right side of (4.2) is positive definite, it must vanish in order for this integral to vanish; hence throughout v, we have:

$$(4.3) \quad \vec{\nabla} \phi_1 = \vec{\nabla} \phi_2, \quad \phi_1 = \phi_2 + C \qquad C = \text{a constant}$$

Hence the two potentials that were assumed to be different yet solutions of the same boundary conditions can differ at most by an additive constant, which vanishes when the gradient is taken; therefore these potentials will give the same electric field distributions.

Solution by Green's Reciprocation Theorem

A large number of theorems that are useful for the solution of electrostatic problems serve to transform the solution of a known, presumably simpler problem, to the solution of another problem whose solution is desired. Of such theorems, one of the most useful is Green's reciprocity theorem. Let us consider a set of point charges where the potentials due to the other charges are given by a set of numbers ϕ_j and let the charges on such regions be q_j. The potential at the point j is related to the charge at the point j or on the region j by the expression:

$$(4.4) \qquad \phi_j = \frac{1}{4\pi k_o} \frac{q_i}{r_{ij}} \sum_{i=1}^{n'}$$

The prime on the summation sign means that the term where i = j is to be omitted from the summation. If, on the other hand, a set of charges q_j' are placed on the same regions, giving rise to the corresponding potentials ϕ_j', these are related by the expression:

$$(4.5) \qquad \phi_j' = \frac{1}{4\pi k_o} \frac{q_i'}{r_{ij}} \sum_{i=1}^{n'}$$

If we multiply Equation (4.4) by q_j' and Equation (4.5) by q_j, and then sum each equation over the index j, and note that r_{ij} is a symmetric function of the coordinates, we obtain:

$$(4.6) \qquad \phi_j q_j' = \frac{1}{4\pi k_o} \sum_{i=1}^{n'} \frac{q_i q_j'}{r_{ij}} \;\Bigg|\; \sum_{j=1}^{n} \phi_j q_j' = \sum_{j=1}^{n} \sum_{i=1}^{n'} \frac{q_i q_j'}{r_{ij}} \frac{1}{4\pi k_o}$$

$$\phi_j' q_j = \frac{1}{4\pi k_o} \sum_{i=1}^{n'} \frac{q_i' q_j}{r_{ij}} \;\Bigg|\; \sum_{j=1}^{n} \phi_j' q_j = \sum_{j=1}^{n} \sum_{i=1}^{n'} \frac{q_i' q_j}{r_{ij}} \frac{1}{4\pi k_o}$$

Since the indices i and j in the expressions on the right are summation dummies we can trade them in one expression and get:

$$(4.7) \qquad \sum_{j=1}^{n} \phi_j q_j' = \sum_{j=1}^{n} \phi_j' q_j$$

We can now generalize this theorem from a set of point charges to a set of n conductors of potentials ϕ_j carrying charges q_j; this generalization follows by combining the points of equal ϕ_j in Equation (4.7) into a single term. Equation (4.7) thus applies directly to such a system of conductors. An application of Equation (4.7) to the solution of a potential problem is given on page 44.

Solution by Green's Function

A great variety of solutions of potential problems can be generated from the knowledge of what is known as a Green's function. The Green's function for a particular geometrical arrangement is the solution of the potential problem for this given geometrical arrangement of grounded conducting boundaries, where the only charge present is a unit point charge located at a point p. It should be noted that the grounded conducting boundaries may be at infinity and the point charge need not be surrounded by a zero potential surface at a finite distance. It may be shown with

the aid of Green's Reciprocation Theorem that the Green's function for a particular geometry is a symmetrical function of the coordinates of the unit charge located at p' and the coordinates of the point of observation p.

Two general types of problems can be solved by the use of Green's function. One type of problem is one in which the potential distribution over a certain conducting boundary is given, and the other type of problem is one in which the charge distribution in a region within a conducting boundary is given. The derivation of the solution of both of these problems can be given together by means of Green's Theorem:

$$(1.18)\iiint (\phi\nabla^2\psi - \psi\nabla^2\phi)dv = \iint (\phi\vec{\nabla}\psi - \psi\vec{\nabla}\phi)\cdot d\vec{S}$$

Let us apply this theorem to the geometry of Figure (4.1)

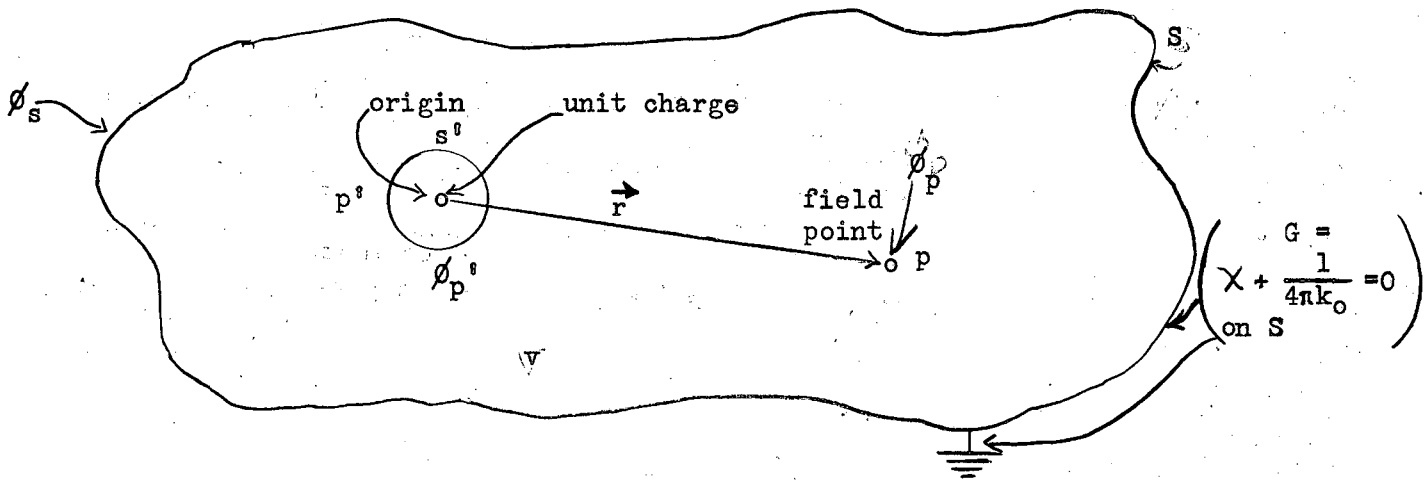


Figure (4.1)

ψ and ϕ in (1.18) are arbitrary functions of position which are required to be non singular throughout v. Let ϕ be the desired solution of a particular potential problem and let $\psi = G$ be the Green's function for the geometry of this problem, that is the solution of the problem of a unit point charge located at r = 0 with S grounded. G will be of the form:

$$(4.8)\quad G = \frac{1}{4\pi k_o r} + \chi = \psi$$

where χ represents the potential due to the induced charges on S. χ is harmonic* in v. G therefore has a singularity only at r = 0 and hence, as in Chapter 1, we must surround p' with a small sphere S' in order to satisfy the requirements associated with Green's theorem for the non singularity of ϕ and ψ. Hence substituting (4.8) into (1.18) we have:

$$(4.9)\quad \iiint (G\nabla^2\phi - \phi\nabla^2 G)dv = \iint_S (G\vec{\nabla}\phi - \phi\vec{\nabla}G)\cdot d\vec{S} - \iint_{S'} (G\vec{\nabla}\phi - \phi\vec{\nabla}G)\cdot d\vec{S}$$

Since v does not include r = 0, $\nabla^2 G = 0$ throughout v. Also by definition G = 0 on S. The integral over S' becomes, in the limit of very small radius:

$$(4.10)\quad \iint_{S'} (G\vec{\nabla}\phi - \phi\vec{\nabla}G)\cdot d\vec{S} \rightarrow -\phi_{p'} \iint_{S'} \vec{\nabla}\left(\frac{1}{4\pi k_o r} + \chi\right)\cdot d\vec{S} \rightarrow \phi_{p'}\ \frac{4\pi r^2}{4\pi k_o r^2} \rightarrow \frac{\phi_{p'}}{k_o}$$

Hence, collecting the non vanishing terms, we have:

$$(4.13)\quad \phi_{p'} = -k_o\left[\iiint G\nabla^2\phi dv + \iint \phi_s \vec{\nabla} G\cdot d\vec{S}\right]$$

Let us now consider two cases:

(1) The surface surrounding the point p' is grounded, making $\phi_S = 0$, and $\nabla^2\phi = -\rho/k_o$, due to the presence of the charge distribution ρ throughout v. Equation (4.13) then reduces to:

$$(4.14)\quad \phi_{p'} = -k_o \iiint G\nabla^2\phi dv = \iiint G\rho\, dv$$

This expression is, of course, fairly obvious, since it merely represents the principle of superposition applied to the density of point sources within the volume v, with each unit source, of which the density ρ consists, contributing its share to the potential $\phi_{p'}$ by the superposition indicated by the integral.

(2) Let there be no sources of ϕ throughout the volume v, $\nabla^2\phi = 0$, but let us

---

* A harmonic function is a function that is a solution of Laplace's equation (1.17).

assume that ϕ is a given function ϕ_s on the surface S. In this case Equation (4.13) reduces to:

$$(4.15) \quad \phi_{p'} = -k_o \iint \phi_s \vec{\nabla} G \cdot d\vec{S}$$

This gives the expression for the potential within a given region enclosed by a conducting boundary where different parts of the boundary are raised to a given set of potentials. This solution gives the potential within this boundary in terms of the surface integral of the potential on the boundary, multiplied by the normal derivative of the Green's function. Physically the normal derivative of the Green's function represents the surface charge density that is induced on the grounded conducting boundary by a unit charge at the point p'. The solution (4.15) then gives the solution of the potential problem corresponding to a given potential on the boundary in terms of the integral of this potential, multiplied by the induced charge produced on the grounded boundary by a unit charge placed at the field point. If we wish to express (4.15) directly in terms of the charge σ_{ls} induced on the grounded boundary we note from Equation (3.16) that:

$$(4.16) \quad \vec{\nabla} G \cdot \frac{d\vec{S}}{dS} = + \frac{\sigma_{ls}}{k_o}$$

and thus (4.15) becomes:

$$(4.17) \quad \phi_{p'} = - \iint \phi_s \sigma_{ls} dS$$

Theorem (4.17) may also be derived directly by the use of Green's Reciprocation Theorem, (4.7). Let us consider for the two cases to be used in the reciprocation theorem:

(1) Let the surface S be grounded and let a single charge $q_{p'}$ be located at the point p', and let this charge induce charges q_{lj} on the j^{th} region of the boundary S.

(2) Let the charge at p' be removed, but let the surface S be divided into

sections, each at a constant potential, the potential of the $j^{\underline{th}}$ section of S being ϕ_{js}. Let $\phi_{p'}$ = potential at p' in case (2).

Relating these two cases by means of (4.7) we obtain:

$$q_{p'}\,\phi_{p'} + \sum_{j=1}^{n} q_{1j}\,\phi_{js} = 0 + 0 \tag{4.18}$$

Note that the two zeros on the right side of (4.18) arise from the fact that on the one hand the potential is zero over the entire boundary in case (1), and that the charge at p' is zero in case (2). If we let $q_{p'}$ be unity in the first case, we obtain:

$$\phi_{p'} = -\sum_{j=1}^{n} \phi_{js} q_{1j} \tag{4.19}$$

This expression is identical with (4.17) but has been obtained, directly in terms of the induced charges, in a way that is more obvious physically. Several examples will be given later of the derivation of Green's function for various conducting boundaries. The solutions to potential problems, of the two types mentioned on page 43, can then be written down immediately.

Solution by Inversion

The process of inversion is a special case, that is, valid in two dimensions as well as in three dimensions, in which a set of solutions of one potential problem can be transformed into the solutions of another potential problem. The inversion transformation is a restricted type of transformation. In two dimensions more general classes of such transformations can be found, than the inversion, but in three dimensions more general classes cannot be obtained.

One of the methods by which the solution of a problem can often by transformed into the solution of a simpler problem, is the inversion transformation on a sphere, as shown in Figure (4.2). It can be shown by direct differentiation that, if $\phi_p = \phi(r,\theta,\phi)$ is a solution of Laplace's equation then $\psi_{p'} = \psi(r',\theta,\phi) =$

$\frac{a}{r'}\,\phi\,(\frac{a^2}{r'},\theta,\phi)$ is also a solution of Laplace's equation. This transformation of the point r into the point r', by the relation $rr' = a^2$, maps the point $p(r,\theta,\phi)$ into its inversion point $p'(a^2/r,\theta,\phi)$, moving the point along the radius vector from a position inside the sphere of radius a to a point outside this sphere, or vice versa. Let a charge q and a charge q' be placed as shown in Figure (4.2). The relations $rr' = a^2$ and $\ell\ell' = a^2$ cause the ratio, $r/\ell' = \ell/r'$

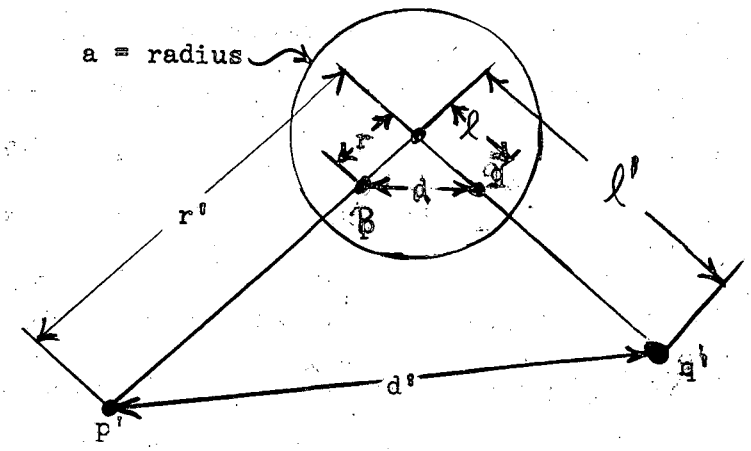


Figure (4.2)

to hold, and thus the triangles, $r\ell d$ and $r'\ell'd'$ are similar. Thus we have:

$$(4.20)\quad \frac{r}{\ell'} = \frac{\ell}{r'} = \frac{d}{d'}$$

The potential at p before inversion is $\phi_p = q/4\pi k_o d$ and the potential at p' after inversion is $\phi_{p'} = q'/4\pi k_o d'$, so we have:

$$(4.21)\quad \frac{\phi_{p'}}{\phi_p} = \frac{q'}{d'}\frac{d}{q} = \frac{q'}{q}\frac{\ell}{r'} = \frac{q'}{q}\frac{r}{\ell'}$$

A suitable law for the inversion of charges must now be formulated. We can verify that the sphere of radius a will be at zero potential if a charge q is placed at a distance ℓ from the center and simultaneously a charge $q' = -qa/\ell$ is placed at a point on the same radius vector and at a distance ℓ' from the center such that $\ell\ell' = a^2$. We know from above that zero potential surfaces transform into zero potential surfaces. Thus the inversion sphere at zero potential under

the influence of the two charges is to remain so after inversion. This is assured if the two charges change places thus:

$$(4.22)\quad q_I = q \text{ at } \ell \text{ becomes } q'_I = aq/\ell \text{ at } a^2/\ell$$

$$q_{II} = -qa/\ell \text{ at } a^2/\ell \text{ becomes } q'_{II} = -q \text{ at } \ell$$

So if a sphere which is originally at zero potential is to remain at zero potential after inversion about its own radius then we must use the relation:

$$(4.23)\quad \frac{q'}{q} = \frac{q\frac{a}{\ell}}{q} = \frac{a}{\ell} = \frac{\ell'}{a} = \sqrt{\frac{\ell'}{\ell}}$$

for the inversion of charges. It is felt that it is more convenient for a charge to retain its original sign and only change its magnitude when it is inverted, although this is not necessary if all charges undergoing an inversion are treated in the same way. We now secure, by substituting (4.23) into (4.21) the rule for the inversion of potentials.

$$(4.24)\quad \frac{\phi_{p'}}{\phi_p} = \frac{a}{r'} = \frac{r}{a} \quad , \text{ In agreement with } \phi_{p'} = \frac{a}{r'}\phi((r,\theta,\phi)$$

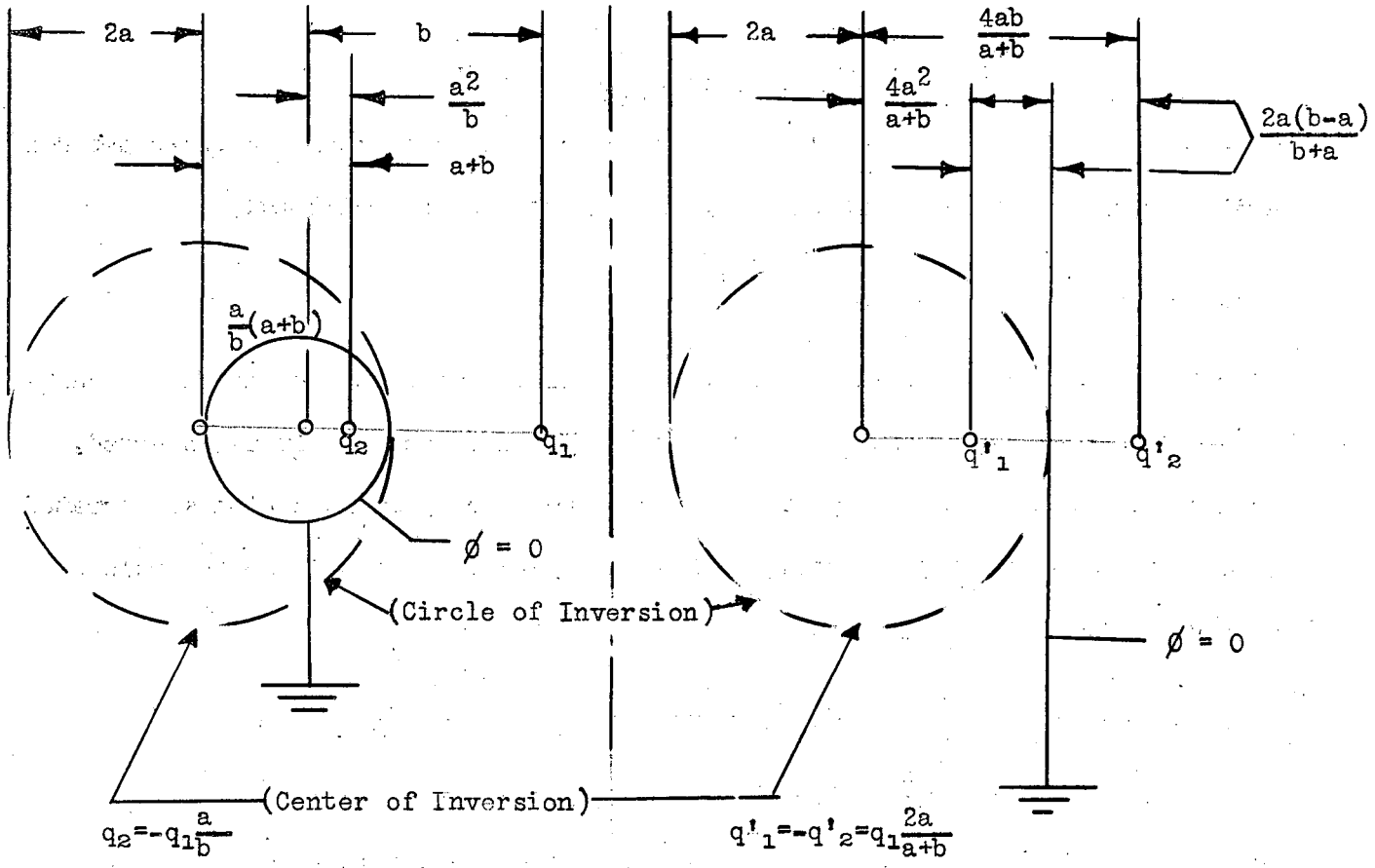
The transformation equations for such quantities as volume or surface charge densities can be obtained by multiplying the charge transformation (4.23) by the transformation of the appropriate geometrical quantities.

In an inversion transformation a point charge will often appear at the center of inversion in the transformed geometry. This point charge arises from the fact that the net charge in the original geometry had electric field lines that terminated on equal and opposite charges located at infinity, and in the inversion, infinity is brought in to the origin.

The main utility of the inversion transformation is that it rectifies spherical boundaries. Two freely charged intersection spheres may be inverted into two intersecting planes, and the plane boundary problem is usually soluble by the method of images.

Solution by Electrical Images.

The method of inversion discussed above justifies the solution of the problem of a point charge opposite a grounded conducting sphere, as shown in Figure (4.3A), by means of the method of electrical images. The sum of the potential from a point



charge q located in free space near a conducting sphere of radius a, and from the charge that is induced on this sphere by q, will be a potential distribution in which this sphere is an equal potential surface. The uniqueness theorem requires that the potential outside a grounded conducting sphere, under the influence of a point charge q, which is geometrically coincident with a sphere formed from the inversion of a grounded conducting plane under the influence of a charge and its electrical image in the plane, will be identical, if the distances of the charges from the sphere and plane obey the rule for inversion derived above. The potential corresponding to the point charge and the image which it makes in the grounded sphere is therefore the correct Green's function in the region bounded by the conducting sphere and by infinity. By means of the method of electrical images and the method

of inversion, the solutions can be immediately written down to problems of the type in which various areas of the surface of a conducting sphere are raised to different arbitrary potentials and to problems of the type in which there is an arbitrary charge distribution in the region outside of a grounded conducting sphere. Similarly by this method corresponding solutions to potential problems involving the inside of a grounded conducting sphere can be obtained.

The solution of problems involving conducting planes can be easily obtained by the method of images, since the sum of the potential due to a point charge and its geometrical mirror image of the opposite sign produces a zero potential on the conducting plane of reflection in which the mirror image is formed. This means that problems involving a plane of which various areas are raised to different arbitrary potentials can be solved by using the potential of the point charge and its image as a Green's function. Figure (4.3E) shows that the plane image solution and the sphere image solution can be generated from one another by inversion.

CHAPTER 5 TWO DIMENSIONAL POTENTIAL PROBLEMS

Potential problems involving geometrical arrangements that may be approximated by a two dimensional geometry with an infinite uniform extent in the third direction are frequently easier to solve than three dimensional problems. Certain mathematical techniques may be applied to them which do not exist in the case of three dimensional problems. The method of complex variable potential description combined with conformal transformation is an especially powerful method of solution.

We shall show that in two dimensions any function W, of a complex variable z_1, which is analytic will have real and imaginary parts each of which individually satisfies Laplace's equation in two dimensions. Thus a suitable function $W = W(z_1)$ can completely describe the potential surfaces and the field lines caused by the geometry of a particular problem. If, $\phi + i\psi = W = W(z_1) = W(x_1 + iy_1)$, we may separate real and imaginary parts and obtain $\phi = \phi(x_1,y_1)$ and $\psi = \psi(x_1,y_1)$. The functions ϕ and ψ will be the equi-potential and field line surfaces or vice versa. Therefore any transformation from one complex variable z_1 to another z_2 will transform the solution of one potential problem described by the first variable to the solution of another potential problem described by the second variable. In general a whole class of two dimensional potential distribution problems can be solved by the following process:

(1) Obtain a transformation $z_2 = f(z_1)$ which will transform the geometric arrangement of the z_1 coordinate system into an arrangement of the z_2 coordinate system which will bring about a simplification in the problem. This coordinate transformation, $z_2 = f(z_1)$ or $z_1 = g(z_2)$, must be so suited that it will carry the complex potential geometry $W = W_1(z_1)$ of the original problem into a simpler complex potential geometry $W = W_1[g(z_2)] = W_2(z_2)$.

The new complex potential W_2 must be that of a more easily soluble potential problem.

(2) Express the potential solution ϕ in the transformed z_2 plane in such a way that $\phi + i\Psi$ is an analytic function of a complex variable.

(3) Transform this solution back to the original z_1 plane.

We shall now discuss the justification for this process. Consider a function $W = \phi + i\Psi = f(z)$ where $z = x + iy$. In order for this functional relationship to be analytic, ϕ and Ψ must fulfill the Cauchy-Riemann differential equations:

$$(5.1)\quad \frac{\partial \Psi}{\partial x} = -\frac{\partial \phi}{\partial y}$$

$$(5.2)\quad \frac{\partial \Psi}{\partial y} = \frac{\partial \phi}{\partial x}$$

By partial differentiation of Eq. (5.1) with respect to x and partial differentiation of Eq. (5.2) with respect to y and then adding the two resulting equations, and then repetition of the process with x and y interchanged, followed by subtraction we have:

$$(5.3)\quad \nabla^2\Psi = \nabla^2\phi = 0$$

Thus both ϕ and Ψ are harmonic functions*. The functional relationship $W = W(z)$ can be demonstrated graphically, as in Figure (5.1) by plotting the lines ϕ = constant,

*Harmonic functions are functions which are solutions of Laplace's equation.

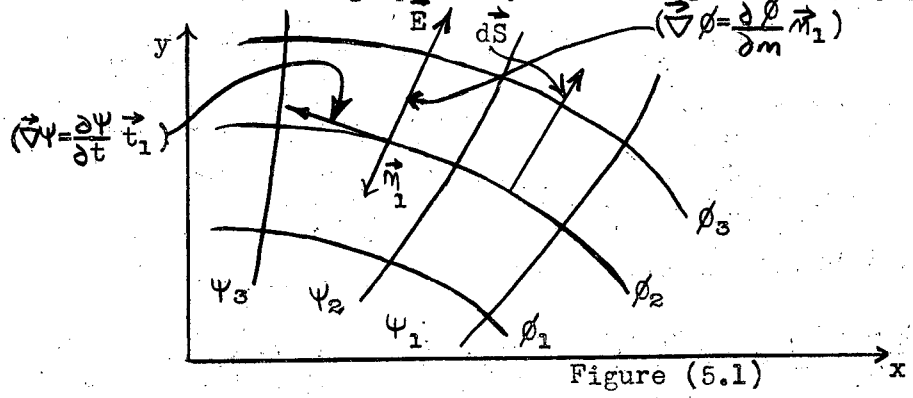


Figure (5.1)

$\vec{n}_1$ = unit vector normal to ϕ curves.

$\vec{t}_1$ = unit vector tangent to ϕ curves.

and the lines Ψ = constant in the $z = x + iy$ plane, after the function W has been separated into its real and imaginary parts.

The ϕ curves secured by giving a succession of values to W may for example represent the potential field of a problem, and the corresponding Ψ curves represent the electric field *, but these latter are usually referred to as the stream curves.

---

*Note that since the functions ϕ and Ψ satisfy the Cauchy-Riemann relations (5.1) and (5.2) that the curves ϕ = constant and Ψ = constant are normal to each other. This may be easily seen, since if we solve for the slopes of $\phi = \phi(x,y)$ and $\Psi = \Psi(x,y)$ we have:

$$\left(\frac{dy}{dx}\right)_{\phi = \text{const.}} = -\frac{\frac{\partial \phi}{\partial x}}{\frac{\partial \phi}{\partial y}} \quad \text{and} \quad \left(\frac{dy}{dx}\right)_{\Psi = \text{const.}} = -\frac{\frac{\partial \Psi}{\partial x}}{\frac{\partial \Psi}{\partial y}}$$

Substituting the relations (5.1) and (5.2) we see that:

$$\left(\frac{dy}{dx}\right)_{\phi} = -\left(\frac{dy}{dx}\right)_{\Psi}^{-1}$$

and thus at any common point the curves will be normal to each other.

The flux of the electric field which crosses a surface S, as in Figure (5.1), lying along one of the equi-potential curves, ϕ = constant, between two stream curves Ψ_1 and Ψ_2, and of unit height normal to the z plane is then given by

(5.4)** $\Phi = \Psi_2 - \Psi_1$ $\qquad \Phi = \iint_S \vec{E} \cdot d\vec{S}$ the electric field flux crossing S

**The derivation of this equation is as follows:

$$\Phi = \iint_S \vec{E} \cdot d\vec{S} = -\iint_S \vec{\nabla}\phi \cdot d\vec{S} = -\iint_S \left(\frac{\partial \phi}{\partial x}\vec{i} + \frac{\partial \phi}{\partial y}\vec{j}\right) \cdot d\vec{S} = -\iint_S \frac{\partial \phi}{\partial n}\, dn$$

By substitution from (5.1) and (5.2) we have:

$$\Phi = \iint_S \left(\frac{\partial \Psi}{\partial y}\vec{i} - \frac{\partial \Psi}{\partial x}\vec{j}\right) \cdot d\vec{S} = \iint_S (\vec{\nabla}\Psi \times d\vec{S}) \cdot \vec{n}_1 = \iint_S \frac{\partial \Psi}{\partial t}\, dt = \Psi_2 - \Psi_1$$

Thus the difference between two stream functions Ψ_1 and Ψ_2 represents the electric flux passing between the unit height right cylinders generated by two neighboring lines Ψ_1 and Ψ_2. This means that no lines of force cross the constant Ψ lines. This is the justification for calling Ψ the stream function, since in hydrodynamic problems in two dimensions, the Ψ lines actually do trace the stream lines of the fluid. Since electric flux lines do not cross the stream lines, then the stream lines traced by giving Ψ different constant values will trace the electric field, when ϕ lines are the equipotentials of the field. If, on the other hand, Ψ had been assumed to be the potential, then ϕ would have been the stream function. In fact, this trading of the meaning of ϕ and Ψ is frequently useful in the solution of two dimensional problems.

The above considerations permit us to obtain immediately the capacity between any two conductors whose boundaries coincide with two equipotential lines ϕ_1 and ϕ_2, and extend between two stream lines Ψ_1 and Ψ_2. The capacity is given by:

$$C = \frac{q}{|\Delta \phi|} = \frac{k_o \Phi}{|\Delta \phi|} \qquad (5.5)$$

$\Phi = \iint_S \vec{E} \cdot d\vec{S} = \frac{q}{k_o}$ from Equation (1.6) and the definition of Φ.

C = electrical capacity

and since the flux Φ is the change in the stream function Ψ between the edges of the conductor surfaces being considered, this becomes:

$$C = k_o \frac{\Psi_1 - \Psi_2}{\phi_1 - \phi_2} \qquad (5.6)$$

Note that in general the stream function is multiple valued if charges are present in the field. The net flux from a charged body is the multivaluedness of the stream function in one circuit around the body.

Another quantity which can be calculated from a known function of the form $W = W(z)$, representing a particular geometry as in Figure (5.1), is the absolute magnitude of the field strength. Consider the modulus of the derivative of W:

$$(5.7)\quad \left|\frac{dW}{dz}\right| = \left|\frac{\partial(\phi + i\Psi)}{\partial x}\frac{dx}{dz} + \frac{\partial(\phi + i\Psi)}{\partial y}\frac{dy}{dz}\right|$$

$$\left|\frac{dW}{dz}\right| \left|\frac{\partial\phi}{\partial x}dx + i\frac{\partial\Psi}{\partial y}dy + i\frac{\partial\Psi}{\partial x}dx + \frac{\partial\phi}{\partial y}dy\right| \left|\frac{1}{dz}\right|$$

Using the Cauchy-Riemann differential equations (5.1) and (5.2), this becomes:

$$(5.8)\quad \left|\frac{dW}{dz}\right| = \left|\frac{\partial\phi}{\partial x} - i\frac{\partial\phi}{\partial y}\right| = \left|\sqrt{\left(\frac{\partial\phi}{\partial x}\right)^2 + \left(\frac{\partial\phi}{\partial y}\right)^2}\right| = \left|\vec{\nabla}\phi\right| = \left|\vec{E}\right|$$

The real and imaginary parts of dW/dz are thus respectively the x and y components of the gradient of the potential and therefore the modulus of dW/dz is equal to the magnitude of the electric field strength.

We shall now consider the solution of several simple potential problems expressed as functions of a complex variable. A much wider group of solutions can then be obtained from these simple cases by any complex variable transformation.

The Potential of a Line Charge.

The Coulomb field around a line charge with a linear charge density q is found by the use of Gauss' electric flux theorem, Eq. (1.6), carrying out the integration over a circular cylinder of radius r, and of unit length, arranged coaxial with the line charge. This yields the field:

$$(5.9)\quad \vec{E} = \frac{q\vec{r}}{2\pi k_o r^2}$$

The potential produced by the line charge may be secured by substitution of (5.9) into Eq. (1.11) and then carrying out a direct integration of the two dimensional Coulomb field, we have:

$$(5.10)\quad \phi = -\frac{q}{2\pi k_o}(\ln r - \ln r_o)$$

Note that in two dimensional potential expressions it is not possible to arbitrarily set the potential at infinity equal to zero since the two dimensional expression really represents the potentials due to charge distributions which are of infinite extent along the z axis* and therefore the influence of the distribution does not

*The z axis is normal to the (x,y) or z plane and sould not be confused with the z complex variable z = x + iy.

decrease toward infinity. Consequently another zero point than infinity must be chosen for the potential base. It must be remarked that a two dimensional problem can, at most, only be an approximation to physical reality since it really implies an infinite charge along any charged object, since this object extends to infinity, and thus infinite energies. A physical problem can only be treated by two dimensional methods when it is possible to neglect end effects arising from the finite linear extent of the physical arrangement. In equation (5.10) the cylinder surrounding the line charge at a distance r_o has been arbitrarily set at zero potential.

The complex potential function W corresponding to the line charge potential (5.10) is therefore seen by inspection, or a more complicated use of the Cauchy-Riemann equations to be:

$$(5.11)\quad W = -\frac{q}{2\pi k_o}(\ln z - \ln z_o) = -\frac{q}{2\pi k_o}(\ln r + i\theta - \ln r_o) = \phi + i\psi \begin{cases} z = re^{i\theta} \\ z_o = r_o \end{cases}$$

where ϕ represents the potential and ψ, the stream function. Note that the stream function, as might be expected from the axial symmetry of the arrangement, is proportional to the polar angle θ, where z_o has been arbitrarily taken as real. The complex potential function for any system of line charges can be obtained by the superposition of appropriate expressions like (5.11), one for each line charge.

The Potential of a Uniform Field.

The complex potential corresponding to a uniform field $\vec{E}$, directed in the +x direction can be seen by inspection to be:

$$(5.12)^{*}\quad W = -|E|z = -|E|(x + iy) = \phi + i\psi$$

where the potential function is $\phi = -|E|x$, and the stream function is $\psi = -|E|y$. The above two cases are the basic potentials from which many more general cases may be generated by transformations and superpositions.

* The potential ϕ has been arbitrarily set equal to zero along the y axis.

We now turn to the analysis of the behavior of curves in a small piece of a complex potential plane when a transformation of the plane is made. Consider a transformation from the z_1 plane to the z_2 plane given by the equation $z_2 = f(z_1)$ and let the transformation function f be analytic, except at a finite number of singularities. At all non-singular points such a transformation is conformal. This means that the angle between two intersecting lines in the z_1 plane, as θ_1 in Figure (5.2A), transforms into an equal angle between the transformations of the two lines in the z_2 plane, as θ_2 in Figure (5.2B). This can be demonstrated as follows:

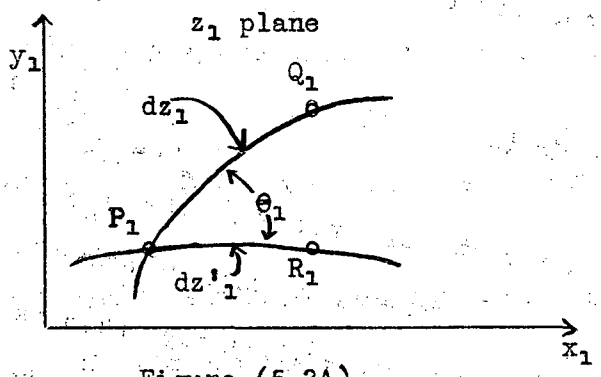


Figure (5.2A)

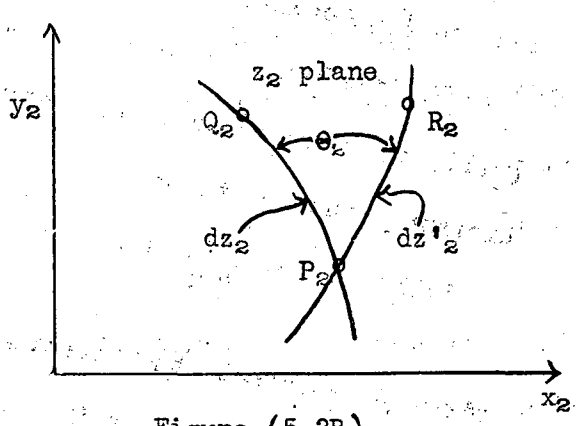


Figure (5.2B)

Since all derivatives of an analytic function of a complex variable exist and are continuous, the derivative of the transformation, dz_2/dz_1 will be finite at all points, excepting the singularities. Taking the derivative $dz_2/dz_1 = f'(z_1)$ and remembering that the argument of a product is equal to the sum of the arguments of the factors, we have for the argument of the differential line element P_2Q_2:

$$(5.13)\quad \arg(dz_2) = \left[\arg f'(z_1)\right]_{P_1} + \arg(dz_1)$$

and taking the argument of P_2R_2:

$$(5.14)\quad \arg(dz'_2) = \left[\arg f'(z_1)\right]_{P_1} + \arg(dz'_1)$$

Subtracting (5.14) from (5.13), and noting that the angles θ_1 and θ_2 are the differences in the arguments of the respective dz's we have:

$$(5.15)\quad \theta_1 = \theta_2$$

In addition the modulus of the derivative $dz_2/dz_1 = f'(z_1)$ represents the scale factor by which all spatial intervals, in the neighborhood of a point,

are multiplied. This may be seen by noting that the modulus of a product is equal to the product of the moduli of the factors. So an infinitesimal triangle will transform into a similar infinitesimal triangle in the new system. Thus:

$$(5.16) \quad \left|dz_2\right| = \left|f'(z)_p\right| \cdot \left|dz_1\right|$$

The similarity of this transformed infinitesimal trangle is an alternate way of seeing that angles are preserved in complex transformations. This means that the othogonality between stream functions and equi-potentials is invariant under a complex variable transformation.

Schwarz Transformation.

The transformation required to reduce a complicated set of boundaries in the z_1 plane to a single straight line boundary in the z_2 plane can be derived in the case of rectlinear boundaries by the use of the Schwarz transformation. The Schwarz transformation will map the inside of a polygon in the z_1 plane into the upper half of the z_2 plane.* This transformation will be introduced by illustrating the bending of one angle whose vertex is at the origin, as seen in Figures (5.3A) and (5.3B). Consider the simple transformation:

$$(5.17) \quad z_1 = z_2^{\beta}$$

where β is not necessarily an integral or a rational number. By this transformation,

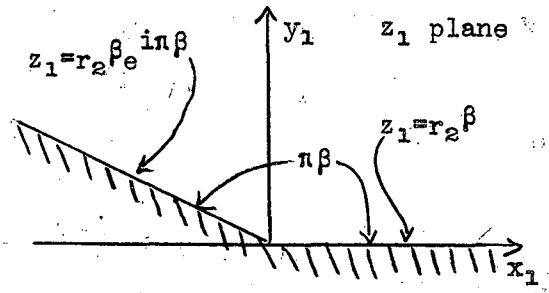


Figure (5.3A)

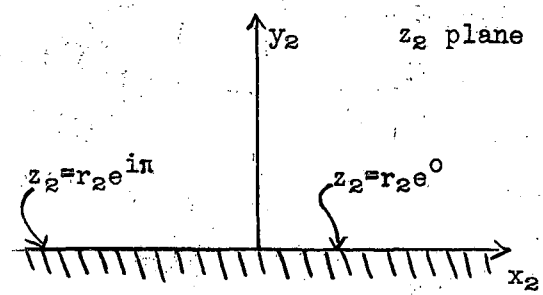


Figure (5.3B)

points lying on the positive real axis in the z_2 plane are mappings of the points on the positive real axis in the z_1 plane with the scale along the axis changed by raising the x_1 coordinate to the $1/\beta$ power, or at least a branch of the transformation can be chosen where this is so.

* A complementary transformation can be derived which will map the outside of a polygon into the upper half plane.

On the other hand, for points lying on the negative real axis in the z_2 plane, that is if $z_2 = r_2 e^{i\pi}$ z_1 will be complex, for by the transformation $z_1 = r_2^{\beta} e^{i\pi\beta}$. Hence the negative real axis of the z_2 plane will be the mapping of a straight line in the z_1 plane, as is required by the conformal properties of the transformation, but this line will make an angle $\pi\beta$ with the positive real z_1 axis. The transformation (5.17) is therefore the transformation which maps the area of the upper half of the z_1 plane lying between $\theta_1 = 0$ and $\theta_1 = \pi\beta$ into the entire upper half of the z_2 plane. The transformation (5.17) has a branch point at $z_2 = 0$, but is analytic everywhere else.

Now let us consider the more general case shown in Figure (5.4) in which we have a number of points b_i in the z_1 plane which are the corners of a polygon whose interior angles are α_i. We wish to map the interior of the polygon into

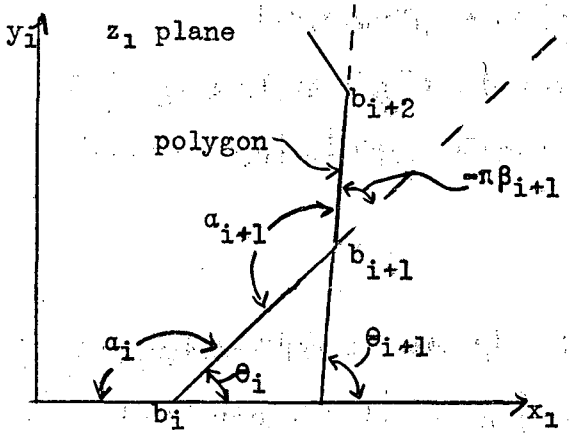


Figure (5.4A)

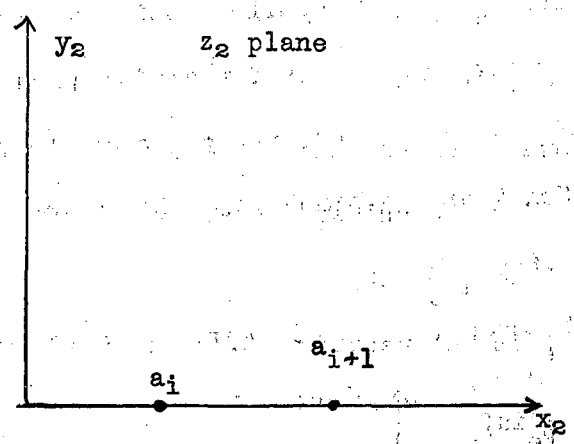


Figure (5.4B)

the upper half of the z_2 plane. Consider a transformation defined by the differential equation:

$$(5.18)\quad \frac{dz_1}{dz_2} = C_1 \prod_{i=1}^{n} (z_2 - a_i)^{\beta_i} \qquad C_1 = \text{a constant, possibly complex.}$$

This transformation will by analytic everywhere except at the points $z_1 = a_i$. Hence by the conformal properties of such a transformation the real z_2 axis, $z_2 = x_2$, will be the mapping of straight line segments in the z_1 plane. The angles which each of these straight line segments make with the real axis will be given by the argument of dz_1/dz_2 evaluated in the segment in question. If we take the

argument of Eq. (5.18) we get:

$$(5.19)\quad \arg\left(\frac{dz_1}{dz_2}\right) = \arg C_1 + \beta_1 \arg (z_2-a_1) + \beta_2 \arg (z_2-a_2) + \cdots + \beta_n \arg (z_2-a_n)$$

We may also express:

$$(5.20)\quad \arg\left(\frac{dz_1}{dz_2}\right) = \arg\left(\frac{dx_1+idy_1}{dx_2+idy_2}\right)$$

When dz_2 lies along the real axis in the z_2 plane, $dy_2 = 0$. So Eq. (5.20) becomes, when evaluated in the $i^{\underline{th}}$ interval:

$$(5.21)\quad \arg\left(\frac{dz_1}{dz_2}\right) = \arg\left(\frac{dx_1}{dx_2} + i\frac{dy_1}{dx_2}\right) = \mathrm{Tan}^{-1}\left(\frac{dy_1}{dx_1}\right) = \theta_i$$

Now when z_2 lies on the real axis between a_i and a_{i+1}, the $\arg(z_2-a_{r\leq i}) = 0$ and the $\arg(z_2-a_{r>i}) = \pi$. So substituting these values and (5.21) into (5.19) we get:

$$(5.22)\quad \theta_i = \arg C_1 + (\beta_{i+1} + \beta_{i+2} + \cdots + \beta_n)\pi$$

Thus all points of the real axis segment $a_{i+1}-a_i$ are mappings of a line segment with slope θ_i in the z_1 plane. Now subtracting (5.22) from a similar expression to (5.22) for θ_{i+1} we obtain:

$$(5.23)\quad \theta_{i+1} - \theta_i = -\pi\beta_{i+1}$$

From the geometry of Figure (5.4A) we see that this angle difference of $-\pi\beta_{i+1}$ at the point b_{i+1} is related to the interior angle α_{i+1} at each point by the relation:

$$(5.24)\quad \alpha_{i+1} = \pi + \pi\beta_{i+1}$$

Or transposing and changing the subscript by -1 we get:

$$(5.25)\quad \beta_i = \frac{\alpha_i}{\pi} - 1$$

Hence equation (5.18) becomes:

$$(5.26)\quad \frac{dz_1}{dz_2} = C_1 \prod_{i=1}^{n} (z_2-a_i)^{(\frac{\alpha_i}{\pi} - 1)}$$

where the scale factor C_1 gives both the relative scale and the relative angular orientation of the two geometries.

In general, the Schwarz transformation is a useful one provided that Equation (5.26) is integrable in terms of elementary functions. This is true, with the exception of special cases, only when the angles are multiples of 90° and not more than two corners are involved. In case the angles are not multiples of 90° and more than one corner is involved then (5.26) is usually not integrable in terms of elementary functions. One further difficulty in the practical application of the Schwarz transformation is the fact that the resultant transformation is given in terms of $z_1 = f(z_2)$ with the coordinates along the real axis in the z_2 plane as the independent variable, rather than in terms of the coordinates of the z_1 plane, or the desired polygon as independent variable. Therefore, with the exception of the most simple cases, considerable computational labor is often necessary to find out what the coordinates a_i of the mappings in the z_2 plane of the corners of the polygon in the z_1 plane actually are in terms of the geometry of the given problem. Once the a_i are determined, the remainder of the solution of the potential problem is usually simple. Let us now consider some special simple cases of the Schwarz transformation.

(1) Single Angle Transformations; $a_i = 0$ in each case.

(a) $\alpha = \pi$. The integration of Equation (5.26) gives:

$$z_1 = C_1 z_2 + C_2 \tag{5.27}$$

This is simply a uniform translation and rotation and is of no physical interest.

(b) $\alpha = \pi/2$. The integration of Equation (5.26) gives:

$$z_1 = C_3 z_2^{1/2} + C_2 \tag{5.28}$$

We will omit the constant of translation C_2. This will map the first quadrant of the z_1 plane into the upper half of the z_2 plane. If, for example we assume that the complex potential in the z_2 plane is given by the uniform field complex potential solution:

$$W = -|E| z_2, \quad W = \phi + i\psi, \begin{cases} \phi = -|E| x_2 \\ \psi = -|E| y_2 \end{cases} \tag{5.12}$$

The W function (5.12) will transform with the relation (5.28) to:

$$(5.29)\quad W = -C_4 |E| z_1^2,\ W = \phi + i\psi \qquad \begin{cases} \phi = -C_4 |E| (x_1^2 - y_1^2) \\ \psi = -C_4 |E| (2x_1y_1) \end{cases}$$

$$W = -C_4 |E| (x_1^2 - y_1^2 + 2ix_1y_1)$$

and will solve the problem of a charged rectangular boundary, see Figure (5.5), or also problems involving charged hyperbolic cylinders.

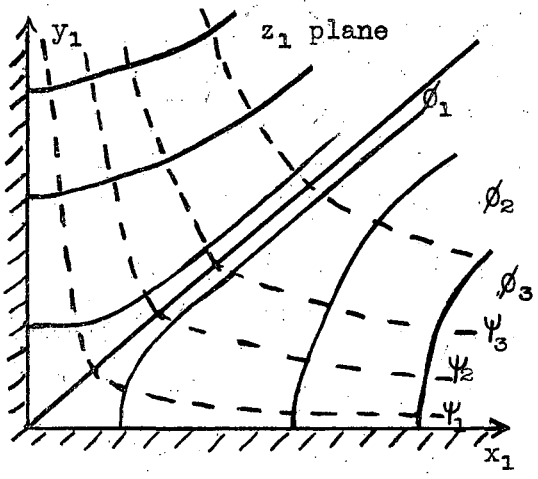


Figure (5.5A)

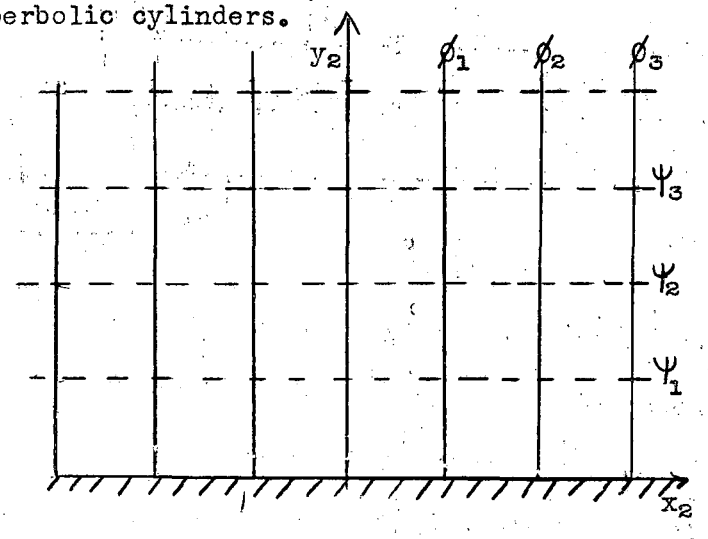


Figure (5.5B)

If the complex potential in the z_2 plane is taken to be the logarithmic potential corresponding to a line charge, Equation (5.11), and if the transformation (5.28) is then applied, we obtain the two dimensional Green's function for an inside rectangular corner, if we have translated the line charge into the upper half of the z_2 plane. This same transformation will give the Green's function for a problem having hyperbolic cylindrical boundaries, and thus problems involving such geometries are amenable to solution.

(c) $\alpha = 0$. Equation (5.26) integrates into:

$$(5.30)\quad z_1 = C_3 \ln z_2 + C_2$$

Omitting C_2 we have:

$$(5.31)\quad z_1 = C_3 \ln z_2 = C_3 \ln r_2 + C_3 i\theta_2$$

If C_3 is real, the real part of z_1 is then $C_3 \ln r_2$, and the positive real z_2 axis is the mapping of the whole real z_1 axis, and the upper half z_2 plane maps into a

strip of width $C_3\pi$, as in Figure (5.6). The transformation can be visualized by considering the origin in the z_2 plane to be pushed to minus infinity, and the negative real axis of the z_2 plane to be revolved clockwise to a position parallel to the positive real axis but located above it a distance $C_3\pi$, as seen in Figure (5.6A). This transformation thus results in a periodic configuation in the z_1 plane, of which the strip $C_3\pi$ wide is the first repeat. The upper half of the z_2 plane is the mapping of the first strip of this configuration in the z_1 plane. The lower half of the z_2 plane is the mapping of the strip lying between $y_1 = C_3\pi$ and $y_1 = 2C_3\pi$ in the z_1 plane, and so on. This transformation is a very useful one in the solution of potential problems involving grids, repeating condenser plates, and other geometries that repeat in one direction.

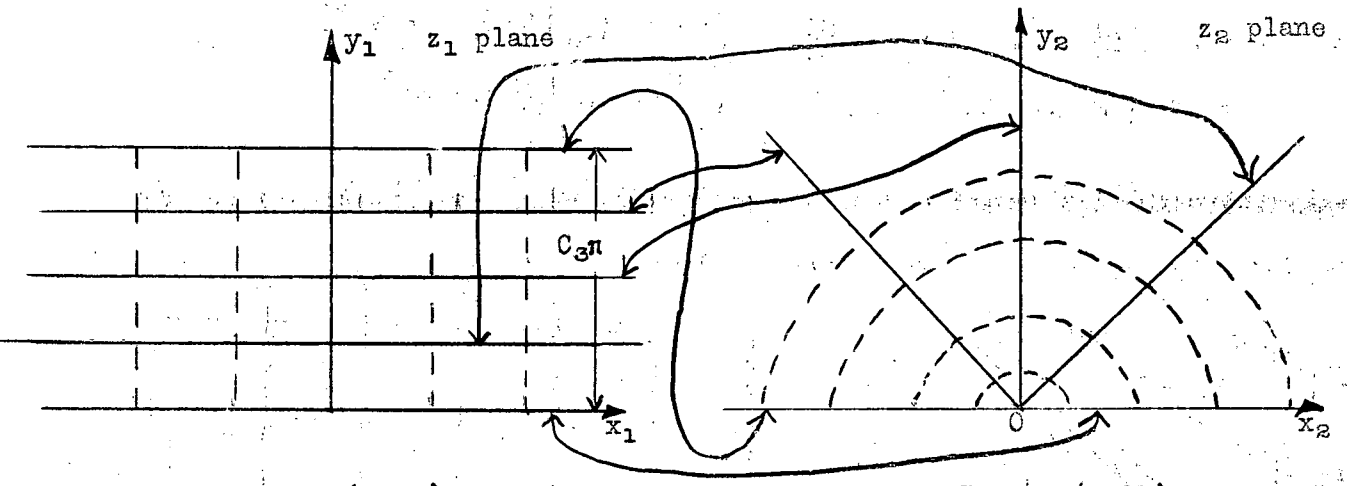


Figure (5.6A) Figure (5.6B)

(2) Multiple Angle Transformations.

If two upward 90° bends are made in the real axis of the z_2 plane at $\pm a$, then the z_2 plane will be the mapping of the vertically oriented, semi-infinite strip seen in Figure (5.7A). The differential equation (5.26) becomes:

$$(5.32)\quad \frac{dz_1}{dz_2} = C_1(z_2-a)^{-1/2}(z_2+a)^{-1/2} = \frac{C_1}{\sqrt{z_2^2-a^2}} = \frac{C_2}{\sqrt{a^2-z_2^2}}$$

The relation $C_1 = iC_2$ was introduced to rotate the figure 90° to place it in the orientation shown. Equation (5.32) integrates into:

$$(5.33)\quad z_1 = C_2\sin^{-1}\left(\frac{z_2}{a}\right),\quad z_2 = a\sin\left(\frac{z_1}{C_2}\right)$$

In practice, this transformation is mainly used as a transformation of the z_1 plane into the z_2 plane. If we consider a uniform complex potential field W in the z_1 plane:

(5.12) $W = - |E| z_1$

and if we map this field into the z_2 plane by the relation (5.33) then the cross section is produced of a field distribution corresponding to the potential around a charged conducting strip of width 2a, or trading potential and stream functions, the potential due to a slot, in a conducting sheet, of width 2a along the real z_2 axis. The major axis of the slot or strip will lie normal to the plane of Figure (5.7B). Equation (5.12) into which (5.33) has been substituted can be resolved

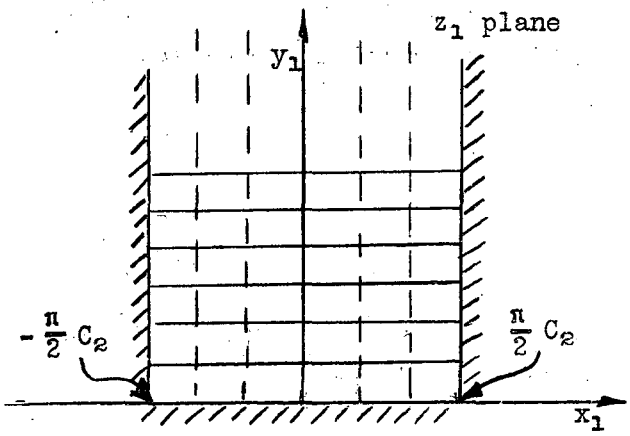


Figure (5.7A)

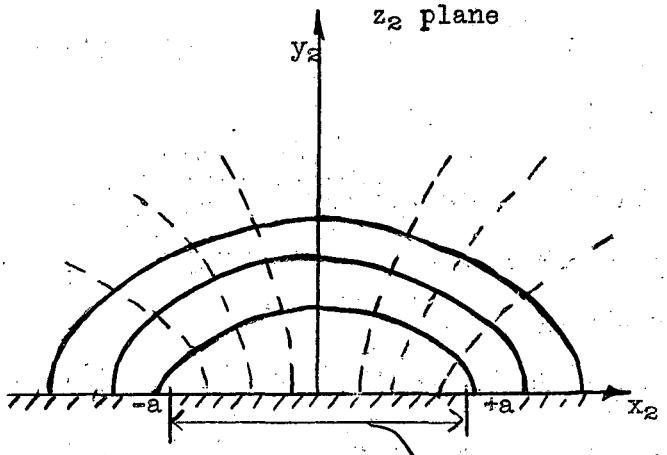


Figure (5.7B) (slot or strip)

into its real and imaginary parts, and if the real part ϕ and the imaginary part ψ are given a set of constant values, they will characterize the equations of the arrangement. These equations turn out to be the equations of confocal elliptic and hyperbolic cylinders, as in Figure (5.7B).

If for instance the fields at each of a series of condenser plates of alternating potentials are given, by an additional zero angle transformation the above solution for a slot can be used for the calculation of multiple condenser plate end effects. There are a great many examples of cases where the Schwarz transformation results in integrable differential equations. Frequently the solution appears in a form which is deceptively simple since it is usually

stated in the terms of the coordinates of the complex geometry, z_1, as a function of the coordinates of the simple geometry, z_2, as independent variable, rather than the other way around. It is sometimes very difficult to solve for the z_2 coordinates as a function of the z_1 coordinates.

Direct Solution of Laplace's Equation in Two Dimensions by the Method of Harmonics.

Laplace's equation possesses a set of separable solutions in certain coordinate systems. By separable we mean that the potential can be expressed in terms of the product of functions of each of the coordinates separately. Potential solutions expressed in such separable coordinate systems are of course particularly useful in case the geometrical boundaries of the problem being considered coincide with one of the coordinate surfaces in that coordinate system. Let us consider this method in two dimensions. Laplace's Equation (1.17) expressed in plane polar coordinates is:

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

To achieve separation, we let:

$$(5.35)\quad \phi = R(r)\,\Theta(\theta)$$

Upon substitution of (5.35) into (5.34) and division by ϕ we get:

$$(5.36)\quad \frac{r}{R}\frac{\partial}{\partial r}\left(r\frac{\partial R}{\partial r}\right)+\frac{1}{\Theta}\frac{\partial^2 \Theta}{\partial \theta^2}=0$$

Since Equation (5.36) is separated into two terms which are respectively functions of r and θ only and therefore must be individually constant, we may arbitrarily set the first term of (5.36) equal to k_m^2. This gives:

$$(5.37)\quad r\frac{\partial}{\partial r}\left(r\frac{\partial R}{\partial r}\right)-k_m^2 R=0,\quad \frac{\partial^2 \Theta}{\partial \theta^2}+k_m^2\,\Theta=0$$

Note that separation of the differential equation in general results in a set of solutions that are characterized by a parameter k_m. These solutions are obtained by integrating Equations (5.37), giving:

$$(5.38)\quad k_m \neq 0,\ R_m = C_m r^{k_m} + D_m r^{-k_m},\ \Theta_m = A_m \cos k_m\theta + B_m \sin k_m\theta$$

$$k_m = 0,\ R_0 = G + H \ln r,\ \Theta_0 = E + F\theta$$

and hence a general solution.

$$(5.39)\quad \phi = \sum_{m=0}^{\infty} R_m \Theta_m = \sum_{m=1}^{\infty} (A_m \cos k_m\theta + B_m \sin k_m\theta)(C_m r^{k_m} + D_m r^{-k_m}) + (E + F\,\theta)(G + H \ln r)$$

is obtained from a linear superposition of the individual solutions of Equations (5.38).

That Equation (5.39) is actually a general solution and that we have generated a complete orthogonal set of solutions, subject to certain conditions, can be shown by a more general analysis of the series given in Equation (5.39). In order to apply a solution of the type of Equation (5.39) to the solution of a practical problem we must first express certain already known potential problem solutions in the same form as (5.39), and then attempt to superimpose on such solutions additional potentials, with undetermined coefficients, of the general form of (5.39). Then the coefficients are to be determined by the use of the boundary conditions of the given problem.

Let us illustrate this general method by considering the problem of a line charge located at a distance r_o from the axis of a dielectric cylinder of radius a and specific inductive capacity k, as seen in Figure (5.8B). Let us first express the logarithmic potential of the line charge alone, as seen in Figure (5.8A), in the same form as Equation (5.39). This amounts to shifting the origin of the logarithmic potential. Since (5.39) represents the general solution of Laplace's equation in plane polar coordinates and is in general non-singular, except at r = 0, the potential of a line charge from (5.10), with the arbitrary

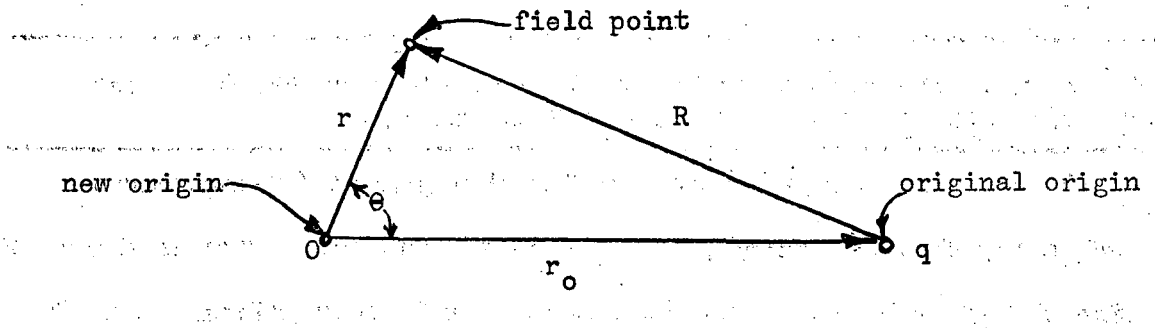


Figure (5.8A)

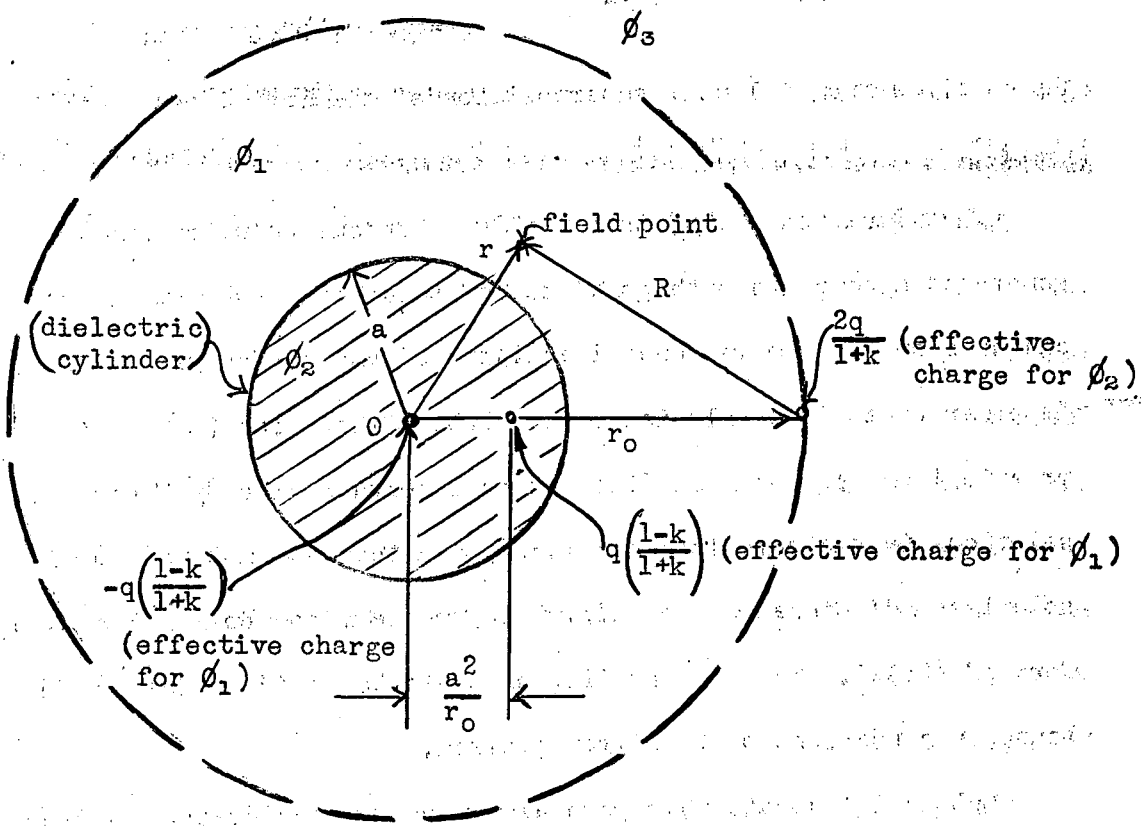


Figure (5.8B)

potential base, $\frac{q}{2\pi k_o} \ln r_o$, omitted:

$$(5.40)^* \quad \phi = -\frac{q}{2\pi k_o} \ln R$$

cannot in general be expressed by a single expansion of the type in (5.39), but must be expressed in terms of two different solutions, one valid in the region where $r < r_o$ and one valid in the region where $r > r_o$.

---

* Note that the r_o in the omitted arbitrary potential base of (5.10) is not necessarily the same r_o as that in Figures (5.8A and B).

These two different expansions of ϕ for the line charge must fit together at $r = r_o$ in such a way that the derivative shall be discontinuous only at the point where the line charge is located, but continuous at all other points. The discontinuity is such that the total flux emerging from that point corresponds to

the value of the line charge per unit length.

The logarithmic potential (5.40) of a line charge at the origin can be put into the form of (5.39) of an isolated line charge located at $\theta = 0$ and $r = r_o$, as in Figure (5.8A), by expressing the radial distance R by the law of cosines, $R = (r^2 + r_o^2 - 2rr_o\cos\theta)^{1/2}$, and then expanding in a power series in r/r_o, for use when $r < r_o$, and in r_o/r, for use when $r > r_o$. This process will generate a Fourier-Laurent expansion of the logarithmic potential for a line charge not located at the origin. The series will be convergent within their respective ranges of validity. Thus for the potential due to a line charge only we have from the expansion of (5.40):

$$(5.41)\quad \phi_{o<r<r_o} = \frac{q}{2\pi k_o}\left\{\sum_{m=1}^{\infty}\frac{1}{m}\left(\frac{r}{r_o}\right)^m \cos m\theta - \ln r_o\right\}$$

$$\phi_{r_o<r<\infty} = \frac{q}{2\pi k_o}\left\{\sum_{m=1}^{\infty}\frac{1}{m}\left(\frac{r_o}{r}\right)^m \cos m\theta - \ln r\right\}$$

We will choose the origin of our polar coordinate system, for the potential of the line charge and dielectric cylinder combined, at the center of the dielectric cylinder with the radius vector corresponding to $\theta = 0$ passing through the charge. Thus the coordinate system will agree with that of Equations (5.41) and will be as shown in Figure (5.8B). To satisfy the boundary condition at the surface of the cylinder $r = a$, we shall consider the line charge solution (5.41) valid for $0 < r < r_o$, and superimpose on it a general solution of the type of (5.39), with $k_m = n$ and, with undetermined coefficients A_n, B_n, E and F, to account for the effect of the polarization of the dielectric cylinder, and make a separation of the potential into two parts ϕ_1 and ϕ_2 to be valid outside and inside the cylinder respectively. This separation is made in order to assure a finite value for the potential at the origin and convergence of the second series in each expression:

$$(5.42)\quad \underset{a<r<r_o}{\phi_1} = \frac{q}{2\pi k_o}\left\{\sum_{m=1}^{\infty}\frac{1}{m}\left(\frac{r}{r_o}\right)^m \cos n\theta - \ln r_o\right\} + \sum_{m=1}^{\infty} B_m r^{-m}\cos m\theta + F$$

$$\underset{o<r<a}{\phi_2} = \frac{q}{2\pi k_o}\left\{\sum_{m=1}^{\infty}\frac{1}{m}\left(\frac{r}{r_o}\right)^m \cos n\theta - \ln r_o\right\} + \sum_{m=1}^{\infty} A_m r^m \cos m\theta + E$$

Since the effect of the induced polarization charges in the cylinder will be non-singular, both at the origin and at infinity, the logarithmic terms of (5.39) have been omitted and only the negative powers of r have been used in the solution ϕ_1 valid outside the cylinder, and only the positive powers of r have been used in the solution ϕ_2 valid inside the cylinder. Using the boundary conditions (3.16) and (3.23) at r = a:

$$(5.43)\quad \phi_1 = \phi_2, \quad \frac{\partial \phi_1}{\partial r} = k \frac{\partial \phi_2}{\partial r}$$

We can evaluate the coefficients A_n, B_n, E and F, by substituting (5.42) into (5.43) and then equating the coefficients of equal order in θ, term by term, to zero. This procedure is justified since these Fourier series form a complete orthogonal set. The resultant solution is:

$$(5.44)\quad \underset{a<r<r_o}{\phi_1} = \frac{q}{2\pi k_o}\left\{\sum_{m=1}^{\infty}\frac{1}{m}\left[\left(\frac{r}{r_o}\right)^m + \left(\frac{1-k}{1+k}\right)\left(\frac{a^2}{r_o}\right)^m \frac{1}{r^m}\right]\cos m\theta - \ln r_o\right\}$$

$$\underset{o<r<a}{\phi_2} = \frac{q}{\pi k_o(1+k)}\sum_{m=1}^{\infty}\frac{1}{m}\left(\frac{r}{r_o}\right)^m \cos m\theta - \frac{q}{2\pi k_o}\ln r_o$$

The potential ϕ_1 outside the cylinder is seen to correspond to an effective line charge arrangement, with the role of the dielectric cylinder taken by two effective line charges, and the cylinder absent; these consist of an effective charge $-q\left(\frac{1-k}{1+k}\right)$ located at the origin, and an effective charge $q\left(\frac{1-k}{1+k}\right)$ located at the inversion point of the actual external line charge, and the actual charge. The inversion point lies on the vector $\vec{r}_o$ at a distance a^2/r_o from the origin. On the other hand the potential ϕ_2 inside the cylinder is seen to correspond to an effective line charge arrangement, with the dielectric cylinder absent, of one effective charge placed at the position of the actual charge but of strength, $\frac{2q}{1+k}$. Therefore this problem could have been solved by the method of images mentioned in Chapter 4. This fact can be verified directly by the use of the logarithmic potentials. Note that the corresponding three-dimensional problem of a point charge and a dielectric

sphere does not have a solution by the method of images.

Let us take another example of the solution of a problem in terms of cylindrical harmonics. Consider a wedge-shaped region bouded by grounded conducting surfaces intersecting at the origin with an interior angle α, as in Figure (5.9). And consider a line charge of strength q per unit length located at the point (r_o, β) within the wedge. The solution of this problem will give the Green's function

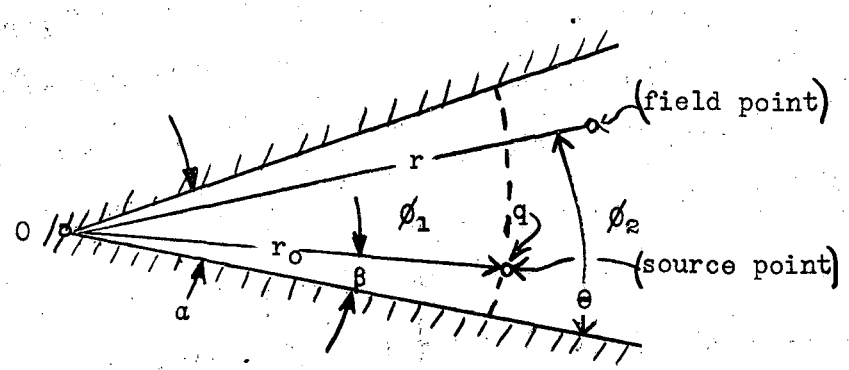


Figure (5.9)

for the region bounded by the intersecting conducting planes. It is again clear that we cannot hope to express a solution by means of a single equation valid throughout the region from r = 0 to $r = \infty$ since the derivative of the potential will be discontinuous at the point occupied by the line charge. Thus we must again construct the solution out of two solutions, one valid in the region $r < r_o$ and the other valid in the region $r > r_o$. We shall join these potential expressions to each other on the cylindrical surface $r = r_o$ by the flux condition corresponding to the charge q.

Since the potential must vanish on the boundary where $\theta = 0$ and where $\theta = \alpha$ the angular part of the solution must be of the form $\sin\left(\frac{m\pi\theta}{\alpha}\right)$. Thus in (5.39) $k_m = m\pi/\alpha$. In order to fulfill Laplace's Equation we set $A = E = F = G = H = 0$ in Equation (5.39) and have for the potentials everywhere inside the wedge:

$$(5.45)\quad \underset{r<r_o}{\phi_1} = \sum_{m=1}^{\infty} C_m \left(\frac{r}{r_o}\right)^{\frac{m\pi}{\alpha}} \sin\frac{m\pi}{\alpha}\theta$$

$$\underset{r>r_o}{\phi_2} = \sum_{m=1}^{\infty} D_m \left(\frac{r}{r_o}\right)^{\frac{m\pi}{\alpha}} \sin\frac{m\pi}{\alpha}\theta$$

The coefficients C_n and D_n must be equal in order to assure continuity of the potentials across the cylindrical surface $r = r_o$. Let us evaluate C_n by integrating the total flux from q over a small surface composed of two cylinders, one of radius slightly larger and the other of radius slightly smaller than r_o, and two radial planes closing the ends of the cylindrical arcs to make the infinitesimal surface completely surround the line charge, as seen in Figure (5.10). The cylindrical surfaces are to be larger than the plane surfaces by an order of magnitude so that the flux across the plane surfaces may be neglected relative to that across the cylindrical surfaces. We can represent the line charge q by an equivalent charge density σ within S:

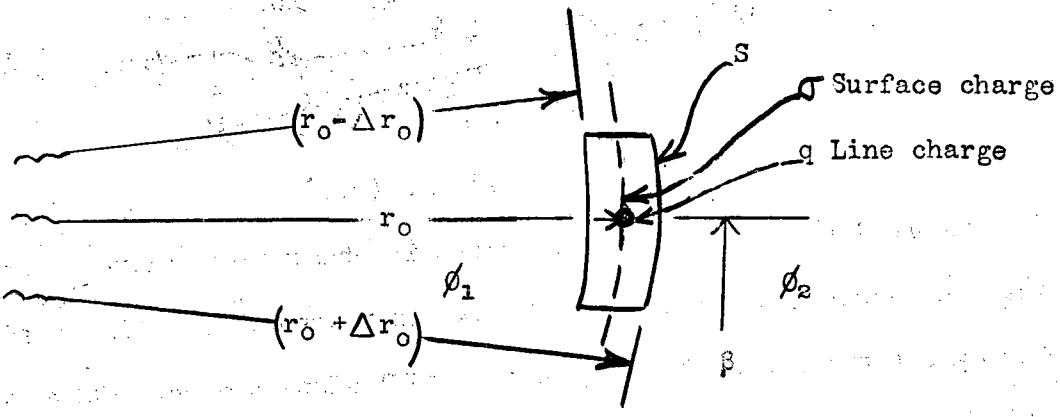


Figure (5.10)

From Equation (3.16) we have:

$$(5.46)\quad \sigma = -k_o\left(\frac{\partial \phi_2}{\partial r} - \frac{\partial \phi_1}{\partial r}\right)$$

Differentiating (5.45) we obtain:

$$(5.47)\quad \frac{\partial \phi_1}{\partial r} = \sum_{m=1}^{\infty} C_m \frac{m\pi}{a} \frac{1}{r_o} \left(\frac{r}{r_o}\right)^{\left(\frac{m\pi}{a}-1\right)} \sin \frac{m\pi}{a}\theta$$

$$\frac{\partial \phi_2}{\partial r} = -\sum_{m=1}^{\infty} C_m \frac{m\pi}{a} \frac{1}{r_o} \left(\frac{r_o}{r}\right)^{\left(\frac{m\pi}{a}+1\right)} \sin \frac{m\pi}{a}\theta$$

The total surface charge within the surfaces is equal to the line charge q. The surface charge density distribution can be expressed by $q\,\delta(\theta - \beta)$ where $\delta(\theta - \beta)$ is a Dirac δ function defined to be equal to zero at points where

$\theta \neq \beta$ and to be infinite at points where $\theta = \beta$, in such a way that:

$$r_o \int_0^\alpha \delta(\theta - \beta)d\theta = 1 \qquad 0 < \beta < \alpha \tag{5.48}$$

If we substitute $\sigma = q\delta(\theta - \beta)$ into (5.46) and also substitute the difference between the potential derivatives from (5.47) into (5.46) we can determine the coefficients C_n by multiplying both sides of the resulting equation by

$$\sin\frac{m\pi\theta}{\alpha}\, d\theta$$

and integrating from 0 to α. The orthogonality of the sine functions in the region in question will cause all terms except the term where n = m to vanish. The integral over the δ function can be evaluated immediately, since the sine function varies slowly relative to the variation of the δ function and the sine factor can be taken out of the integral when it is evaluated at $\theta = \beta$. The equation, characteristic of the δ function, expressing this mathematically, is:

$$\int_a^b \delta(x - c)f(x)dx = f(c)/r_o \qquad a < c < b \tag{5.49}$$

We thus obtain, noting that $r = r_o$:

$$q\int_0^\alpha \delta(\theta - \beta)\sin\frac{m\pi\theta}{\alpha}\, d\theta = 2k_o C_m \frac{m\pi}{\alpha}\,\frac{1}{r_o}\int_0^\alpha \sin^2\left(\frac{m\pi\theta}{\alpha}\right)d\theta \tag{5.50}$$

Integrating and replacing m by n.

$$C_n = \frac{q}{n\pi k_o}\sin\left(\frac{n\pi\beta}{\alpha}\right) \tag{5.51}$$

and hence the complete solution is:

$$\underset{r<r_o}{\phi_1} = \frac{q}{\pi k_o}\sum_{n=1}^{\infty}\frac{1}{n}\left(\frac{r}{r_o}\right)^{\frac{n\pi}{\alpha}}\sin\frac{n\pi\beta}{\alpha}\sin\frac{n\pi\theta}{\alpha} \tag{5.52}$$

$$\underset{r<r_o}{\phi_2} = \frac{q}{\pi k_o}\sum_{n=1}^{\infty}\frac{1}{n}\left(\frac{r}{r_o}\right)^{\frac{n\pi}{\alpha}}\sin\frac{n\pi\beta}{\alpha}\sin\frac{n\pi\theta}{\alpha}$$

This is the desired Green's function, when q is set equal to unity. This method is a general one which is useful for deriving the Green's function within any set of boundaries corresponding to equi-coordinate planes in a particular coordinate system.

CHAPTER 6 THREE DIMENSIONAL POTENTIAL PROBLEMS

The Solution of Laplace's Equation in Spherical Coordinates.

Let us now consider the solution of some examples of three dimensional potential problems where Laplace's equation is separable. Laplace's equation, expressed in spherical polar coordinates (r, θ, ϕ), is given by:

$$(6.1)\quad \nabla^2\Phi = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial \Phi}{\partial r}\right) + \frac{1}{r^2\sin\theta}\frac{\partial}{\partial \theta}\left(\sin\theta\frac{\partial \Phi}{\partial \theta}\right) + \frac{1}{r^2\sin^2\theta}\frac{\partial^2 \Phi}{\partial \phi^2} = 0$$

In order to achieve the separation of (6.1) let us put:

$$(6.2)\quad \Phi = R(r)Y(\theta,\phi)$$

where $Y(\theta,\phi)$ is known as a spherical harmonic. Substituting (6.2) into equation (6.1) and letting:

$$(6.3)\quad n(n+1) = \frac{1}{R}\frac{\partial}{\partial r}\left(r^2\frac{\partial R}{\partial r}\right)$$

we obtain the separated equations:

$$(6.4)\quad \frac{\partial}{\partial r}\left(r^2\frac{\partial R}{\partial r}\right) - n(n+1)\,R = 0$$

$$(6.5)\quad \frac{\partial}{\partial \theta}\left(\sin\theta\frac{\partial Y}{\partial \theta}\right) + \frac{1}{\sin\theta}\frac{\partial^2 Y}{\partial \phi^2} + n(n+1)\sin\theta Y = 0$$

Again as in Chapter 5 we have introduced a separation constant n. This has been possible since the differential operation of ∇^2 on Φ results in two terms which are respectively either a function only of r or a function of θ and ϕ, and therefore must be individually constant. The differential equation for the radial part of Φ, (6.4), has the solution:

$$(6.6)\quad R(r) = A_n r^n + B_n r^{-n-1}$$

The set of functions $Y_n(\theta,\phi)$ which are the solutions of (6.5) have orthogonality properties that are similar to the orthogonality properties of the Fourier series, which we discussed in the chapter on two dimensional solutions. This may be

---

In Chapter 6 there are three similar symbols: Φ is the potential; Φ is the function used to separate variables in $\Phi = R(r)\,\Theta(\theta)\,\Phi(\phi)$; and ϕ is the angular coordinate.

demonstrated directly by the use of Green's theorem:

$$(1.18)\quad \iiint_V (\phi\nabla^2\psi - \psi\nabla^2\phi)dv = \iint_S (\phi\vec{\nabla}\psi - \psi\vec{\nabla}\phi)\cdot d\vec{S}$$

Let S be the surface of a sphere.

If we put $\phi_1 = R_1Y_1$ and $\phi_2 = R_2Y_2$ for the two potentials ϕ and ψ, appearing in (1.18) we obtain:

$$(6.7)\quad \iiint_V (R_1Y_1\nabla^2R_2Y_2 - R_2Y_2\nabla^2R_1Y_1)dv = \iint_S (R_1Y_1\vec{\nabla}R_2Y_2 - R_2Y_2\vec{\nabla}R_1Y_1)\cdot d\vec{S}$$

The left side vanishes since ϕ_1 and ϕ_2 are solutions of Laplace's equation.

Carrying out the indicated differentiation and noting that the component of the $\vec{\nabla}$ operator that is parallel to $\vec{S}$ does not operate on the function Y, we have:

$$(6.8)\quad \left(\frac{R'_2}{R_2} - \frac{R'_1}{R_1}\right)\iint Y_1Y_2\,dS = 0$$

If the two radial functions correspond to different values of the separation constant n, they will have a different dependence on r and thus their logarithmic derivatives will be unequal and therefore the left term will vanish only if the integral vanishes. If this is so, then the two spherical surface harmonics Y_1 and Y_2 must be orthogonal to each other when integrated over an element of spherical surface, or an element of solid angle. This proof is independent of the particular nature of the coordinate system used, as long as Laplace's equation is separable in this coordinate system.* We can therefore conclude that in general, orthogonal functions are generated in the solution of Laplace's equation.

The spherical surface harmonics can be further separated by letting:

$$(6.9)\quad Y(\theta,\phi) = \Theta(\theta)\,\Phi(\phi)$$

This results in the following two equations, when m is introduced as a separation parameter:

$$(6.10)\quad \frac{d}{d\mu}\left[(1-\mu^2)\frac{d\Theta}{d\mu}\right] + \left[n(n+1) - \frac{m^2}{(1-\mu^2)}\right]\Theta = 0 \quad ; \quad \mu = \cos\theta$$

$$(6.11)\quad \frac{d^2\Phi}{d\phi^2} + m^2\Phi = 0$$

*The proof is general if $\vec{S}$ is taken as one of the coordinate planes.

These equations are solved by:

$$(6.12)\quad \Theta = C_n P_n^m(\mu) + D_n Q_n^m(\mu)$$

$$\Phi = E_m \cos m\phi + F_m \sin m\phi \qquad \text{when } m \neq 0$$

$$(6.13)\quad \Phi = G\phi + H \qquad \text{when } m = 0$$

The functions $P_n^m(\cos\theta)$ and $Q_n^m(\cos\theta)$ are the associate Legendre functions of the first and second kind respectively. Their mathematical properties can be found in numerous references.

The Potential of a Point Charge.

Let us now consider the application of these solutions of Laplace's equation to problems whose geometry has azimuthal symmetry, that is problems for which $m = 0$. The surface harmonic $Y(\theta,\phi)$ then has the form:

$$(6.14)\quad Y(\theta,\phi) = \Theta(\theta) = C_n P_n(\mu) + D_n Q_n(\mu)$$

We may obtain the potential of a point charge, expressed in terms of a series expansion in the radial and angular functions obtained in the above separation of the coordinates, from the Coulomb potential (1.24), by expanding the cosine law expression for $1/R$:

$$(6.15)\quad \frac{1}{R} = \frac{1}{r_o}\left[\left(\frac{r}{r_o}\right)^2 + 1 - 2\frac{r}{r_o}\cos\theta\right]^{-1/2} = \frac{1}{r}\left[\left(\frac{r_o}{r}\right)^2 + 1 - 2\frac{r_o}{r}\cos\theta\right]^{-1/2}$$

in powers of r/r_o and r_o/r.

$$(6.16)\quad \frac{1}{R} = \frac{1}{r_o}\sum_{n=0}^{\infty}\left(\frac{r}{r_o}\right)^n P_n(\mu) = \frac{1}{r}\sum_{n=0}^{\infty}\left(\frac{r_o}{r}\right)^n P_n(\mu)$$

For the same reasons as were used in the solution of the wedge problem in Chapter 5, we must use two potentials, one valid in the region were $r < r_o$ and one valid in the region where $r > r_o$, respectively. The physical arrangement is shown in Figure (6.1). The two potentials are:

$$(6.16)\quad \underset{r<r_o}{\phi_1} = \frac{q}{4\pi k_o r_o}\sum_{n=0}^{\infty}\left(\frac{r}{r_o}\right)^n P_n(\mu)$$

$$\underset{r>r_o}{\phi_2} = \frac{q}{4\pi k_o r}\sum_{n=0}^{\infty}\left(\frac{r_o}{r}\right)^n P_n(\mu)$$

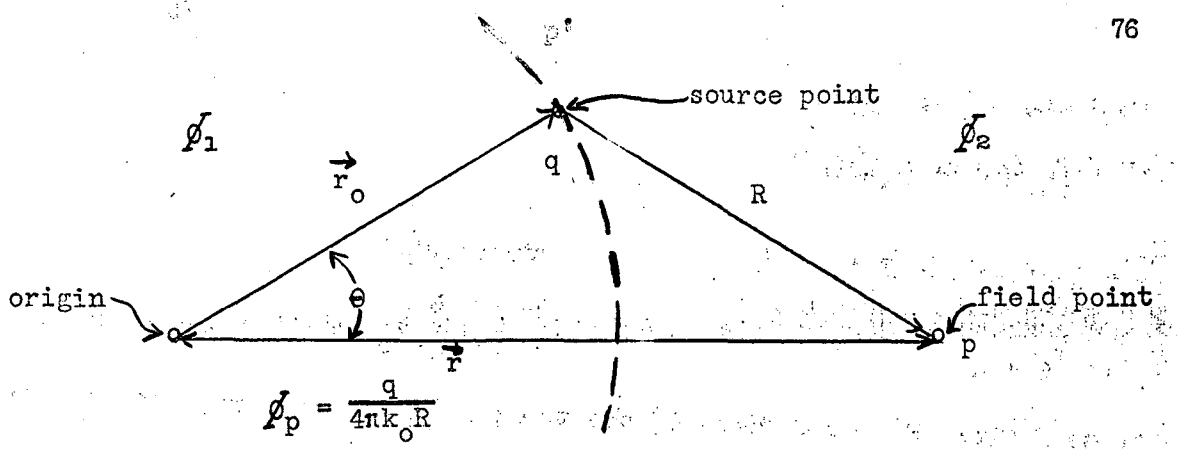


Figure (6.1)

The resulting potential of the point charge is therefore a Taylor-Laurent series in r and a series of Legendre polynomials in cos θ.

The Potential of a Dielectric Sphere and a Point Charge.

The set of point solutions derived above can be used to generate the solutions of problems involving a number of point charges and having boundaries which possess spherical symmetry. If we consider for example the simple problem of a point charge and a dielectric sphere of radius a, as shown in Figure (6.2), with r_o being the distance from the center of the sphere to the point charge. We will

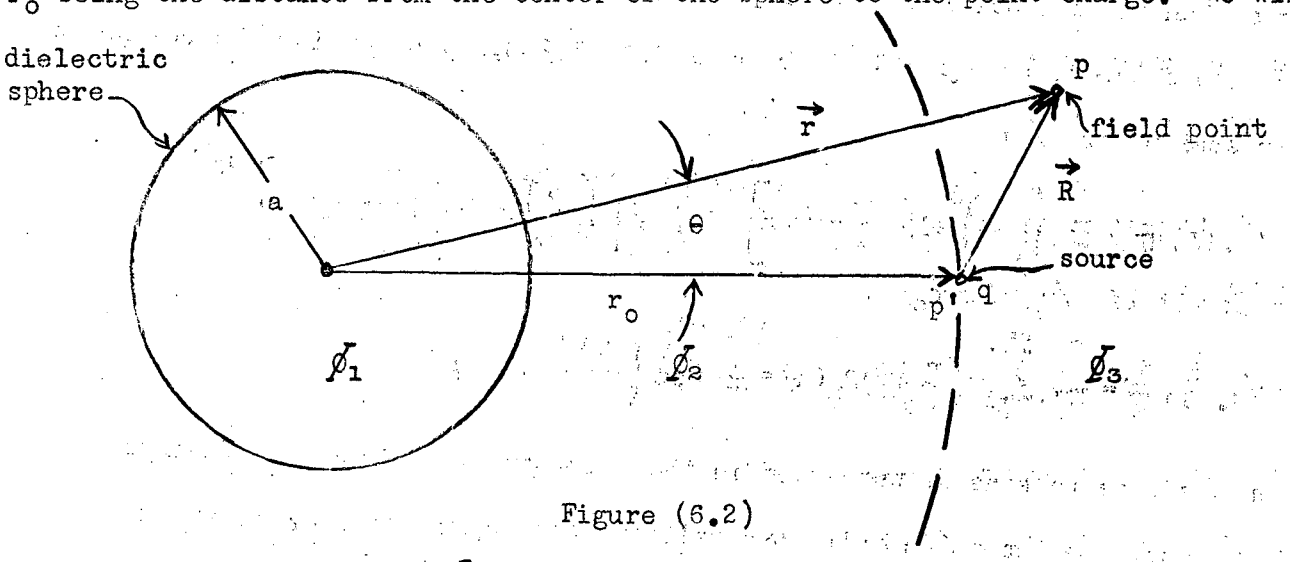


Figure (6.2)

need three expressions for ϕ which are valid in the following ranges of r:

(6.17) $o<r<a$ $\quad \phi_1 = \sum_{n=0}^{\infty} A_n r^n P_n(\mu)$

$a<r<r_o$ $\quad \phi_2 = \frac{q}{4\pi k_o r_o} \sum_{n=0}^{\infty} \left(\frac{r}{r_o}\right)^n P_n(\mu) + \sum_{n=0}^{\infty} B_n r^{-n-1} P_n(\mu)$

$r_o<r<\infty$ $\quad \phi_3 = \frac{q}{4\pi k_o r} \sum_{n=0}^{\infty} \left(\frac{r_o}{r}\right)^n P_n(\mu) + \sum_{n=0}^{\infty} B_n r^{-n-1} P_n(\mu)$

The fit of ϕ_1 to ϕ_2 at r = a can be carried out in the same way in which it was in the two dimensional case. The fit of ϕ_2 to ϕ_3 at $r_o = r$ is inherent from the nature of the solutions in Equation (6.16). The fact that the boundary conditions (3.16) and (3.20) must be fulfilled for all values of the angle θ and the fact that the angular functions are orthogonal, makes it possible to equate the terms of the series separately and this equating of ϕ_1 to ϕ_2 will determine the coefficients A_n and B_n. The resulting solution is identical to the solution obtained by the inversion process outlined in Chapter 4, for $k = \infty$.

The Potential of a Dielectric Sphere in a Uniform Field.

As a second example of a problem with spherical symmetry, consider a dielectric sphere, of specific inductive capacity k, in a uniform field whose force lines are parallel to the x axis as shown in Figure (6.3). The lines of electric displacement $\vec{D}$ are shown.

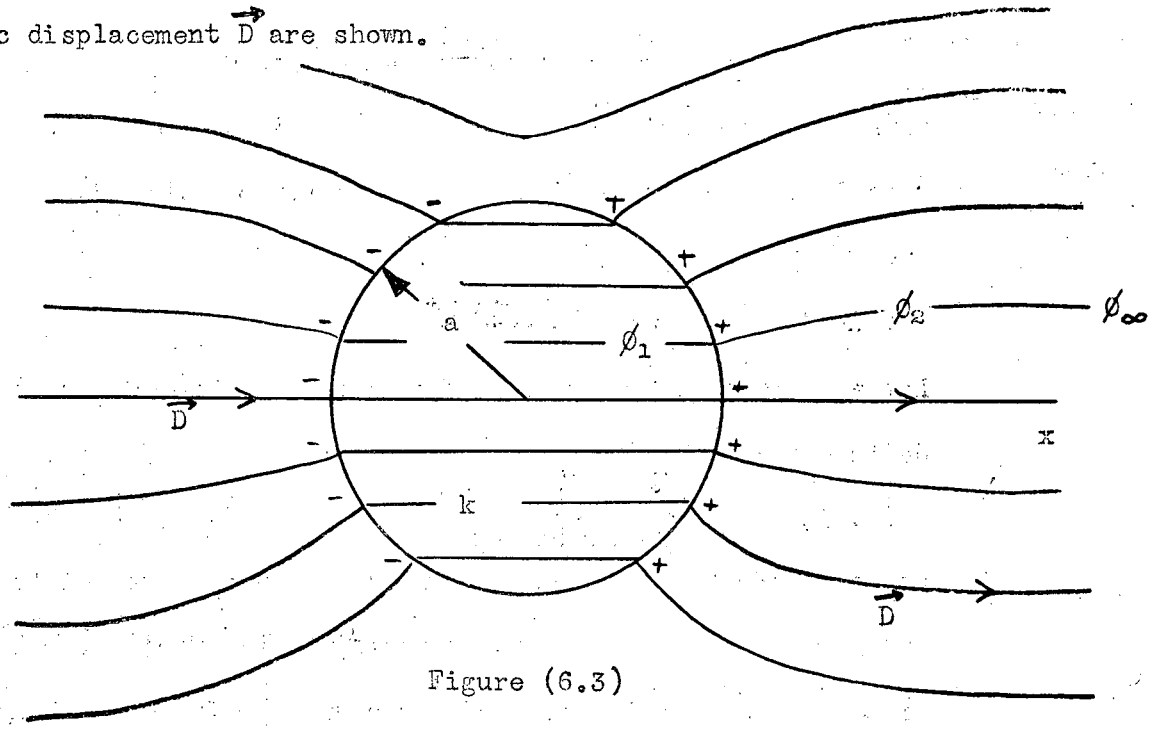


Figure (6.3)

The potential at infinity being uniform is given by:

(6.18) $\phi_\infty = -E_o x = -E_o r \cos\theta = -E_o r\,\mu = -E_o r P_1(\mu)$

From (6.6) and (6.12) we have, for the potentials inside and outside the sphere,

by inspection, the following expressions:

$$(6.19)\quad \phi_1 = \sum_{n=0}^{\infty} A_n r^n P_n(\mu)$$

$$\phi_2 = \sum_{n=0}^{\infty} B_n r^{-n-1} P_n(\mu) - E_o r \cos\theta \qquad D_n = E_m = F_m = G = H = 0$$

The boundary conditions, $\phi_1 = \phi_2$ and $k\left(\frac{\partial \phi_1}{\partial r}\right) = \left(\frac{\partial \phi_2}{\partial r}\right)$ at $r = a$, must hold for all values of the angle θ. We therefore evaluate the constants A_n and B_n by equating the coefficients of equal powers of $\cos\theta$ in the expansions and find that:

$$(6.20)\quad \begin{cases} A_o = B_o = 0 \\ A_n = B_n = 0 \text{ for } n > 1 \end{cases} \qquad \begin{cases} A_1 = \frac{-3E_o}{k+2} \\ B_1 = \frac{(k-1)E_o a^3}{k+2} \end{cases}$$

This gives for the potentials:

$$(6.21)\quad \phi_1 = \frac{-3E_o r}{k+2} \cos\theta$$

$$\phi_2 = \frac{(k-1)}{(k+2)} \frac{E_o a^3 \cos\theta}{r^2} - E_o r \cos\theta$$

Note that the field $\vec{E}$ inside the sphere is uniform, but is smaller than the field outside the sphere at infinity by the ratio $3/(k+2)$. Also the induced field of the sphere in the region outside the sphere is that of a dipole whose moment is:

$$(6.22)\quad \vec{p} = 4\pi k_o a^3 \left(\frac{k-1}{k+1}\right) \vec{E}_o$$

Let a quantity L be known as the depolarization factor for a dielectric body, defined as:

$$(6.23)\quad L = \frac{|\vec{E}| - |\vec{E} \text{ inside}|}{|k_o| |\vec{P} \text{ inside}|} \qquad (3.12)\quad \vec{P} = k_o(k-1)\vec{E}$$

For a sphere $L = 1/3$, for a thin rod oriented parallel to the field, $L \simeq 0$ and for a thin disk oriented normally to the field, $L = 1$. Thus the electric field within a dielectric body in a uniform field is always smaller than the field at a large distance, while the dielectric displacement is always larger.

The Potential of a Spherical Arbitrary Potential Distribution.

As a third example, let us consider a spherical surface, of radius a, over which the potential distribution is a given function, $\phi(a, \theta)$, of the angle θ.

We will have two potentials from Equations (6.6) and (6.12), one valid inside and one valid outside the surface:

$$(6.24)\quad \underset{r<a}{\phi} = \sum_{n=0}^{\infty} A_n r^n P_n(\mu), \qquad B_n = 0 \quad (D_n=E_m=F_m=G=H=0)$$

$$\underset{r>a}{\phi} = \sum_{n=0}^{\infty} B_n r^{-n-1} P_n(\mu), \qquad A_n = 0$$

The constants A_n and B_n may be determined by equating the expressions in (6.24), with r = a, to $\phi(a,\theta)$, and then multiplying the resulting equality by $P_m(\mu)$ and taking advantage of the orthogonality condition:

$$(6.25)\quad \int_{-1}^{+1} P_n(\mu)\, P_m(\mu)\, d\mu = \frac{2}{2n+1}\,\delta_{mn}$$

we have:

$$(6.26)\quad A_n = \frac{2n+1}{2a^n}\int_{-1}^{+1} \phi(a,\theta) P_n(\mu)\, d\mu \qquad B_n = a^{2n+1} A_n$$

So the potentials become:

$$(6.27)\quad \underset{r<a}{\phi} = \sum_{n=0}^{\infty} \frac{2n+1}{2a^n}\, r^n P_n(\mu) \int_{-1}^{+1} \phi(a,\theta') P_n(\mu')\, d\mu'$$

$$\underset{r>a}{\phi} = \sum_{n=0}^{\infty} \frac{2n+1}{2}\, a^{n+1} r^{-n-1} P_n(\mu) \int_{-1}^{+1} \phi(a,\theta') P_n(\mu')\, d\mu'$$

The Potential of a Charged Ring.

As a fifth example, let us consider the potential of a charged ring, of total charge q, possessing a half angle θ_o at the origin, and located at a distance r_o from the origin, as seen in Figure (6.4). The potential along the

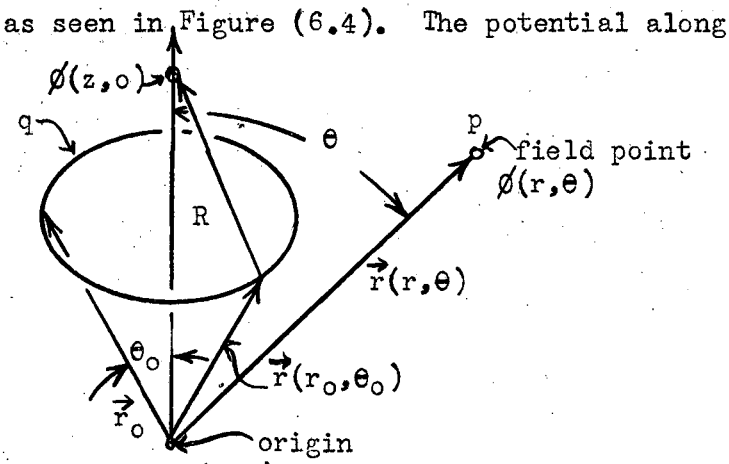


Figure (6.4)

z axis, the axis of symmetry, is found by expanding 1/R in the Coulomb potential, $\phi = q/4\pi k_o R$, by Equation (6.16):

$$(6.31)\qquad \underset{r<r_o}{\phi}(z,o) = \frac{q}{4\pi k_o r_o}\sum_{n=0}^{\infty}\left(\frac{z}{r_o}\right)^n P_n(\cos\theta_o)$$

$$\underset{r>r_o}{\phi}(z,o) = \frac{q}{4\pi k_o r_o}\sum_{n=0}^{\infty}\left(\frac{r_o}{z}\right)^{n+1} P_n(\cos\theta_o)$$

The potential at a general point, not lying on the z axis, may be found by multiplying the n^{th} term in the series by $P_n(\mu)$ and writing r for z:

$$(6.32)\qquad \underset{r<r_o}{\phi}(r,\theta) = \frac{q}{4\pi k_o r_o}\sum_{n=0}^{\infty}\left(\frac{r}{r_o}\right)^n P_n(\cos\theta_o)P_n(\cos\theta)$$

$$\underset{r>r_o}{\phi}(r,\theta) = \frac{q}{4\pi k_o r_o}\sum_{n=0}^{\infty}\left(\frac{r_o}{r}\right)^{n+1} P_n(\cos\theta_o)P_n(\cos\theta)$$

The uniqueness theorem is essential to justify the argument that led to the above result.

We shall not discuss problems that involve the associated Legendre functions $P_n^m(\mu)$ which appear in the potential expressions in cases of azimuthal asymmetry, and we shall also not discuss problems that involve the Legendre functions of the second kind $Q_n^m(\mu)$ which are singular at $\mu = \pm 1$, that is along the polar axis. The Legendre functions of the second kind are used in problems such as those involving conical boundaries, where the z axis is excluded from the range of validity of the potentials.

The Solution of Laplace's Equation in Cylindrical Coordinates.

Let us now consider the separation of Laplace's equation in cylindrical coordinates. Laplace's equation in (r,ϕ,z) is:

$$(6.33)\qquad \frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial\phi}{\partial r}\right) + \frac{1}{r^2}\frac{\partial^2\phi}{\partial\phi^2} + \frac{\partial^2\phi}{\partial z^2} = 0$$

Separating by means of the product functions:

$$(6.34)\quad \phi(r,\phi,z) = R(r)\,\Phi(\phi)Z(z)$$

we obtain:

$$(6.35)\quad r\frac{d}{dr}\left(r\frac{dR}{dr}\right) + (k^2r^2-n^2)R = 0$$

$$\frac{d^2\Phi}{d\phi^2} + n^2\Phi = 0$$

$$\frac{d^2Z}{dz^2} - k^2Z = 0$$

where k and n are the separation parameters. The character of the solution will differ markedly whether n or k are real or imaginery. If solutions are desired which are single valued in the azimuth angle ϕ, then the solution must be periodic in ϕ and hence n must be real. If k is real, the solution along the z axis will be exponential and the radial solutions will be in terms of the Bessel's functions J_n and Y_n. The integrals are therefore of the form:

$$(6.36)\quad R(r) = A_nJ_n(kr) + B_nY_n(kr) \qquad k\neq 0$$

$$R(r) = Ar^n + Br^{-n} \qquad k=0$$

$$\Phi(\phi) = C_n\cos\phi + D_n\sin n\phi \qquad n\neq 0$$

$$\Phi(\phi) = C\phi + D \qquad n=0$$

$$Z(z) = E_ke^{kz} + F_ke^{-kz} \qquad k\neq 0$$

$$Z(z) = Ez + F \qquad k=0$$

If k and n are both zero:

$$\phi = (A\ln r + B)(C\phi + D)(Ez + F)$$

If the cylindrical solution is required to be periodic in the z direction, then k must be imaginary, and the solutions of the radial equation will be Bessel functions of an imaginary variable which are usually designated by I_n and K_n.

Let us illustrate the use of these functions by one example. Let us consider a problem possessing cylindrical symmetry which is uniform radially and which has azimuthal symmetry. Thus m = 0. In general such a problem can be solved by the use of integrals in place of a summation. The integral will have the form:

$$(6.37)\quad \int_0^\infty e^{\pm kz} f(k) J_0(kr)\,dk$$

Instead of determining the set of coefficients in the sum, we must determine the value of the function f(k). The potential of a point charge may be expressed by this integral and is given with the aid of the identity:

$$(6.38)\quad \frac{1}{R} = \frac{1}{\sqrt{r^2+z^2}} = \int_0^\infty e^{\pm kz} J_0(kr)\,dk$$

(Where the + sign is used for $z<0$ and the - sign for $z>0$.) The Coulomb potential $\phi = q/4\pi k_0 R$ is:

$$(6.39)\quad \phi = \frac{q}{4\pi k_0} \int_0^\infty e^{\pm kz} J_0(kr)\,dk$$

The potential of (6.39) can then be used in combination with the induced potential of the form of (6.38) to form the solution of a problem corresponding to plane boundaries normal to the z axis and under the influence of a point charge located at the origin. This layer structure, shown in Figure (6.5), composed of several layers of varying specific inductive capacities, k_1, k_2, etc., has the potentials shown in the figure. If we apply the boundary conditions at all of the interfaces and equate the functions under the integral sign, there will be a sufficient number of equations to determine the functions, and therefore the solution.

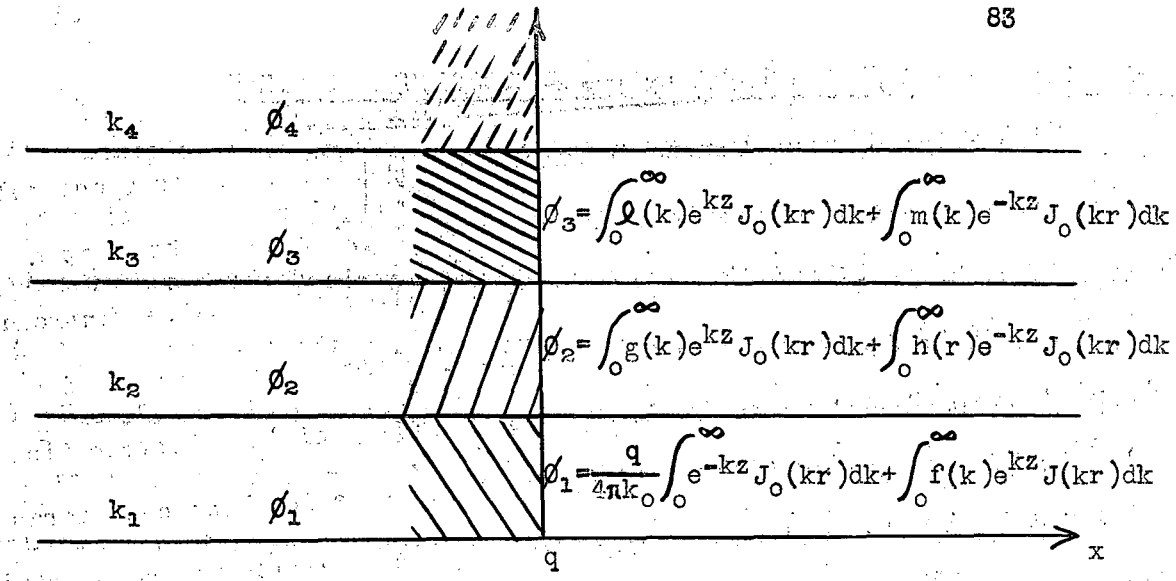


Figure (6.5)

The discussion of boundary value problems given in the last three chapters is in no way a general one and an attempt has been made to give only a few examples of the methods used in their solution.

CHAPTER 7 ENERGY RELATIONS IN THE ELECTROSTATIC FIELD

Our discussion of the electrostatic field has thus far been based entirely on a single experimental law, namely Coulomb's law (1.2) for the action at a distance force between two point charges. The electric field has been introduced as an intermediate agent whose purpose is to simplify the description of the interaction between charges. The question of the reality of the electric field as an independent physical entity therefore does not arise in these considerations. Maxwell attempted to ascribe a larger degree of physical reality to the electric field than will be necessary for our purposes. The fundamental reason for attributing a physical reality to the electric field will actually not become apparent until non-static effects are discussed. However, if we assume the reality of Maxwell's electric field, it is necessary that all of the mechanical properties, of a system which is interacting electrically, can be described either in terms of the sources which partake of the interaction or in terms of the fields themselves which are produced by the sources.

This means that the detailed nature of the sources should not influence the action of a field on a given system of charges. The description of the electric field alone must be a sufficient description to determine what interaction occurs if a number of charges are introduced at given points in the field. This interaction must be independent of the configuration of the charges which are causing the field.

Therefore it should be possible to develop a field theory in which we can describe the overall mechanical properties such as energy, equivalently in terms of the charges which are the sources of the field, or in terms of integrals over the field produced by the charges. The only criterion for the correctness of such overall relations when expressed in terms of the field theory shall be that the results are equivalent to those which are obtained from a direct consideration of the action at a distance interaction of the charges responsible for the field.

Let us consider a set of charges q_i located in free space in regions where the potentials are ϕ_i. The work done in the course of the physical assembly of these already created charges, which are initially located an infinite distance apart is given by:

$$(7.1) \quad W = 1/2 \sum_{i=1}^{n} q_i \phi_i$$

By assembling these charges we have changed the energy of the system, and since all of the forces are conservative, we can identify this expression for the work of assembly with the energy of the system. This energy must be stored somewhere. However, the location that one selects as the place of energy storage is a function of one's point of view.

For example, if we consider two masses on the end of a compressed spring, we have a system which possesses potential energy which will be released if the spring is allowed to expand. In the expansion the masses will acquire kinetic energy. The physical location of the energy in this mechanical system when it is in its initial condition is not necessarily in the spring. Phenomenologically the masses may be considered to be initially in regions of higher potential energy than they are in after the expansion of the spring. Equation (7.1) corresponds to the latter point of view. We shall now try to transform (7.1) to an expression which would make it appear as if the electrical energy resides in the so to speak "elastic" quality of the electric field, as would be required in order to correspond to the point of view that the energy of the mass-spring system resides in the spring.

The expression obtained by Maxwell for the energy in an electric field, expressed as a volume integral over the field is:

$$(7.2) \quad U = \frac{k_o}{2} \iiint E^2 \, dv$$

The integral is carried over all space. We shall now show that the field energy U is in fact the same as the assembly work W. In order to show this let us

introduce the partial fields $\vec{E}_i$, each being the Coulomb field of only one of the point charges that are responsible for the field. $\vec{E}$ and E^2 are then given by:

$$(7.3)\quad \vec{E} = \sum_{i=1}^{n} \vec{E}_i; \quad E^2 = \sum_{i=1}^{n} E_i^2 + \sum_{i=1}^{n} {\sum_{j=1}^{n}}' \vec{E}_i \cdot \vec{E}_j$$

where the prime on the summation indicates that the term for which $i = j$ has been omitted from the summation, since such terms are grouped separately in the first summation. If point charges are considered, the first term in the summation makes an infinite contribution to the integral over E^2 in (7.2). However, this infinite term is independent of the relative position of the charges and therefore it must represent the work necessary to create the charges from an arbitrary zero point of energy. Although there is no obvious reason why this energy of creation term should be infinite, the reason for its infiniteness is irrelevant since this energy does not enter into problems of electrostatic interaction. We will therefore designate:

$$(7.4)\quad U_s = \frac{k_o}{2} \iiint \sum_{i=1}^{n} E_i^2 dv$$

as the self energy of the system, and assume that this term, for reasons not contained in electrostatic theory, will ultimately be found to be finite. The Maxwell field energy expression (7.2) then becomes:

$$(7.5)\quad U = U_s + \frac{k_o}{2} \sum_{i=1}^{n} \iiint \vec{E}_i \cdot (-\vec{\nabla} {\sum_{j=1}^{n}}' \phi_j) dv$$

where ${\sum_{j}}' \phi_j$ denotes the potential at the position of the $i^{\underline{th}}$ charge due to all of the charges except the $i^{\underline{th}}$ charge itself. Using the vector expression:

$$(7.6)\quad \vec{\nabla} \cdot (\vec{A}\phi) = \phi \vec{\nabla} \cdot \vec{A} + \vec{A} \cdot \vec{\nabla} \phi$$

to perform an integration by parts, we obtain:

$$(7.7)\quad U = U_s - \frac{k_o}{2} \sum_{i=1}^{n} \iiint \left[\vec{\nabla} \cdot (\vec{E}_i {\sum_{j=1}^{n}}' \phi_j) - {\sum_{j=1}^{n}}' \phi_j \vec{\nabla} \cdot \vec{E}_i \right] dv$$

and by using Gauss' theorems (1.6 and 1.7), and noting that $\vec{\nabla}\cdot\vec{E}_i$ is zero except at the position of the i^{th} charge so that $\sum_j \phi_j$ may be removed from the integral and the divergence evaluated in terms of the source, we obtain:

$$(7.8)\quad U = U_s - \frac{k_o}{2}\sum_{i=1}^{n}\iint \vec{E}_i\left(\sum_{j=1}^{n}\phi_j\right)\cdot d\vec{S} + \frac{k_o}{2}\sum_{i=1}^{n}\phi_i^{\,o}\,\frac{q_i}{k_o}$$

where $\phi_i^{\,o} = \sum{}' \phi_j$ is the potential at the i^{th} charge due to the other charges. The surface term can be made arbitrarily small by letting the boundary surface go to infinity. This is true since the fields decrease at least as the inverse second power, and the potential at least as the inverse first power of the distance, while the differential area of integration increases only as the square of the distance. If we consider that the integral in (7.2) covers all of space where there is a field, then this integral, as a result of (7.8), reduces to:

$$(7.9)\quad U = U_s + 1/2\sum_{i=1}^{n}\phi_i^{\,o} q_i$$

The second term of this equation is identical to the expression for the work of assembly of the charges from infinity (7.1), while the first term U_s is the self energy corresponding to the energy used in the creation of the charges themselves. This analysis shows that (7.2) and (7.9) correspond to the same energy; however, (7.2) expresses the energy as a volume integral over an energy density $k_o E^2/2$ extending over all of space. No experiment can directly ascertain whether the energy resides in the field or is associated with the charge which produces the field.

In case dielectric bodies are present in the field, we shall show that Equation (7.2) becomes:

$$(7.10)\quad U = 1/2\iiint \vec{E}\cdot\vec{D}\; dv$$

In the case of continuous charges, the self energy problem disappears.

Let us consider the change of energy when a small increment of true charge $\delta\rho$ is added to the field. The work done is given by:

$$(7.11)\quad \delta W = \iiint \phi\,\delta\rho\, dv = \iiint \phi\,\delta(\vec{\nabla}\cdot\vec{D})dv = \iiint \phi(\vec{\nabla}\cdot\delta\vec{D})dv$$

Using the vector relation (7.6) and Gauss' theorem, we have:

$$(7.12)\quad \delta W = \iiint \vec{\nabla}\cdot(\delta\vec{D}\phi)dv - \iiint \delta\vec{D}\cdot\vec{\nabla}\phi dv = \iint \delta\vec{D}\phi\cdot d\vec{S} - \iiint \delta\vec{D}\cdot\vec{\nabla}\phi dv$$

Dropping the surface term, as we did in the derivation of (7.9), we obtain:

$$(7.13)\quad \delta W = -\iiint \delta\vec{D}\cdot\vec{\nabla}\phi dv = \iiint \vec{E}\cdot\delta\vec{D}\, dv$$

This increment of work usually cannot be integrated unless $\vec{E}$ is a given function of $\vec{D}$. If, for example, $\vec{E}$ and $\vec{D}$ are related by a dielectric constant, as in (3.11), which is a function of position but not of $\vec{E}$, then the energy resulting from the integration of the work increment from $\vec{D} = 0$ to $\vec{D} = \vec{D}$ gives:

$$(7.14)\quad U = \int_0^D \delta W = \int_0^D \iiint \vec{E}\cdot\delta\vec{D}\, dv = \iiint \int_0^D \frac{kk_0\,\delta(E^2)}{2}\, dv = 1/2 \iiint kk_0E^2 dv$$

$$= 1/2 \iiint \vec{E}\cdot\vec{D}\, dv$$

which is the same as we obtained before.

The assumption of a dielectric constant k that does not change with time and is only a function of position implies that the process of change of field is an isothermal process, since the dielectric constant is usually a function of the temperature. If energy enters the dielectric it may heat the dielectric and cause a variation of k with time. In order to assure isothermal behavior, the dielectric material in question must be in contact with a heat bath which can abstract heat from it to maintain a constant temperature. Thus we cannot equate the increment of work done:

$$(7.15)\quad \delta W = \iiint \vec{E}\cdot\delta\vec{D}\, dv$$

to the increase in total energy, since heat changes are also involved. The work increment done, as given by (7.12), does however represent the maximum work which can be extracted at a later time from the total electric field energy.

Thermodynamically the maximum work which can be obtained from a system under isothermal conditions is not the total energy but is the free energy F of the system. This means that in the presence of dielectrics, the expression $U=1/2\iiint \vec{E}\cdot\vec{D}\,dv$ cannot be identified with the total energy of the system, but can only be identified with the thermodynamic free energy. This distinction, of course, vanishes when no materials with temperature dependent dielectric properties are present in the field. In the thermodynamic sense the electric field $\vec{E}$ is analogous to gas pressure and the displacement $\vec{D}$ is analogous to volume.

The other thermodynamic functions can easily be derived. Since, if the total energy is U and the entropy S:

(7.16) $F = U - TS \quad ; \quad dF = dU - TdS - SdT$ T = temperature

we obtain, since $dU - TdS = \iiint \vec{E}\cdot d\vec{D}\,dv$ = increment of work at constant temperature

$$dF = \iiint \vec{E}\cdot d\vec{D}\; dv - SdT \text{ and } dF\Big|_{T\,\text{constant}} = \iiint \vec{E}\cdot d\vec{D}\; dv \text{ as before.}$$

Hence:

$$ST = -T\frac{\partial F}{\partial T}\Big|_D = \iiint \frac{k_o}{2} TE^2 \frac{dk}{dt}\,dv = \iiint \frac{\vec{E}\cdot\vec{D}}{2}\frac{T}{k}\frac{dk}{dt}\,dv$$

(7.17)
$$U = \iiint \frac{k_o}{2} E^2 \frac{d}{dT}(Tk)\,dv = \iiint \frac{\vec{E}\cdot\vec{D}}{2}\frac{1}{k}\frac{d}{dT}(Tk)\,dv$$

The heat absorbed during application of the field is thus:

(7.18) $$\delta Q = TdS = \iiint \vec{E}\cdot\delta\vec{D}\,\frac{T}{k}\left(\frac{dk}{dT}\right)dv$$

If for example, the specific inductive capacity has the form:

(7.19) $k = 1 + \chi = 1 + A/T$ A = constant

which applies to gases composed of molecules with a permanent dipole moment, then $\frac{dk}{dT} < 0$, and hence $\delta Q < 0$ if $\delta D > 0$. Hence heat will be given off when the field is applied, and conversely.

From now on we will use only the free energy density in our considerations. This will enable us to equate changes in the free energy directly to the mechanical work quantities responsible for them without making it necessary to include thermal quantities in the energy balance. The free energy expression, applicable even

in the presence of dielectrics:

$$(7.10)\quad U = 1/2\iiint \vec{E}\cdot\vec{D}\, dv$$

behaves in electrical problems in the same manner as the chemical free energy does in chemical kinetics, in the sense that a reaction will proceed until the free energy takes on a minimum value. In the electric case, charges on a conductor will redistribute themselves in such a way that the over all free field energy will be minimized. We can show this directly. Let us consider a virtual process in which charges in equilibrium on a conductor are displaced by an infinitesimal amount along the constant potential conductor surfaces in such a way that the total charge remains unchanged. The variation of free energy is given by:

$$(7.20)\quad \delta U = 1/2\iiint kk_o\delta (E^2)dv = \iiint \vec{E}\cdot\delta \vec{D}\, dv$$

using Gauss' theorem and (3.6) and letting the surface term vanish, we have:

$$(7.21)\quad \delta U = \iiint \vec{E}\cdot\delta \vec{D}\, dv = \iiint -\vec{\nabla}\phi\cdot\delta \vec{D}\, dv = \iiint \left[\phi\vec{\nabla}\cdot\delta\vec{D} - \vec{\nabla}\cdot(\delta \vec{D}\phi)\right] dv$$

$$\delta U = \sum \phi_i \iiint \delta\rho_i dv - \iint \delta\vec{D}\phi\cdot d\vec{S} = \sum \phi_i \iiint \delta\rho_i dv = 0$$

The summation extends over each individual conductor which, since it is at equilibrium, is at constant potential ϕ_i. The last term vanishes since the total charge on each conductor is unchanged and thus the variation of the free energy when a system is in equilibrium is zero. This theorem, usually known as Thomson's theorem, shows that the free energy is actually an extremum at equilibrium.

The term:

$$(7.22)\quad U_v = \frac{\vec{E}\cdot\vec{D}}{2}$$

is known as the energy density* of the electrostatic field. It is a density in

---

* More accurately, free energy density.

---

the sense that its volume integral gives the overall energy of the field. On the other hand, in the same sense as it was impossible to localize the energy either in the field, or in the source charges, it is also impossible to associate energy

in a definite way with each specific volume of field in a manner which can be verified by experiment.

In deriving the energy expression, it is assumed that the medium is held at rest and hence no work is done in motion against forces. This implies that the virtual process of assembling the charges in the dielectrics is a process with particular constraints. The resultant energy expression is nevertheless general, since no non-conservative forces are involved. In Chapter 8, we shall consider the more general virtual process permitting mass motion; (7.14) will however continue to apply since the final field energy is independent of history.

CHAPTER 8 FORCES IN THE ELECTROSTATIC FIELD

We shall now derive the force per unit volume that acts on a dielectric body when it is under the influence of an external electrostatic field. This may be derived from the energy principle. The variation in free energy δU when a unit volume of the dielectric undergoes a virtual displacement δx is given in terms of $\vec{F}_v$ the force per unit volume by:*

$$(8.1)\quad \delta U = -\iiint \vec{F}_v \cdot \vec{\delta} x \, dv$$

If we are able to consider virtual displacements within the dielectric, then in view of the fact that the magnitude of the virtual displacement δx is an arbitrary function of space, we can identify the quantity $\vec{F}_v$ in Equation (8.1) with the true volume force. To say this in a different way, if $\vec{u}$ is an arbitrary velocity field within a dielectric, then the rate at which energy is lost by the field is given by:

$$(8.2)\quad \frac{dU}{dt} = -\iiint \vec{F}_v \cdot \vec{u} \, dv$$

where $\vec{F}_v$ represents the volume force as given above. The total free energy of the dielectric, and the field outside the dielectric, is given by:

$$(7.10)\quad U = \frac{1}{2}\iiint \vec{E}\cdot\vec{D} \, dv = \frac{1}{2}\iiint \frac{D^2}{kk_o} \, dv$$

If a virtual displacement of the dielectric is carried out, there are two factors which can bring about an energy change. One factor is a change in the true charge densities and the other factor is a change in the specific inductive capacity. We can put:

$$(8.3)\quad \delta U = \left(\frac{\partial U}{\partial \rho}\right)_k \delta\rho + \left(\frac{\partial U}{\partial k}\right)_\rho \delta k$$

The charge dependent term is given by:

$$(8.4)\quad \left(\frac{\partial U}{\partial \rho}\right)\delta\rho = \iiint \phi \, \delta\rho \, dv$$

as was shown in the proof of Thomson's theorem (7.21). In computing the term that depends on the change in the dielectric constant we must keep the total true

---

* It is assumed here that the virtual velocities $\vec{u}$ corresponding to the virtual displacements $\delta\vec{x}$ are sufficiently slow that the process is: a) reversible and b) isothermal. Under these conditions the change in free energy can be equated to the mechanical work done.

charge constant. Thus by taking the variation of (7.10) with respect to k we have:

$$(8.5)\quad \left(\frac{\partial U}{\partial k}\right)_\rho \delta k = \frac{\delta}{2k_o}\iiint \frac{D^2}{k}\, dv = -\frac{k_o}{2}\iiint E^2 \delta\, k dv + \iiint \vec{E}\cdot\delta\vec{D}\, dv$$

The last term vanishes since the volume integral of the product of an irrotational vector $\vec{E}$ and a solenoidal vector $\delta\vec{D}$ is zero. This fact may be demonstrated from the fact that an irrotational vector field may be represented by a scalar potential, with the aid of Equation (7.6). The field $\vec{E}$ is irrotational from (1.10) and the variation of the displacement is solenoidal from the fact that the charge is constant and the relation (3.4). Hence substituting (8.4) and (8.5) into (8.3) and equating to (8.1) we have:

$$(8.6)\quad \delta U = \iiint \left(\phi\delta\rho - \frac{k_o}{2} E^2 \delta k\right) dv = -\iiint \vec{F}_v\cdot\delta\vec{x}\, dv$$

and dividing (8.6) by δt we have:

$$(8.7)\quad \frac{dU}{dt} = \iiint \left(\phi\frac{\partial\rho}{\partial t} - \frac{k_o}{2}E^2\ \frac{\partial k}{\partial t}\right) dv = -\iiint \vec{F}_v\cdot\vec{u}\, dv$$

We must now express the time dependent derivatives $\partial\rho/\partial t$ and $\partial k/\partial t$ in terms of the arbitrary velocity field $\vec{u}$. This can be done in terms of the hydrodynamic equations of continuity:

$$(8.8)\quad \vec{\nabla}\cdot(\rho\vec{u}) + \frac{\partial\rho}{\partial t} = 0 \qquad \rho = \text{the charge density}$$

$$(8.9)\quad \vec{\nabla}\cdot(g\vec{u}) + \frac{\partial g}{\partial t} = 0 \qquad g = \text{the mass density}$$

which represent respectively the conservation of charge and mass. In order to calculate $\partial k/\partial t$ we must associate the change in dielectric constant with the velocity flow. Since there is a net transport of material in a velocity field the change in dielectric constant can only be associated with changes in geometry if we consider the time history of a volume element that is moving with the velocity $\vec{u}$. The total derivative of a particular quantity, such as k or g, when evaluated so that the observation point for this derivative moves with a chosen volume element in a velocity field, is known as the substantial derivative and is related to the partial derivatives and to the velocity $\vec{u}$ by the relation:

$$(8.10)\quad \frac{Dk}{Dt} = \frac{\partial k}{\partial x}\frac{\partial x}{\partial t} + \frac{\partial k}{\partial y}\frac{\partial y}{\partial t} + \frac{\partial k}{\partial z}\frac{\partial z}{\partial t} + \frac{\partial k}{\partial t} = \vec{\nabla} k\cdot\vec{u} + \frac{\partial k}{\partial t}$$

Hence the desired partial derivatives are:

$$(8.11)\quad \frac{\partial k}{\partial t} = -\vec{\nabla} k\cdot\vec{u} + \frac{Dk}{Dt} \quad \text{or} \quad \frac{\partial g}{\partial t} = -\vec{\nabla} g\cdot\vec{u} + \frac{Dg}{Dt}$$

If we have a dielectric equation of state, that is a relation which gives the dependence of the dielectric constant on the density, such as the Clausius-Mosotti equation (3.44), then the substantial derivative of the dielectric constant can be expressed in terms of the substantial derivative of the density by:

$$(8.12)\quad \frac{Dk}{Dt} = \frac{dk}{dg}\frac{Dg}{Dt}$$

The assumption that the dielectric constant depends on the mass density alone includes of course the assumption that the virtual processes are isothermal as was discussed in Chapter 7. The substantial derivative of the density with respect to time in Equation (8.12) can be evaluated with the aid of (8.11) and the equation of continuity (8.9) giving:

$$(8.13)\quad \frac{Dk}{Dt} = \frac{dk}{dg}\left(\frac{\partial g}{\partial t} + \vec{\nabla} g\cdot\vec{u}\right) = \frac{dk}{dg}\left[\vec{\nabla} g\cdot\vec{u} - \vec{\nabla}\cdot(g\vec{u})\right] = -\frac{dk}{dg}\, g\vec{\nabla}\cdot\vec{u}$$

Substituting this into (8.11) we get:

$$(8.14)\quad \frac{\partial k}{\partial t} = -\frac{dk}{dg}\, g\vec{\nabla}\cdot\vec{u} - \vec{\nabla} k\cdot\vec{u}$$

And then substituting (8.8) and (8.14) into (8.7) we have:

$$(8.15)\quad \frac{dU}{dt} = \iiint \left[-\phi\vec{\nabla}\cdot(\rho\vec{u}) + \frac{k_o}{2}E^2\frac{dk}{dg}g\vec{\nabla}\cdot\vec{u} + \frac{k_o}{2}E^2\vec{\nabla} k\cdot\vec{u}\right] dv$$

This expression must be brought into the form of (8.2), the dot product of an expression and the velocity $\vec{u}$, in order to evaluate the volume force $\vec{F}_v$. The first term can be put in this form by using (7.6) to integrate it by parts, and assuming that the integrals are extended over all space:

$$(8.16)\quad -\iiint \phi\vec{\nabla}\cdot(\rho\vec{u}) = \iiint\left[\rho\vec{u}\cdot\vec{\nabla}\phi - \vec{\nabla}\cdot(\phi\rho\vec{u})\right]dv = \iiint \rho\vec{u}\cdot\vec{\nabla}\phi\, dv - \int \phi\rho\vec{u}\cdot d\vec{s}$$

$$= \iiint \rho\vec{\nabla}\phi\cdot\vec{u}\, dv$$

The surface term vanishes since the boundary surface may be assumed to be outside of the dielectric, and therefore outside of a region of charge density. Similarly, the second term in (8.15) can be put in the desired form by integrating by parts and dropping the surface term:

$$(8.17)\quad \frac{k_o}{2}\iiint E^2 g \frac{dk}{dg}\vec{\nabla}\cdot\vec{u}\,dv = \frac{k_o}{2}\iiint \vec{\nabla}\cdot(E^2\frac{dk}{dg}\vec{gu})\,dv - \frac{k_o}{2}\iiint \vec{\nabla}(E^2\frac{dk}{dg}g)\cdot\vec{u}\,dv$$

Collecting terms (8.15) becomes

$$(8.18)\quad \frac{dU}{dt} = \iiint \left[-\rho\vec{E} + \frac{k_o}{2}E^2\vec{\nabla}k - \frac{k_o}{2}\vec{\nabla}(E^2\frac{dk}{dg}g)\right]\cdot\vec{u}\,dv$$

and by comparing (8.18) and (8.2) we conclude that:

$$(8.19)\quad \vec{F_v} = \rho\vec{E} - \frac{k_o}{2}E^2\vec{\nabla}k + \frac{k_o}{2}\vec{\nabla}(E^2\frac{dk}{dg}g)$$

is the volume force. The first term in (8.19) gives the ordinary electrostatic volume force in agreement with (1.1). The second term gives a force which will appear whenever an inhomogenous dielectric is in an electric field. The last term, known as the electrostriction term, gives a volume force on a dielectric in an inhomogenous electric field. Note that the magnitude of the electrostriction term depends explicitly through $\frac{dk}{dg} = \left(\frac{\partial k}{\partial g}\right)_T$ on the electrical equation of state of the material. It is interesting to note that the last term will never give a net force on a finite region of dielectric if we integrate it over a large enough protion of dielectric so that its extremities are in a field-free region. Under this condition, the electrostriction term, being a pure gradient term, will integrate out. It is for this reason that this term is frequently omitted, since in the calculation of over all total forces on dielectric bodies, it usually does not contribute. In cases where it can be omitted, however, an incorrect pressure variation within the dielectric is obtained, even though the total force is given correctly.

Maxwell Stress Tensor

Before we consider any specific examples of the application of the expression

for the general volume force given in Equation (8.19) we will reformulate this volume force in terms of its "space stress". In accordance with a pure field theory it should be possible to calculate the net force on a given volume element within a dielectric in terms of the field conditions on the surface of this volume element. This implies that the field is the stress transmitting medium in the same sense that a string tying two weights together is the medium that transmits a force from one weight to the other. This was a point that was emphasized by Maxwell in order to bring out the importance and the physical reality of field quantities. Again, in a similar manner to the analogous energy case that we discussed in Chapter 7, we can only give an alternate description of the way in which the forces act, and cannot give a definite physical proof of the validity of the field concept, as exemplified by Equation (8.19), compared to the action at a distance concept. Since the only physical fact that underlies this entire discussion is Coulomb's law, the remainder of the discussion being mathematical, one cannot expect to obtain any physical concept regarding the mechanical interaction of charges which will add any physical facts beyond Coulomb's law. New physical facts based on the field concept will arise only when time dependent effects in the present theory are further investigated in later chapters.

If we consider that the force acting on a given volume is transmitted across the elements of surface bounding that volume, then this transmitting force can be formulated in terms of a quantity known as the stress tensor T. The $ij^{\underline{th}}$ component T_{ij} of the stress tensor T is so constituted that the $i^{\underline{th}}$ component dF_i of the force $\vec{dF}$ transmitted across a surface element $\vec{dS}$ whose component in the $j^{\underline{th}}$ direction is dS_j, is given by:

$$(8.20) \quad dF_i = \sum_{j=1}^{3} T_{ij}\, dS_j$$

It can be shown by the consideration of the equilibrium of a rectangular solid under surface stresses that the stress tensor T must be symmetric. We shall

adopt the so-called Einstein summation convention which ordinarily states that a summation is to be carried out over indicies that are repeated in any single term. So Equation (8.20) can be written as:

$$(8.21)\quad dF_i = T_{ij}dS_j$$

If we integrate (8.21) to give the $i^{\underline{th}}$ component of the total force acting on a given volume, then this force is given by:

$$(8.22)\quad F_i = \iint T_{ij}dS_j$$

If this force is to be expresable in terms of a volume force, whose $i^{\underline{th}}$ component is F_{vi} then:

$$(8.23)\quad F_i = \iint T_{ij}dS_j = \iiint F_{vi}dv$$

and hence, by Gauss' theorem (1.7) expressed in tensor notation:

$$(8.24)\iiint \frac{\partial T_{ij}}{\partial x_j}\,dv = \iint T_{ij}dS_j$$

the relation between the stress tensor and the volume force is:

$$(8.25)\quad F_{vi} = \frac{\partial T_{ij}}{\partial x_j}$$

Thus if we express the volume force (8.19) as the tensor divergence of a certain quantity T then the quantity T can be identified with the surface stress tensor T that gave the stress transmitted by the field across the surface of the volume in Equation (8.20). Let us write down Equation (8.19) in tensor notation:

$$(8.26)\quad F_{vi} = E_i \frac{\partial D_j}{\partial x_j} - \frac{k_o}{2} E_j E_j \frac{\partial k}{\partial x_i} + \frac{k_o}{2} \frac{\partial}{\partial x_i}\left(E_j E_j g \frac{dk}{dg}\right)$$

If we let:

$$(8.27)\quad \beta = \frac{g}{k}\frac{dk}{dg}$$

and consider the tensor relation:

$$(8.28)\quad \frac{k_o}{2}\frac{\partial}{\partial x_i}(E_j E_j k) = \frac{2kk_o}{2} E_j \frac{\partial E_j}{\partial x_i} + \frac{k_o}{2} E_j E_j \frac{\partial k}{\partial x_i}$$

and use the tensor form of $\vec{\nabla} \times \vec{E} = 0$:

$$(8.29)\quad \frac{\partial E_j}{\partial x_i} = \frac{\partial E_i}{\partial x_j}$$

We have on substituting (8.27), (8.28) and (8.29) into (8.26):

$$(8.30)\quad F_{vi} = E_i \frac{\partial D_j}{\partial x_j} + D_j \frac{\partial E_i}{\partial x_j} - \frac{k_o}{2} \frac{\partial}{\partial x_i} \left[E_j E_j k(1-\beta) \right]$$

which can be written as:

$$(8.31)\quad F_{vi} = \frac{\partial}{\partial x_j} (E_i D_j) - \frac{\delta_{ij}}{2} \frac{\partial}{\partial x_j} \left[(1-\beta) E_k D_k \right]$$

By comparing (8.31) with (8.25) we see that:

$$(8.32)\quad T_{ij} = E_i D_j - \frac{\delta_{ij}}{2} (1-\beta) E_k D_k = kk_o \left[E_i E_j - \frac{\delta_{ij}}{2} (1-\beta) E_k E_k \right]$$

is the complete expression for the Maxwell stress tensor. Writing out explicitly the matrix corresponding to this tensor, we obtain:

$$(8.33)\quad T = kk_o \begin{pmatrix} \frac{1}{2}(E_x^2 - E_y^2 - E_z^2) & E_xE_y & E_xE_z \\ E_xE_y & \frac{1}{2}(E_y^2 - E_z^2 - E_x^2) & E_yE_z \\ E_xE_z & E_yE_z & \frac{1}{2}(E_z^2 - E_x^2 - E_y^2) \end{pmatrix}$$

where β has been set equal to zero, for simplicity. Note that additional terms will appear in the stress tensor if the field is not irrotational as was assumed above. The Maxwell tensor is a symmetric tensor of the second rank and can therefore be reduced to three components by transformation to principal axes, by the standard method. The principal values of the matrix can be obtained by solving the secular determinant:

$$(8.34)\quad \left| T_{ij} - \delta_{ij} \lambda \right| = 0$$

which gives two equal and one unequal value for λ. The principal values of the tensor, when $\beta = 0$ are:

$$(8.35)\quad \lambda_1 = \frac{kk_o}{2} E^2 \quad , \quad \lambda_{2,3} = -\frac{kk_o}{2} E^2$$

Therefore the Maxwell stress tensor T' when expressed in terms of principal coordinates takes the simple form:

$$(8.36) \qquad T' = \frac{kk_0}{2} \begin{pmatrix} E^2 & 0 & 0 \\ 0 & -E^2 & 0 \\ 0 & 0 & -E^2 \end{pmatrix}$$

The principal axes are so oriented that the coordinate axis corresponding to the single root of the secular determinant λ_1 is parallel to $\vec{E}$, while the two axes corresponding to the double roots λ_2 and λ_3 are perpendicular to $\vec{E}$. This fact is often expressed qualitatively by stating that the electric field transmits a tension $kk_0E^2/2$ parallel to the direction of the field and a contraction of magnitude $kk_0E^2/2$ transverse to the direction of the field.

Let us choose a coordinate system in which the x axis is parallel to the direction of the field so that $E_y = E_z = 0$. In this coordinate system consider the stress across a surface element as shown in Figure (8.1) whose normal makes

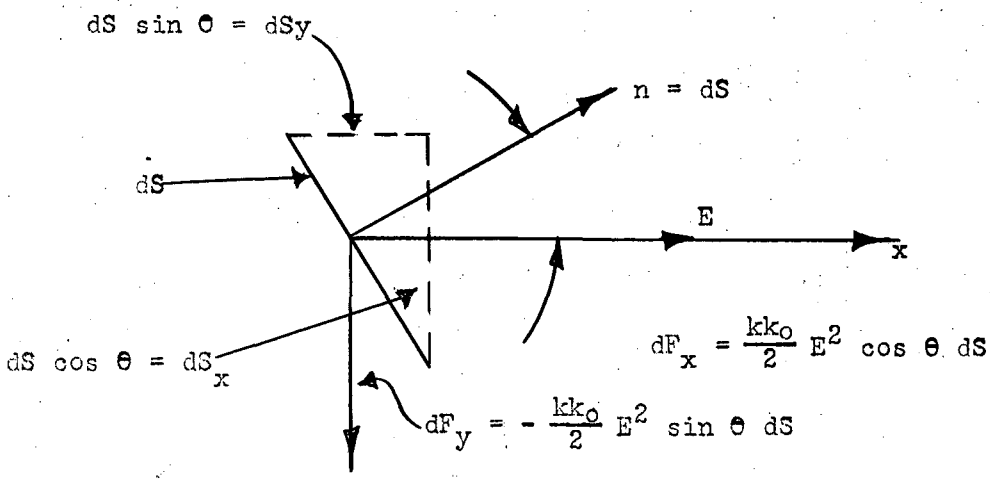


Figure (8.1)

an angle θ with the x axis. The stress will then have two components, one parallel to $\vec{E}$, and one perpendicular to $\vec{E}$, lying in the plane of $\vec{E}$ and $\vec{n}$ the normal to the surface. The magnitudes of these stresses are then the stress components given by the matrix in Equation (8.36), multiplied by the surface element components as indicated in the figure. The resultant stress on $\vec{dS}$ can then be obtained by taking the vector sum of the two stress components as shown in Figure (8.2). It is seen that the electric field bisects the angle between

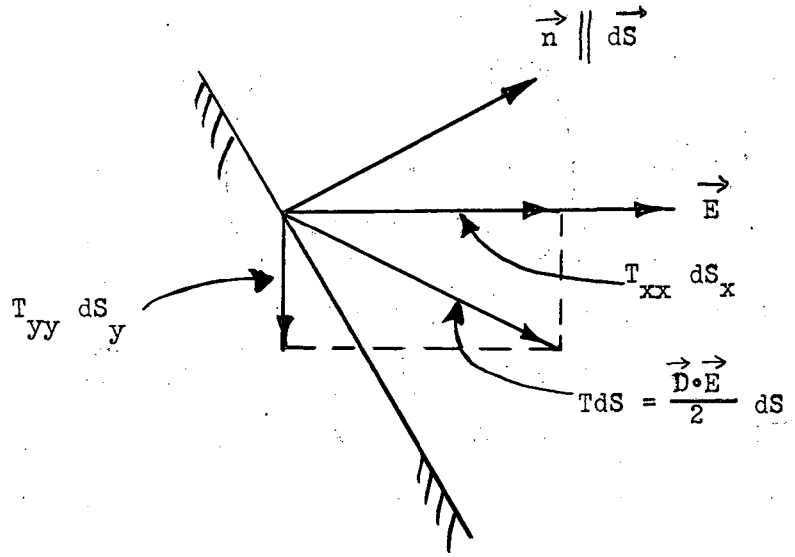


Figure (8.2)

the normal to the surface and the direction of the resultant stress acting on the surface. This construction is frequently a useful one in the graphical evaluation of the forces on a charged region if an experimental field plot is available, or in the analogous magnetic case to be discussed later, this construction is useful for the computation of forces on magnetized materials, or on current-carrying conductors.

In the special case of stress transmitted across surfaces either parallel or normal to the electric field, we have the simple situation indicated in Figure (8.3) where the field transmits a pull of magnitude ED/2 across a surface that is normal to the field and a push of magnitude ED/2 across a surface that is

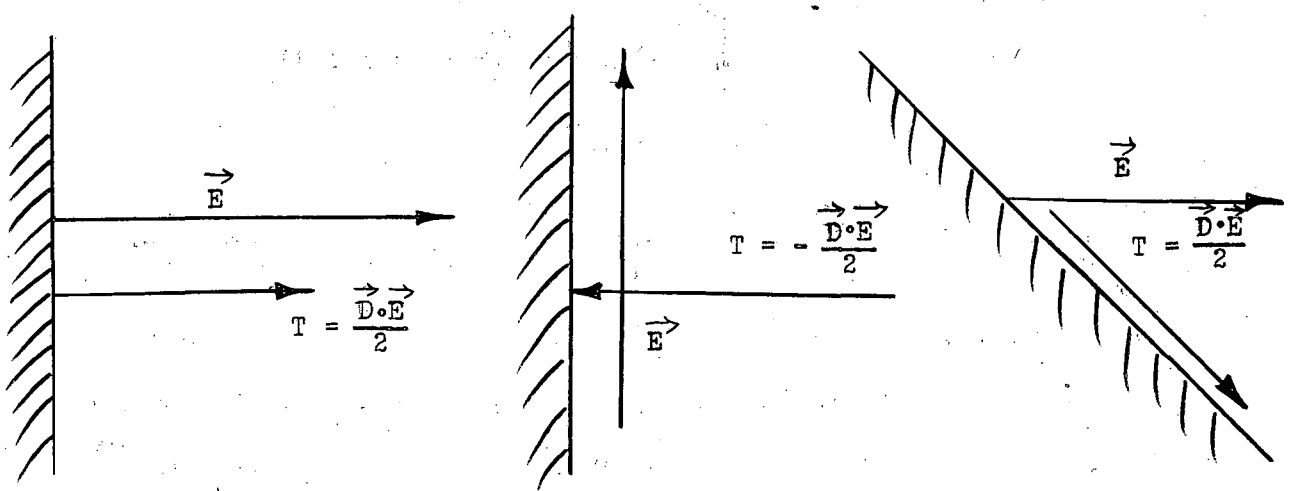


Figure (8.3)

tangential to the field. A surface that is oriented at 45° to the direction of

the field as seen in Figure (8.3) will receive a force that acts parallel to the surface, also of magnitude ED/2 per unit area, of the surface. These relations can be demonstrated for simple cases such as the attraction and repulsion between two charges of opposite or equal sign. If we consider, for example, two charges of equal magnitude, but opposite sign, then the lines of force are distributed as in Figure (8.4).

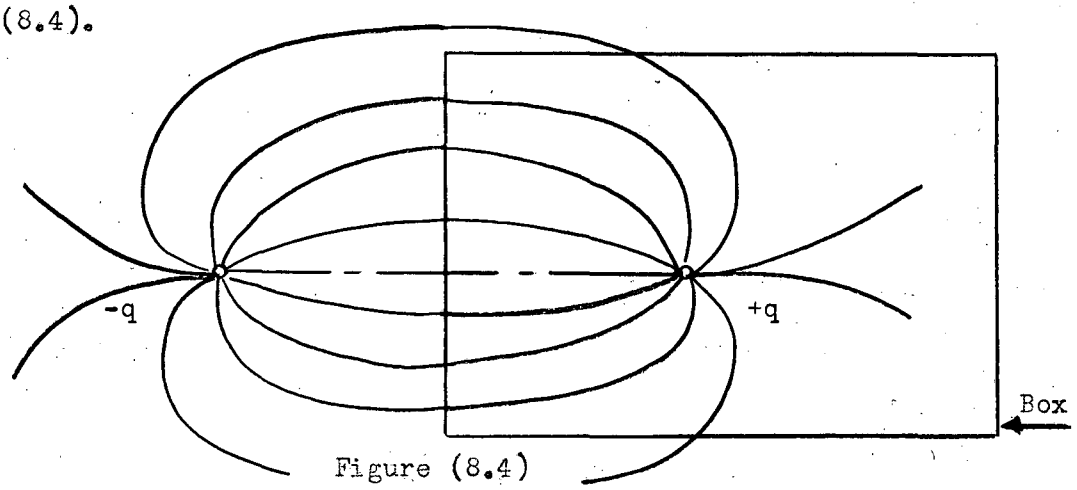


Figure (8.4)

If we integrate the stress tensor over the surfaces of a box one of whose faces is the plane of symmetry between the two charges, considering the other faces of the box to be at infinity, we will get an expression that is in agreement with the Coulomb attraction (1.2). If we consider the two equal charges as in Figure (8.5) and the same box as in Figure (8.4) then the lines of force are

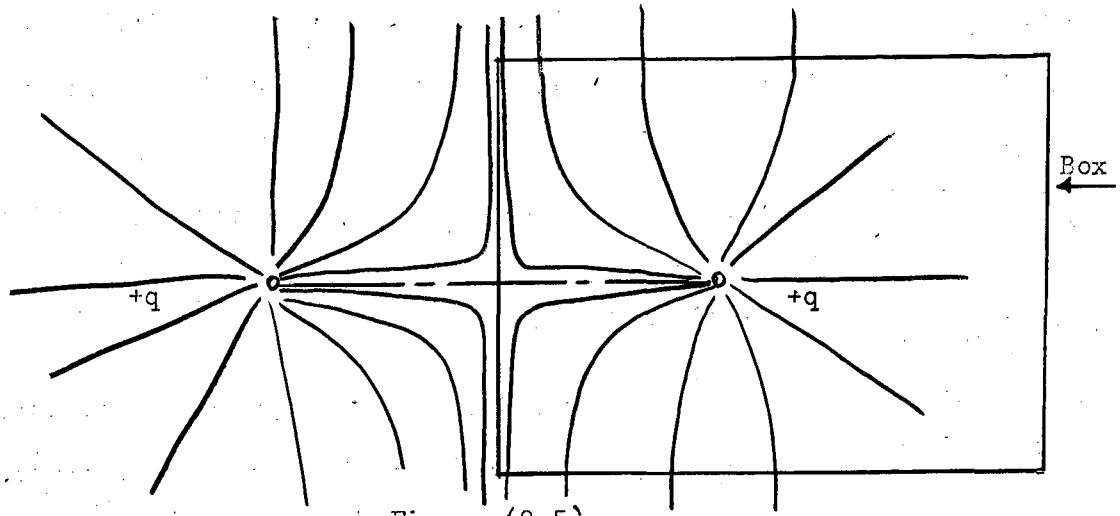


Figure (8.5)

parallel at the plane of symmetry between the charges resulting in a repulsion whose magnitude can, by integration across the plane of symmetry, be shown to also be in accordance with the Coulomb repulsion.

The Behavior of Dielectric Liquids in an Electrostatic Field.

We shall now turn to some applications of these expressions for the forces in an electrostatic field. We will first treat the behavior of a dielectric liquid in an electric field. If we consider an uncharged dielectric liquid acted on by the volume force given by Equation (8.19), then the pressure gradient at any point within the liquid is given by:

$$(8.37)^{\bigstar} \quad \vec{\nabla} p = \vec{F}_v = -\frac{k_o E^2}{2}\vec{\nabla} k + \frac{k_o}{2}\vec{\nabla}\left(E^2 g \frac{dk}{dg}\right)$$

This can be written as:

$$(8.38) \quad \vec{\nabla} p = \frac{k_o g}{2}\vec{\nabla}\left(E^2 \frac{dk}{dg}\right)$$

and integrating this, assuming a definite equation of state of the liquid, we obtain:

$$(8.39) \quad \int_{p_1}^{p_2} \frac{dp}{g} = \frac{k_o}{2}\left\{\left[E^2 \frac{dk}{dg}\right]_2 - \left[E^2 \frac{dk}{dg}\right]_1\right\}$$

This equation denotes the important fact that the pressure within the dielectric liquid is a unique function of the electric field at a given point, the function depending on the electrical and the mechanical equation of state of the liquid. Equation (8.39) also indicates that the net pressure difference between two points outside the region of the electric field, resulting from electrical forces, in a dielectric liquid will vanish. A situation that involves boundaries will be analyzed later.

If we consider the liquid to be incompressible, Equation (8.39) reduces to:

$$(8.40) \quad p_2 - p_1 = \frac{g k_o}{2}\left[E^2 \frac{dk}{dg}\right]_1^2$$

from which the magnitude of the pressure difference can be estimated numerically in terms of the Clausius-Mosotti relation or a similar equation of state. If the Clausius-Mosotti relation is valid, Equation (8.40) for an incompressible

---

$\bigstar$ We are putting here $\vec{F}_v = +\vec{\nabla} p$ where p is the mechanical pressure in the liquid when in equilibrium with the electrical volume force $\vec{F}_v$. As a result of the pressure gradient a mechanical force $\vec{F}_{v(mech)} = -\vec{\nabla} p$ is set up; thus $\vec{F}_v + \vec{F}_{v(mech)} = 0$, which is the equilibrium condition.

fluid becomes:

$$(8.41)\quad p_2 - p_1 = \frac{Ng\alpha}{18M}\left[E^2\,(k+2)^2\right]_1^2 = \left[\frac{k_o\,E^2\,(k-1)\,(k+2)}{2\qquad 3}\right]_1^2$$

Let us now consider the stresses that act across the boundary between two dielectrics or simply the case of a boundary between a dielectric of specific inductive capacity k and a vacuum. We assume that the transition from dielectric to vacuum takes place in a continuous manner, as indicated by Figure (8.6). Consider as an example a two dimensional problem involving a pair of condenser plates as shown in Figure (8.7) that dip into a dielectric liquid. If only the net pressure difference from A to D is desired, then this can be computed by

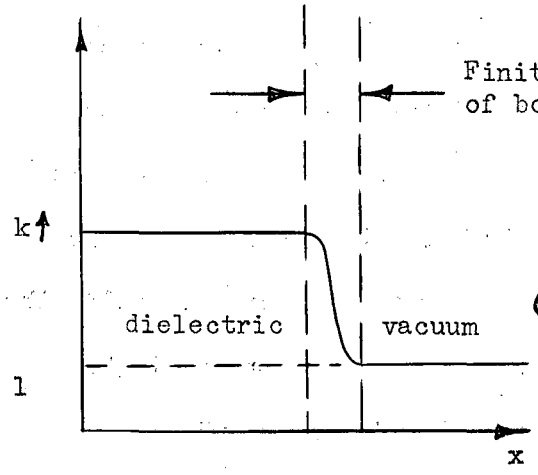


Figure (8.6)

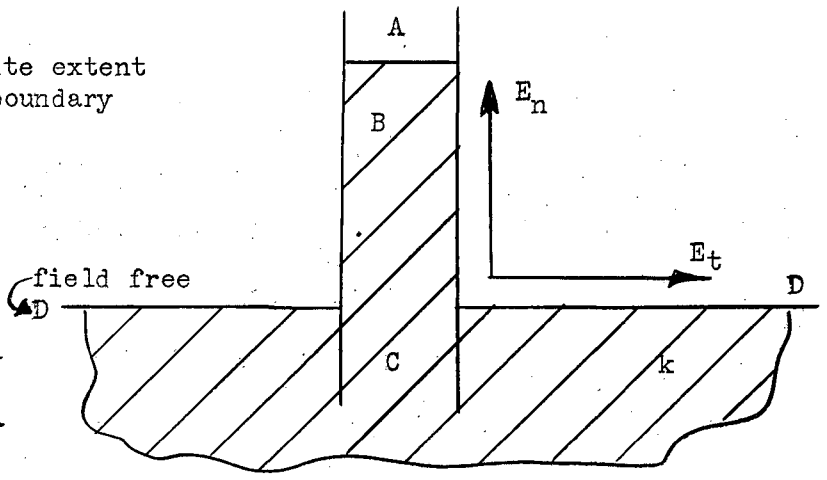


Figure (8.7)

integrating only the second term, the term that is proportional to the gradient of the dielectric constant, of the force equation (8.19). The resultant electrical pressure difference, which has to be balanced by hydrostatic effects is therefore:

$$(8.42)\quad p_A - p_D = \frac{k_o}{2}\int_A^D E^2 \vec{\nabla} k \cdot d\vec{x}$$

or:

$$(8.43)\quad p_A - p_D = \frac{k_o}{2}\int_A^D (Et^2 + En^2)\,\frac{dk}{dx}\,dx$$

Treating the tangential and the normal components of the electric field in accordance with the boundary conditions expressed by Eqs. (3.15) and (3.19), we obtain:

$$(8.44) \qquad p_A - p_D = \frac{k_o}{2}\left[E_{tB}^2\,(k-1) + k^2 E_{nB}^2 \int_A^D \frac{dk}{k^2}\right]$$

$$p_A - p_D = \frac{k_o(k-1)}{2}\left[Et^2 + kEn^2\right]_B$$

Note that the field quantities in equation (8.44) refer to the field inside the liquid. This formula will give directly the rise of the liquid in the condenser plates. However, this formula is insufficient to describe the detailed pressure behavior of the liquid from A to D, since the detailed pressure behavior also depends on the electrostriction term. In fact, the pressure change as shown in Figure (8.8) from A to B is actually of opposite sign from the pressure change from A to D. As the field decreases from B to D, the pressure decreases below the outside value at A by an amount which is larger than the pressure rise at the surface A-B. The net difference computed in Equation (8.44) gives only the difference in pressure between A and D. The pressure which forces the liquid up is actually exerted at the region C where the field is inhomogeneous and not at the surface of the liquid, as is seen in Figure (8.8). The physical reason for

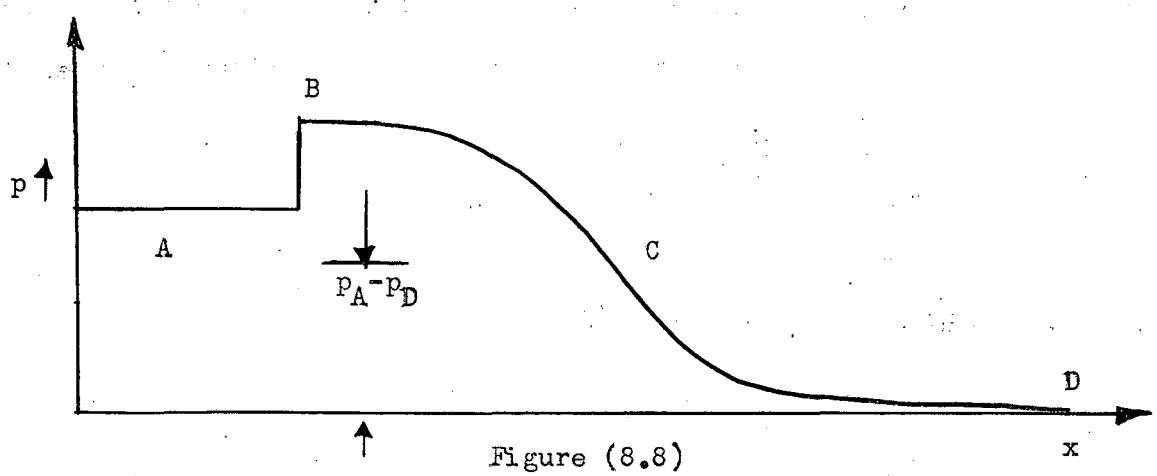


Figure (8.8)

this is that the energy of dipoles in an electric field is lower than their energy in field-free space and therefore the dipoles in the liquid are drawn into the regions of higher field in order to satisfy the criterion for mechanical equilibrium which requires a minimum potential energy. This action on the dipoles takes place in the region C where the field commences to build up. This minimum energy effect is partially counterbalanced by the electrostriction drop at AB resulting in a net pressure drop as given by Equation (8.44). This example shows that considerable care is necessary in applying the force equations in dielectrics.

Let us consider another example which appears to be extremely simple, but which actually leads to an apparent paradox. If we consider a set of charged conductors so arranged that they may be immersed in a dielectric liquid, then if the true charges on these conductors are kept constant as a liquid is introduced between them, the free energy of the system as given by Equation (7.19) will drop in the ratio 1/k, since D remains constant but E is reduced, in this ratio. If on the other hand the voltages had been maintained at their initial values as the liquid was introduced, then the free energy would be increased by a factor k, since in this case E remains constant, while D increases by a factor k. Of course, these arguments only pertain if all of the space between the conductors and outside of them is filled with a dielectric liquid* . Otherwise we cannot

---

*At least all space where there are electric fields must be so filled.

---

assume, as is implied in the above statements, that the distribution of E and D remains constant as the dielectric material is introduced between the plates. This means that if a system maintained at constant charge is totally surrounded by a dielectric liquid all mechanical forces will drop in the ratio 1/k. This is the reason why a factor 1/k is frequently included in the expression for Coulomb's law (1.2) to indicate this decrease in force. The physical significance of this decrease in force, which is required by energetic considerations, is often somewhat mysterious since it is hard to see on the basis of a

field theory why the interaction between two charges should be dependent upon the nature or condition of the intervening material, and therefore the inclusion of an extra factor $1/k$ in Coulomb's law lacks a physical explanation.

Let us consider this problem in the simple two dimensional geometry of a parallel plate condenser as in Figure (8.9). Let $\pm q_s$ be the surface charge per unit area on each condenser plate and $\pm q_p$ be the polarization charge on the

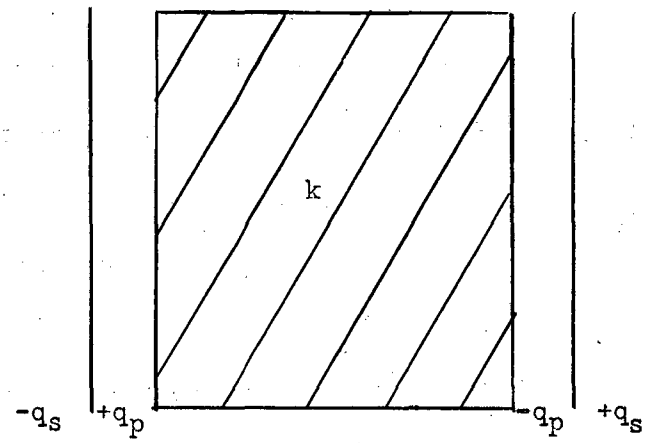


Figure (8.9)

outer surface of the intervening dielectric slab. Let the true surface charge q_s be assumed to be constant during the introduction of the dielectric slab. The purely electrical force acting on each condenser plate will be the sum of the force due to the fields produced by the charges q_s on the opposite plates and the charges q_p on the two surfaces of the dielectric. The two layers of polarization charge will produce equal and opposite fields on each plate and their effects will therefore cancel each other. From the point of view of electrical interaction alone, it is not obvious why any change in force at all is obtained when the dielectric layer is introduced, since the only direct interaction between the charges q_s on the plates, which are assumed to remain constant, seems to be unaffected by the introduction of the dielectric slab. That is the force per unit area remains:

$$(8.45) \quad F_s = \frac{q_s^2}{2k_o}$$

as long as the dielectric does not touch the plates of the condenser. Therefore the decrease in force to:

$$(8.46)\quad F_s = \frac{q_s^2}{2kk_o}$$

which is experienced when the experiment is performed with a liquid, that wets the plates and also completely surrounds them, can not be explained by electrical forces alone. This apparent paradox can be explained by taking into account the difference in pressure in the liquid in the field-filled space between the condenser plates and in the field-free space outside the condenser plates. By equation (8.44), this difference in pressure is given by:

$$(8.47)\quad p_B - p_A = \frac{k_o-1}{2k_o}\left(1 - \frac{1}{k}\right) D_n^2 = \frac{k_o-1}{2}\left(1 - \frac{1}{k}\right) q_s^2$$

The sum of the force resulting from this pressure and the pure electrical force given in (8.45) gives the total force given in (8.46) which was derived from energy considerations. Thus the decrease in force that is experienced between two charges when they are immersed in a dielectric liquid can be understood only by considering the effect of the pressure of the liquid on the charges themselves. In accordance with the philosophy of the action at a distance theory no change in the purely electrical interaction between the charges takes place.

CHAPTER 9 STATIONARY CURRENTS

In the discussion of electrostatics in preceding chapters all currents were assumed to be zero. We shall now take the flow of charge, or current into account, but shall assume that the currents are varying slowly in time. This means that the current flow will depend entirely on the electric fields that are present and will not depend on magnetic interactions. In the absence of numerical estimates of the relative magnitude of magnetic interactions and of the interaction of currents with the lattice structure of the resistive medium which gives rise to electrical resistance, it is not obvious that a situation ever exists in which currents depend only on the electric fields. But it turns out that the magnetic effects can be neglected when the fields vary at low frequencies, provided that the dimensions of the conductors involved are small compared to the so-called "skin depth" of the currents in the particular conductor. We will also neglect an additional effect, known as the Hall effect, which is present even at zero frequency, and gives rise to redistribution of the equal potential surfaces in a current-carrying conductor. However, in all but very special substances, this effect is extremely small.

The conservation of charge in the flow of current in a medium is expressed by the equation of continuity:

$$(9.1)\quad \vec{\nabla}\cdot\vec{j} + \frac{\partial \rho}{\partial t} = 0 \quad , \qquad \vec{j} = \rho\vec{v} = \text{the current density within the medium (Amperes/Meter}^2\text{)}$$

The current flow is called stationary if there is no accumulation of the charge at any point. This criterion is expressed by:

$$(9.2)\quad \vec{\nabla}\cdot\vec{j} = 0$$

In order to relate the theory of current flow to the theory of the electric field, another equation is necessary which will give the connection between the current and the field that exist at a particular point in the conducting material.

This equation is:

(9.3) $\vec{j} = \sigma\vec{E}$ σ = the electrical conductivity (Mho/Meter)

Equation (9.3) is equivalent to Ohm's law. This relation is an over all phenomenological characteristic, and it may not be valid in many cases. The range of current densities over which Equation (9.3) is valid is called the linear range of the particular material, and can be very large, as in metals, or very small, as in a semi-conductor. Equation (9.3) implies that the conduction is isotropic. In crystals having a low symmetry (9.3) must be replaced by a tensor equation.

Stationary current flow is impossible in a purely irrotational electric field, since in stationary current flow energy is expended at a rate $\vec{j} \cdot \vec{E}$ per unit volume and this energy cannot be provided by an irrotational field.

Hence stationary currents are possible only in case there are present additional sources of the electric field known as electromotive forces, which are not irrotational. Denoting such non-irrotational electromotive fields by $\vec{E}'$, the conduction equation (9.3) will therefore become:

(9.4) $\vec{j} = \sigma(\vec{E} + \vec{E}')$

Defining the electromotive force as:

(9.5) $\mathcal{E} = \oint(\vec{E} + \vec{E}')\cdot d\vec{\ell} = \oint\vec{E}'\cdot d\vec{\ell} = \oint\frac{\vec{j}\cdot d\vec{\ell}}{\sigma}$, $\mathcal{E}$ = electromotive force (Volts)

Note that the conservative part of the field, $\vec{E}$, drops out of the closed line integration. This means that the current flow is due entirely to the non-conservative forces, and is only influenced by the conductivity and the geometry.

In case the current density is nearly constant over major portions of the path of integration, as frequently happens, Equation (9.5) can be written as:

(9.6) $\mathcal{E} = J\oint\frac{d\ell}{\sigma S} = JR$ $J = |j|S$ = the total current, a constant for the circuit (Amperes)

R = the resistance of the conductor (Ohms)

S = the cross sectional area of the conductor where the current density is $|j|$.

This is the form in which Ohm's law is usually stated. Note that in a case where there is no current flow, we obtain, by integrating Equation (9.4) along a line between two points 1 and 2 which traverses all of the region in which there is a non-conservative force or field:

$$(9.7)\quad -\int_1^2 \vec{E}\cdot d\vec{\ell} = \int_1^2 \vec{E'}\cdot d\vec{\ell} = \oint \vec{E'}\cdot d\vec{\ell} = \mathcal{E}$$

This indicates that the open circuit electrostatic voltage between two points is equal to the total electromotive force in the circuit. It follows, in the absence of a current flow, that within a particular region there are non-conservative fields $\vec{E'} = -\vec{E}$. Thus, for example, within a given boudary, the non-conservative fields (e.g. the chemical potentials) are exactly equal to the electrostatic field which is set up by the charges on the boundaries, in the absence of a current.

Formally speaking the current distribution and the field distribution are entirely defined by the non-conservative field and by the conductivity of the medium. Using Equations (1.10), (9.2) and (9.4) we have the following expressions for $\vec{E}$ and $\vec{j}$ in terms of $\vec{E'}$:

$$(9.8)\quad \vec{\nabla}\cdot\vec{j} = 0 \qquad \vec{\nabla}\cdot(\sigma\vec{E}) = -\vec{\nabla}\cdot(\sigma\vec{E'})$$

$$\vec{\nabla}\times\left(\frac{\vec{j}}{\sigma}\right) = \vec{\nabla}\times(\vec{E'}) \qquad \vec{\nabla}\times\vec{E} = 0$$

In the region where there are no non-conservative fields, $\vec{E}$ is derivable from a potential and hence in the case of stationary flow, the potential still obeys Laplace's equation:

$$(1.17)\quad \vec{E} = -\vec{\nabla}\phi \qquad \vec{\nabla}\cdot(\sigma\vec{\nabla}\phi) = 0 \text{ or } \nabla^2\phi = 0 \text{ if } \sigma \text{ is constant.}$$

The boundary conditions are changed, however, since now the conductivities, rather than the dielectric constants define the flux relation across a boundary. From Equations (9.2) and (9.3) we have the relation, in the absence of non-conservative fields:

$$(9.9)\quad \vec{\nabla}\cdot(\sigma\vec{E}) = 0$$

From this we have for the boundary condition between two mediums designated by the subscripts 1 and 2 respectively:

$$(9.10)\quad \vec{n} \cdot (\sigma_2 \vec{E}_2 - \sigma_1 \vec{E}_1) = 0$$

$$\vec{n} \cdot (\sigma_2 \vec{\nabla}\phi_2 - \sigma_1 \vec{\nabla}\phi_1) = 0$$

And as before in Equation (3.19):

$$(9.11)\quad \vec{n} \times (E_2 - E_1) = 0$$

$$\vec{n} \times (\vec{\nabla}\phi_2 - \vec{\nabla}\phi_1) = 0$$

It follows from (1.17), (9.10) and (9.11) that the solution of stationary current distribution problems is mathematically identical to the solution of electrostatic potential distribution problems that have the same geometry, and thus all of the methods that were developed in Chapters 4, 5 and 6 are applicable to these problems. The only difference between the static current problems and the electrostatic problems is that the conductivity in a given region may be zero, while the specific inductive capacity cannot become less than unity. This means that the type of boundary value problems which arise in stationary current flow may, under certain conditions, be quite different from any that can exist in electrostatic cases. As an example, if we consider that the region between a set of parallel condenser plates is filled with a medium of conductivity σ, then the current in the stationary current range will be exactly uniform over the entire area of the plates, within the conducting medium, while in the analogous electrostatic case, the field distribution will be only approximately uniform and will be disturbed by the fringing field at the edges of the plates.

In general, if electrostatic methods permit the calculation of the capacity between two electrodes, then one can conclude immediately what the resistance would be between these electrodes if all of the space in which they are located is filled with a homogeneous resistive medium. The capacity between two electrodes 1 and 2 is given by:

$$(9.12)\quad C = \frac{kk_0 \iint \vec{E} \cdot d\vec{S}}{\int_1^2 \vec{E} \cdot d\vec{\ell}} \qquad (C = \text{capacity in farads})$$

The numerator is the charge on each electrode by Gauss' flux theorem (1.6) and the denominator is the potential difference between the electrodes. The resistance between the two electrodes is given by:

$$(9.13)\quad R = \frac{\int_1^2 \vec{E}\cdot d\vec{\ell}}{\sigma \iint \vec{E}\cdot d\vec{S}}$$

where the denominator gives the net flux of current between the two electrodes and the numerator gives the potential difference between them. Comparing Equations (9.12) and (9.13) we have:

$$(9.14)\quad \frac{1}{R\sigma} = \frac{C}{kk_o}$$

or:

$$(9.15)\quad RC = \frac{kk_o}{\sigma}$$

Note that the product of the resistance and the capacity is a constant that depends only on the conductivity of the conductor and on the specific inductive capacity of the material between the condenser plates, and not on the geometry. However, it is not always possible to find an electrostatic problem which will be fully analogous to the corresponding stationary current problem, since the difference in the range of dielectric constants and resistances mentioned above causes different field patterns in the two cases. The formula for the capacity of a parallel plate condenser, when edge effects are neglected is:

$$(9.16)\quad C = \frac{kk_oS}{\ell}$$

ℓ = the distance between the plates.

S = the area of the plates

which when substituted into (9.15) gives:

$$(9.17)\quad R = \frac{\ell}{\sigma S}$$

ℓ = the distance between the ends of the wire.

S = the cross sectional area of the wire.

for the resistance of a wire of known length and cross sectional area. The formula for the resistance of a wire is of course applicable for large values of ℓ since the zero conductivity of the surrounding medium prevents fringing of the current flow lines, which is in strong contrast to the shape of the field lines that would arise from condenser plates that coincided with the ends of a

wire of ordinary length.

The equation of continuity:

$$(9.1)\quad \vec{\nabla}\cdot\vec{j} + \frac{\partial \rho}{\partial t} = 0$$

in combination with the conductivity equation:

$$(9.3)\quad \vec{j} = \sigma \vec{E}$$

and the source equation derived from (3.4) and (3.11):

$$(9.18)\quad \vec{\nabla}\cdot\vec{E} = \frac{\rho}{kk_o}$$

can be integrated with respect to the time in case we are dealing with a homogeneous dielectric or a homogeneous conductor. The integration gives:

$$(9.19)\quad \rho = \rho_o e^{-\frac{t}{\tau}}$$

where the characteristic time τ is given by:

$$(9.20)\quad \tau = kk_o/\sigma$$

τ is usually known as the relaxation time of the dielectric. Note that it applies only in the case of a homogeneous medium, for if it were not homogeneous, then the spatial dependence of the conductivity and the dielectric constant would have to be taken into account in the integration of Equation (9.1). The relaxation time is a characteristic time for a medium in that it gives an indication of the time in which essentially stationary conditions will be reached after the initiation of a particular current flow. The criterion which must be used to determine whether or not the stationary current equations will be applicable in a particular case is whether the time of observation following the inception of such currents exceeds the relaxation time τ by a sufficiently large amount.

CHAPTER 10 TYPES OF CURRENTS

We began the discussion of Electrodynamics with the treatment of steady currents since the original direct experimental observations of the magnetic interaction of currents were made with currents of this type. We must now consider the additional modifications that must be introduced into the theory if non-stationary currents are to be treated.

We shall discuss currents in a vacuum in a manner that is similar to the one which we used to treat charges in a vacuum and will then derive the magnetic effects that occur within an arbitrary medium. We shall classify currents in two categories, those that are true currents, that is may be identified with the motion of true charges, and other currents which are associated with the medium itself. This separation which is analogous to the separation that was made in the electrostatic theory between the potentials of true charges and the potentials of polarization charges, will lead us to consider two types of magnetic fields, one derived from true currents, and one derived from the combined effects of all the currents whatever may be their origin. It is this latter field, namely the magnetic field of induction $\vec{B}$, which can be considered to be the space-time average of the interatomic fields. Before proceeding to discuss magnetic interaction, let us classify the type of currents which we will consider in a medium.

1. True Currents - $\vec{j}$ - These currents are identical to the physical transportation of true charges.

2. Polarization Currents - $\frac{\partial \vec{P}}{\partial t}$ - These currents arise from the change of the polarization with time.

3. Magnetization Currents - $\vec{j}_m$ - These currents are stationary currents that flow within regions that are inaccessible to observation but which might give rise to net currents on boundaries, due to imperfect orbit cancellation on an atomic scale. We shall describe these magnetization currents in terms of

the magnetic moment per unit volume:

(10.1) $\vec{M} = \frac{1}{2}(\vec{r} \times \vec{j}_m)$ $\qquad\qquad$ $\vec{M}$ = magnetization (Weber/Meter^2)

corresponding to these currents. The magnetic moment of a particular volume is then:

(10.2) $\vec{m} = \iiint \vec{M}\, dv = \frac{1}{2}\iiint (\vec{r} \times \vec{j}_m)\, dv$

Note that this is analogous to the expression for the mechanical angular momentum in terms of the velocity of a volume when the charge density is associated with the mass. In the special case of a single "stationary current" loop that encloses a given area, $\vec{m}$ becomes simply the product of the current in the loop multiplied by the area of the loop and directed normal to the loop in a direction that agrees with the righthand rule for the current circulation. This is in agreement with the elementary definition of the magnetic moment of a current loop. We shall show later that:

(10.3) $\vec{j}_m = \vec{\nabla} \times \vec{M}$ $\qquad\qquad$ is the inverse of Equation (10.1)

4. Convective Currents: If a material medium in motion contains charges of various types, additional currents will be obtained which arise from convective effects. These convective currents will be derived both from the motion of the true and the polarization charges contained in the medium. There will also be changes in the net polarization current due to the motion of the medium. The convective currents will be discussed in a later chapter.

Let us consider the second and third sources of current in more detail. When expressed in terms of molecular coordinates, ξ, the electrical moment, $\vec{p}$, of a polarized molecule is defined by:

(10.4) $\vec{p} = \iiint \vec{P}\, dv = \iiint \rho \vec{\xi}\, dv$, $\vec{P} = \rho \vec{\xi}$

If the charge density ρ within the molecule is changing in time,* the polarization will change in time by an amount:

* The time variation of electrical quantities due to charge motions can be described either by considering changes in charge density as a function of time, or changes in coordinates of fixed charges; this discussion shows the equivalence of the two descriptions.

$$(10.5) \qquad \frac{\partial \vec{P}}{\partial t} = \iiint \frac{\partial \rho}{\partial t} \vec{\xi}\, dv$$

Substituting from the equation of continuity (9.1) this reduces to:

$$(10.6) \qquad \frac{\partial \vec{P}}{\partial t} = -\iiint \vec{\nabla} \cdot (\rho \vec{u})\, \vec{\xi}\, dv$$

Integrating by parts and dropping a surface term, which is justified by choosing the surface of integration so that it lies outside the region where there are molecular charges, we obtain:

$$(10.7) \qquad \frac{\partial \vec{P}}{\partial t} = \iiint \rho \vec{u} dv$$

or on a large scale:

$$(10.8) \qquad \frac{\partial \vec{P}}{\partial t} = \frac{\iiint \rho \vec{u} dv}{\iiint dv} = \overline{\rho \vec{u}}$$

Hence the quantity $\partial \vec{P}/\partial t$ does represent the space-time average value of the molecular currents caused by a varying polarization.

Let us now consider the magnetization current. We defined the magnetic moment $\vec{m}$ of a region by Equation (10.2). Note that this is purely a kinematic definition in the sense that it does not involve any mention of an actual interaction, magnetic or otherwise. This corresponds to the definition of the electric moment of a region given in Equation (10.4) which is also only a kinematic description of a specific alignment of charges, although the net charge of the volume is zero. The magnetic moment definition is a description of a system of currents which do not produce any net flow across a surface which is large enough to be accessible to macroscopic observation. For a given distribution of magnetic moment, Equation (10.2) can be solved for the magnetization current (10.3) which can then be used to compute magnetic interactions.

As an example, consider the neighboring current loops in a rectangular network in the x,y plane and let us consider the z component of the magnetic moment of these current loops as seen in Figure (10.1). If the magnetization is inhomogeneous then there will not be complete cancellation between the boundaries

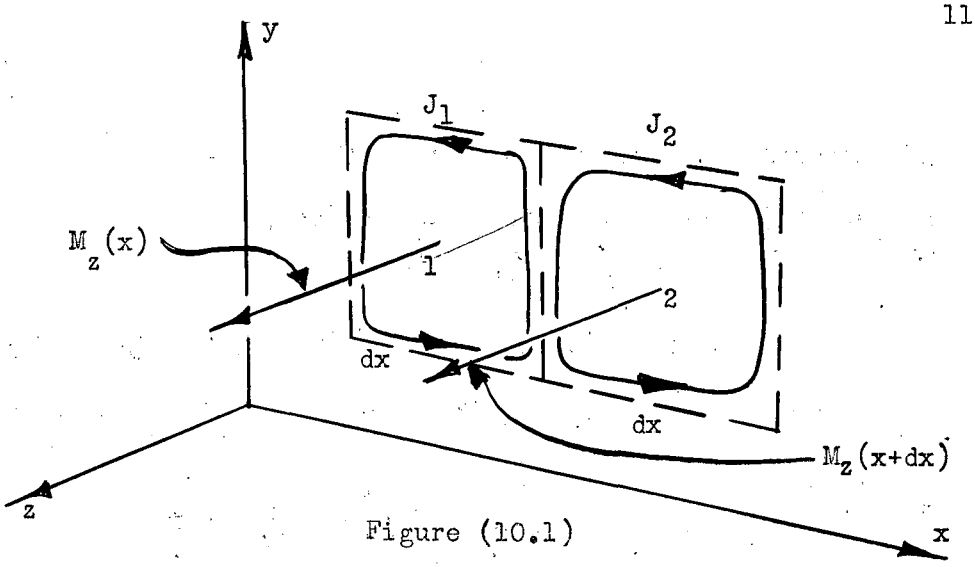


Figure (10.1)

of the loops and a net current will flow. This net current will bring about whatever net effects are ascribable to the currents. From the expression for the magnetic moment of a current loop:

J = the current in the loop.
$\vec{S}$ = the area of the loop.

(10.9) $$\vec{m} = J \vec{S}$$

and the expression for the moment of one of the loops in Figure (10.1):

(10.10) $$\vec{m} = \vec{M}\, dx\, dy\, dz$$

we have for the current in rectangle 1:

(10.11) $$J_1 = \frac{M_z\, dx\, dy\, dz}{dx\, dy}$$

and using Taylor's expansion theorem, the current in the neighboring rectangle 2 is:

(10.12) $$J_2 = \frac{(M_z + \frac{\partial M_z}{\partial x} dx)\, dx\, dy\, dz}{dx\, dy}$$

The difference between J_1 and J_2 results in a net current in the y direction along the mutual boundary of rectangle 1 and rectangle 2, that is:

(10.13) $$J_y = -\frac{\partial M_z}{\partial x}\, dx\, dz$$

This will be recognized as one of the six components of the curl and one sees that in general:

(10.14) $$\vec{j}_m = \vec{\nabla} \times \vec{M}$$

Expression (10.14) is the solution of Equation (10.2) in that it reduces it to an identity. Equation (10.14) gives the net current produced in a region of inhomogeneous magnetization. In a region of discontinuous magnetization, it is easily seen that a surface current equal to the change in the tangetial component of the magnetization will result at such a discontinuity. This follows when Equation (10.14) is applied to a limiting transverse surface bounding such a discontinuity.

Thus in a stationary medium, the total current is given by the sum of the three types of current enumerated on pages 114 and 115.

$$(10.20) \qquad \vec{j}_{total} = \vec{j}_{true} + \frac{\partial \vec{P}}{\partial t} + \vec{\nabla} \times \vec{M}$$

In order to conserve charge it is necessary that this total current obey the equation of continuity (9.1). If we take the divergence of the total current, we obtain:

$$(10.21) \qquad \vec{\nabla} \cdot \vec{j}_{total} = \vec{\nabla} \cdot \vec{j}_{true} + \vec{\nabla} \cdot \left(\frac{\partial \vec{P}}{\partial t}\right) + \vec{\nabla} \cdot (\vec{\nabla} \times \vec{M})$$

Taking the partial derivative of Equation (1.9) with respect to time, we have:

$$(10.22) \qquad \left(\frac{\partial \rho}{\partial t}\right)_{total} = k_o \vec{\nabla} \cdot \left(\frac{\partial \vec{E}}{\partial t}\right)$$

Substituting (10.21) into (9.1) we have:

$$(10.23) \qquad \vec{\nabla} \cdot \vec{j}_{true} + \vec{\nabla} \cdot \left(\frac{\partial \vec{P}}{\partial t}\right) + \left(\frac{\partial \rho}{\partial t}\right)_{total} + \vec{\nabla} \cdot (\vec{\nabla} \times \vec{M}) = 0$$

Substituting (10.22) into this and using the relation (3.3) we have:

$$(10.24) \qquad \vec{\nabla} \cdot \vec{j}_{true} + \vec{\nabla} \cdot \left(\frac{\partial \vec{P}}{\partial t}\right) + k_o \frac{\partial \vec{\nabla} \cdot \vec{E}}{\partial t} + \vec{\nabla} \cdot (\vec{\nabla} \times \vec{M}) = 0$$

$$\vec{\nabla} \cdot \left(\vec{j}_{true} + \frac{\partial \vec{D}}{\partial t} + \vec{\nabla} \times \vec{M}\right) = 0$$

The divergence of the total current (10.21) is not zero. This means that the total current is not solenoidal. However, the quantity:

$$(10.25) \qquad \vec{c} = \vec{j}_{true} + \frac{\partial \vec{D}}{\partial t} + \vec{\nabla} \times \vec{M} = \vec{j}_{total} + k_o \frac{\partial \vec{E}}{\partial t}$$

generated by adding the term $k_o \partial \vec{E}/\partial t$ to the total current is a solenoidal current. The extra term which has been added to the total current to form the solenoidal current $\vec{c}$ is known as the vacuum displacement current. The need for the addition of this term to produce a solenoidal net current vector was recognized by Maxwell. $k_o \partial \vec{E}/\partial t$ does not have the significance of a current in the sense of being the motion of charges. We shall see later that the magnetic effects of currents can only be formulated in terms of solenoidal currents, and therefore that the vacuum displacement current term must be

introduced in order to be able to apply the formulas which will be developed for the magnetic interaction of solenoidal currents to cases involving the magnetic interaction of non-stationary currents.

The geometrical significance of the solenoidal current $\vec{c}$ is that at points where there is an accumulation of charge, the current is assumed to be continuous across the discontinuity in the form of the rate of change of the field resulting from the accumulation of the charges on the boundaries of the discontinuity. As an example, a battery charging a condenser produces a closed current loop in terms of $\vec{c}$.

CHAPTER 11 THE MAGNETIC INTERACTION OF CURRENTS

The magnetic interaction of currents is best described in terms of an experimentally established interaction in vacuum that is analogous to the electrostatic Coulomb law. The mathematical generalization of the experiment of Ampère which gave the force between two current-carrying elements as seen in Figure (11.1) is given by the expression:

$$(11.1)\quad \vec{F}_2 = \frac{\mu_0}{4\pi} J_1 J_2 \oint\oint_{12} \frac{d\vec{\ell}_2 \times (d\vec{\ell}_1 \times \vec{r}_{12})}{r^3}, \quad \mu_0 = (4\pi \times 10^{-7} \text{ Henry/Meter})$$

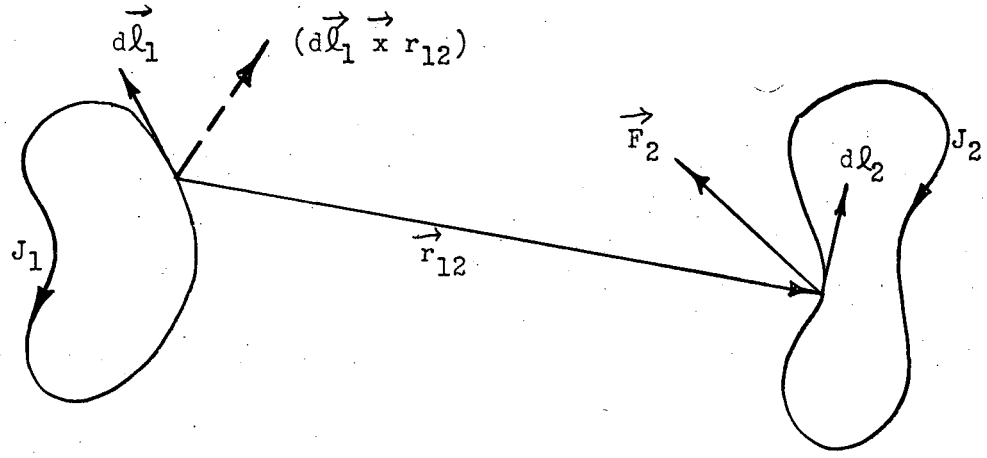


Figure (11.1)

$\vec{F}_2$ is the force on the circuit that carries the current J_2 and that has the line element $d\vec{\ell}_2$. Due to the geometry that is involved in expressing the relative directions of $\vec{F}_2$, $d\vec{\ell}_1$, $d\vec{\ell}_2$ and $\vec{r}$, this force equation appears to be more complicated than the Coulomb force equation (1.2). Also it appears, superficially, to violate Newton's third law of the equality of action and reaction. The integrand of Equation (11.1) is in fact asymmetrical as it stands. However, when the integral is carried out over two closed circuits the resulting force is symmetrical in terms of the geometry of the two interacting current loops. This can be shown as follows: If we expand the integrand

by the double vector product rule:

$$(11.2)\quad \vec{A} \times (\vec{B} \times \vec{C}) = (\vec{A} \cdot \vec{C})\,\vec{B} - (\vec{A} \cdot \vec{B})\,\vec{C}$$

we obtain:

$$(11.3)\quad \vec{F}_2 = \frac{\mu_0}{4\pi} J_1 J_2 \oint_1 \oint_2 \left\{ \frac{(d\vec{\ell}_2 \cdot \vec{r}_{12})\, d\vec{\ell}_1}{r^3} - \frac{(d\vec{\ell}_1 \cdot d\vec{\ell}_2)\, \vec{r}_{12}}{r^3} \right\}$$

Since the $d\vec{\ell}_2$ integrand in the first term is an exact differential,* this expression vanishes as the integration is carried out over closed loops. The other integral:

$$(11.4)\quad \vec{F}_2 = -\left[\frac{\mu_0}{4\pi} \oint_1 \oint_2 \frac{(d\vec{\ell}_1 \cdot d\vec{\ell}_2)\, \vec{r}_{12}}{r^3} \right] J_1 J_2$$

is symmetric in terms of loops 1 and 2.

The reason that we are taking Equation (11.1) rather than Equation (11.4) as the starting point for the discussion of the magnetic interactions is that Equation (11.1) is in such a form that the interaction expression can be separated into a field produced by loop 1 and a force exerted by this field on loop 2. Expression (11.4), on the other hand, which implicitly contains the cosine of the angle between the elementary current elements does not permit such a separation and therefore does not lead directly to a vector field formulation of magnetic interactions. The separation of Equation (11.1) into a field and a field force can be carried out by putting:

$$(11.5)\quad \vec{F}_2 = J_2 \oint_2 d\vec{\ell}_2 \times \vec{B}_2$$

where:

$$(11.6)\quad \vec{B}_2 = -\frac{\mu_0}{4\pi} J_1 \oint_1 d\vec{\ell}_1 \times \vec{\nabla}_2 \left(\frac{1}{r_{12}} \right), \qquad \vec{B}_2 = \frac{\mu_0}{4\pi} J_1 \oint_1 \frac{d\vec{\ell}_1 \times \vec{r}_{12}}{r^3}$$

$\vec{B}_2$ = the magnetic field of induction caused by circuit 1 at the position of circuit 2. (Weber/Meter^2)

$\vec{B}$ is analogous to $\vec{E}$ in the electrostatic theory in that it determines the force

---

*Being the line integral of a gradient: $\oint \vec{\nabla} \left(\frac{1}{r} \right) \cdot d\vec{\ell}_2$

that acts on a circuit element. Equation (11.6) is a generalization of the Biot and Savart law. It should be noted that thus far we have no differential form of this law. The Biot and Savart law when expressed in terms of volume currents becomes:

$$(11.7)\quad \vec{F} = \iiint (\vec{j} \times \vec{B})dv$$

$$(11.8)\quad \vec{B} = \frac{\mu_o}{4\pi}\iiint \left(\frac{\vec{j} \times \vec{r}}{r^3}\right)dv'$$

Note that $\vec{r}$ in (11.8) is directed from the point of integration or source point (where $\vec{j}$ is located) toward the field point where $\vec{B}$ is being determined.

Let us inquire under what conditions the magnetic induction field $\vec{B}$ can be derived from a scalar potential by the relation:

$$(11.9)\quad \vec{B} = -\vec{\nabla}\phi_m \cdot \mu_o$$

Let us consider a closed loop carrying a current J as in Figure (11.2)

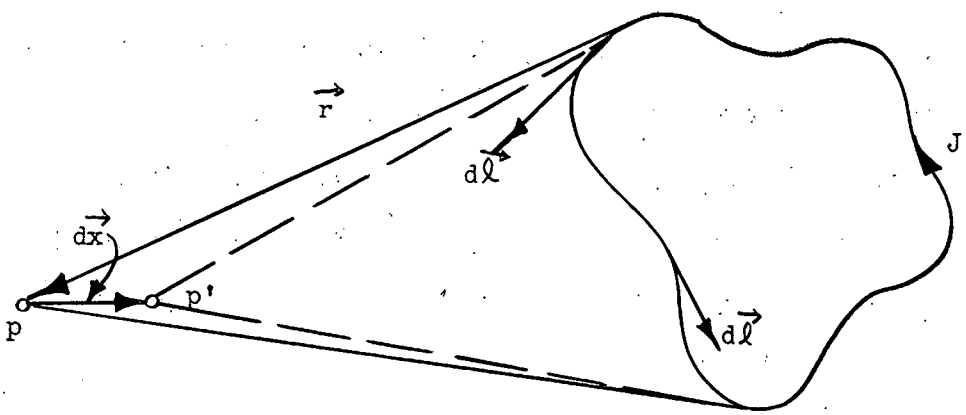


Figure (11.2)

Let the field of the current loop be measured at a point p. If the field $\vec{B}$ were derivable from a scalar magnetic potential ϕ_m and if the point of observation were moved through a distance dx from p to p', then the increment in the scalar magnetic potential ϕ_m would be given by:

$$(11.10)\quad d\phi_m = -\frac{\vec{dx} \cdot \vec{B}}{\mu_o}$$

which, using the Biot and Savart expression (11.6), becomes:

$$(11.11) \quad d\phi_m = -\frac{1}{4\pi} J \oint \frac{\vec{dx} \cdot (\vec{d\ell} \times \vec{r})}{r^3} = -\frac{1}{4\pi} J \oint \frac{\vec{r} \cdot (\vec{dx} \times \vec{d\ell})}{r^3}$$

The mixed vector-scalar product permits cyclic permutation, since it is the volume of the parallelepiped whose edges are the three vectors in the product. Equation (11.11) is of course equal to the change in the scalar magnetic potential which is obtained if the point of observation were held stationary and the loop were moved by an amount $-\vec{dx}$, as was discussed in the derivation of Equation (2.19) for the potential of a dipole sheet. Using the relation given in Equation (2.19), the change in the scalar magnetic potential can be written as:

$$(11.12) \quad d\phi_m = \frac{J}{4\pi} \, d\Omega$$

where $d\Omega$ is the change in the solid angle subtended by the loop at the point of observation brought about by an infinitesimal displacement $-\vec{dx}$ of all of the points of the loop.

This scalar potential has the same mathematical properties as the solution of the electrostatic potential of the surface dipole layer discussed earlier, as seen from the similarity of Equation (2.19) for the static potential of a dipole and Equation (11.12) for the magnetic induction of a current loop. In addition this means that the scalar potential of a current loop is multiple-valued in the sense that it appears to undergo a discontinuity of magnitude J when a surface bounded by the loop is crossed. In the case of the electric dipole sheet this surface has a physical significance, however, in the magnetic case, this surface can be chosen in any arbitrary position. Since the choice of the surface is arbitrary the magnetic field derived from such a potential outside the current-carrying region is non-ambiguous. However, line integrals of the magnetic field of a current loop will be correct only if the path of integration does not pass through the arbitrary surface. The line integral of the magnetic field of induction $\vec{B}$ around a closed path threading the current J is exactly

equal to the magnitude of the discontinuity in the magnetic scalar potential ϕ_m across the arbitrary reference surface, and hence we have:

(11.13) $$\oint \vec{B} \cdot d\vec{\ell} = \mu_o I_{total}$$

For a graphical representation of the magnetic scalar potential of a current loop, see Figure (2.5A). The dipole layer of Figure (2.5A) corresponds to the arbitrarily located surface of discontinuity of the current loop.

Equation (11.13) is the integral representation of the differential relation that gives the total circulation of the magnetic field vector in terms of the current that causes the magnetic field. Since (11.13) is valid for any arbitrary closed path of integration, we can convert it into a differential expression by substituting (11.13) into Stokes' theorem (1.14) and reducing the size of the surface of integration to a differential, thus securing:

(11.14) $$\vec{\nabla} \times \vec{B} = \mu_o \vec{j}$$

We conclude that $\vec{B}$ cannot in general be derived from a single-valued magnetic scalar potential. The concept of the magnetic scalar potential is of practical utility provided it is only used to derive magnetic fields in the absence of continuous current distributions. However, the magnetic scalar potential ϕ_m cannot be used if line integrals are considered, which loop current-carrying regions, or if the field within current-carrying media are being considered.

Equation (11.14) defines the circulation density of the magnetic field of induction at a particular point in terms of the current density at that point. In order to completely define a vector field it is necessary to specify not only the circulation density but also the source density of the field. This means that the value of the $\vec{\nabla} \cdot \vec{B}$ must be ascertained. To obtain its value we must make assumptions as to the nature of the sources of the magnetic field. If we assume that currents are fundamentally the only sources of the magnetic field, and that the magnetic field of such currents is only given by the law of Biot-Savart (11.6) and (11.8), then we shall show that:

(11.15) $$\vec{\nabla} \cdot \vec{B} = 0$$

This will follow immediately from vector considerations if we can show that the specific form of the magnetic field given by the Biot-Savart law permits $\vec{B}$ to be derived from a vector potential $\vec{A}$ by taking the curl:

$$(11.16) \qquad \vec{B} = \vec{\nabla} \times \vec{A}$$

If we re-express the law of Biot-Savart (11.8) as:

$$(11.17) \qquad \vec{B} = \frac{\mu_o}{4\pi} \iiint \frac{\vec{j}' \times \vec{r}}{r^3} dv' = -\frac{\mu_o}{4\pi} \iiint \vec{j}' \times \vec{\nabla}\left(\frac{1}{r}\right) dv'$$

The operator $\vec{\nabla}$ in (11.17) does not operate on the variable of integration but on the variable of field position so we can take the $\vec{\nabla}$ operator outside of the integration sign. The current vector $\vec{j}'$ is a function only of the variables of integration so $\vec{\nabla}(1/r)\times\vec{j}' = \vec{\nabla}\times(\vec{j}'/r)$. This operation thus gives:

$$(11.18) \qquad \vec{B} = \frac{\mu_o}{4\pi} \vec{\nabla} \times \iiint \frac{\vec{j}' dv'}{r}$$

Hence $\vec{B}$ is actually in the form of Equation (11.16) if we let:

$$(11.19) \qquad \vec{A} = \frac{\mu_o}{4\pi} \iiint \frac{\vec{j}' dv'}{r}$$

and hence (11.15) holds for all current-produced fields as a vector consequence of (11.16). The explicit expression for the vector potential $\vec{A}$ in terms of the current is therefore correctly given by (11.19). The expression for the vector potential of a linear current distribution corresponding to the expression for the vector potential of a continuous current distribution given in (11.19) is:

$$(11.20) \qquad \vec{A} = \frac{\mu_o}{4\pi} J \int \frac{d\vec{\ell}}{r}$$

The fields produced by currents can therefore be computed by first computing the vector potential $\vec{A}$, using (11.19) or (11.20) and then obtaining the magnetic field by the relation (11.16). The vanishing of the divergence of $\vec{B}$ follows from (11.16).★

★Equations (11.14) and (11.15) can also be derived by direct differentiation of Eq. (11.8):

$$\vec{\nabla} \times \vec{B} = -\frac{\mu_o}{4\pi} \iiint \vec{\nabla} \times \left[\vec{j} \times \vec{\nabla}\left(\frac{1}{r}\right)\right] dv' = -\frac{\mu_o}{4\pi} \iiint \left\{\vec{j}\left[\vec{\nabla}^2\left(\frac{1}{r}\right)\right] - (\vec{j}\cdot\vec{\nabla})\vec{\nabla}\left(\frac{1}{r}\right)\right\} dv'$$

$= \mu_o \vec{j}$ since $\nabla^2\left(\frac{1}{r}\right) = -4\pi\delta(r)$ and the second term vanishes if $\vec{\nabla}'\cdot\vec{j} = 0$.

$$\vec{\nabla} \cdot \vec{B} = -\frac{\mu_o}{4\pi} \iiint \vec{\nabla} \cdot \left[\vec{j} \times \vec{\nabla}\left(\frac{1}{r}\right)\right] dv' = \frac{\mu_o}{4\pi} \iiint \vec{j} \cdot \left\{\vec{\nabla} \times \left[\vec{\nabla}\left(\frac{1}{r}\right)\right]\right\} dv' = 0$$

CHAPTER 12 MAGNETIC MATERIALS

Thus far only magnetic fields in a vacuum have been treated. If material media are introduced into the magnetic fields, then the current $\vec{j}$ in Equation (11.14) must be replaced by the total current including the magnetization and the polarization currents, as given in Equation (10.20). Since we have shown that in vacuo that $\vec{\nabla} \times \vec{B}$ is proportional to the total stationary current density, the relation $\vec{\nabla} \times \vec{B} = \mu_o \vec{j}$ can only be generalized to non-stationary cases if either the current remains solenoidal or if the relation used in deriving the magnetic field from the current is modified. The choice between these alternatives made by Maxwell was to retain the relations that derive the magnetic field from the current, (11.6), (11.8) or (11.18) but to use, in the general case of non-stationary fields, the general current $\vec{c}$ in Equation (10.25), which includes the displacement current and which remains solenoidal. The total current $\vec{c}$ for media at rest, in terms of all of the components, is thus given by:

$$(12.1)\quad \vec{c} = \vec{j}_{true} + \vec{\nabla} \times \vec{M} + \frac{\partial \vec{D}}{\partial t} = \vec{j}_{true} + \frac{\partial \vec{P}}{\partial t} + \vec{\nabla} \times \vec{M} + k_o \frac{\partial \vec{E}}{\partial t}$$

and hence for stationary media the equations defining the vector field $\vec{B}$ are:

$$(12.2)\quad \vec{\nabla} \cdot \vec{B} = 0$$

and using $\vec{c}$ for $\vec{j}$ in (11.14) we have:

$$(12.3)\quad \vec{\nabla} \times \vec{B} = \mu_o \left(\vec{j}_{true} + \vec{\nabla} \times \vec{M} + \frac{\partial \vec{D}}{\partial t} \right) = \mu_o \left(\vec{j}_{true} + \frac{\partial \vec{P}}{\partial t} + \vec{\nabla} \times \vec{M} + k_o \frac{\partial \vec{E}}{\partial t} \right)$$

Note that each of the four terms on the righthand side of (12.3) has a distinct physical meaning. The polarization current term $\frac{\partial \vec{P}}{\partial t}$ and the term $k_o \frac{\partial \vec{E}}{\partial t}$ have a superficial similarity and their sum is equal to the displacement current $\frac{\partial \vec{D}}{\partial t}$. However, the polarization current represents a space-time average of actual charge motions within a polarized medium, while the term $k_o \frac{\partial \vec{E}}{\partial t}$ is a mathematical supplement which is necessary to bring Equation (12.3) and (10.23) into agreement with the equation of continuity and does not represent a current having a physical reality

in the sense of charge motion. The similarity of these two terms has led to a great deal of speculation as to whether some motion of charge might not actually be involved in the vacuum displacement current $k_o\frac{\partial \vec{E}}{\partial t}$, but such charge motion could only be measured by its magnetic effect and it is therefore unnecessary to introduce such charges as a physical concept. As will be shown later, by relativistic considerations, it is not possible to measure any properties of the field-carrying medium by physical means, and therefore one cannot ascribe any material significance to the vacuum displacement current term.

In the treatment of the polarization of dielectrics in Chapter 3 it was found to be mathematically convenient to separate the field whose sources were true charges only, from the total field whose sources were the true charges plus the polarization charges. In a similar manner it is convenient to separate the magnetic field whose circulation density arises from true currents only, from the total field whose circulation density arises from true currents plus atomic magnetization currents. Therefore if we write Equation (12.3) in the form:

$$(12.4)\quad \vec{\nabla} \times (\vec{B} - \mu_o \vec{M}) = \mu_o\left(\vec{j}_{true} + \frac{\partial \vec{D}}{\partial t}\right)$$

and define a new field $\vec{H}$ by:

$$(12.5)\quad \vec{H} = \frac{1}{\mu_o}(\vec{B} - \mu_o \vec{M}) = \frac{\vec{B}}{\mu_o} - \vec{M}$$

$\vec{H}$ = the magnetic field intensity: (Ampere-Turn/Meter)

then (12.4) will reduce to:

$$(12.6)\quad \vec{\nabla} \times \vec{H} = \vec{j}_{true} + \frac{\partial \vec{D}}{\partial t}$$

Equation (12.6) means that the circulation density of $\vec{H}$ arises from the true current plus the total displacement current $\frac{\partial \vec{D}}{\partial t}$. By the total displacement we mean the sum of the polarization current $\frac{\partial \vec{P}}{\partial t}$ and the vaccum displacement current $k_o\frac{\partial \vec{E}}{\partial t}$. Under stationary conditions or quasi-stationary conditions* we have:

---

*Here we consider quasi-stationary conditions to be those in which the magnetic effect of the displacement current is negligible compared to the magnetic effect of the true current.

(12.7) $\frac{\partial \vec{D}}{\partial t} = 0$, $\qquad \vec{\nabla} \times \vec{H} = \vec{j}_{true}$

and in integral form:

(12.8) $\oint \vec{H} \cdot d\vec{\ell} = J_{true}$

Note that in the sense of the separation between the effect that is produced by the total and the true charges and the total and the true currents respectively that $\vec{B}$ plays a role that corresponds to $\vec{E}$ while $\vec{H}$ plays a role that corresponds to $\vec{D}$, as we can see by comparing Equations (11.13) and (12.8) with:

(12.9) $\iint \vec{E} \cdot d\vec{S} = \frac{q_{total}}{k_o}$

$\iint \vec{D} \cdot d\vec{S} = q_{true}$

The discussion of dielectrics in Chapter 3 was limited to the case of linear media, namely media in which the polarization was proportional to the applied electric field. In ferro-magnetic substances, however, the case of nonlinear behavior is most common and therefore we must discuss some of the properties of the magnetic field which arise in cases when the magnetization $\vec{M}$ is not a linear and often not even a unique function of the external fields. At first we will assume $\vec{M}$ to be a given function of the material medium independent of external fields. In the most extreme case, that of a permanent magnet, there will be a magnetic moment $\vec{M}$ per unit volume, even in the absence of any true currents. In this case, as we see from (12.7), $\vec{H}$ will be irrotational and will therefore behave mathematically like an electrostatic field, while $\vec{B}$ remains, of course, solenoidal:

(12.10) $\vec{\nabla} \times \vec{H} = 0, \vec{\nabla} \times \vec{B} = \mu_o \vec{\nabla} \times \vec{M} \neq 0$

However, the magnetic field will have sources as we see by taking the divergence of (12.5) and using (12.2):

(12.11) $\vec{\nabla} \cdot \vec{H} = - \vec{\nabla} \cdot \vec{M} = \rho_m$, $\qquad \rho_m$ = magnetic source density

The "magnetostatic field" $\vec{H}$ can therefore be derived from a magnetic source

density ρ_m which is equal to the negative divergence of the magnetization. One unit of this equivalent magnetic charge density is usually known as a magnetic pole. In terms of this description a magnetic pole has no physical reality other than that the mathematical description of the resultant magnetic field of a permanent magnet is formally the same as the mathematical description of the resultant electric field of electric charges. Since the magnetic field $\vec{H}$ of a permanent magnet is irrotational, it can be derived from a magnetic scalar potential ϕ_m in the same way that $\vec{E}$ may be derived from the electrostatic potential ϕ. If we put:

$$(12.12) \quad \vec{H} = -\vec{\nabla}\phi_m$$

then the resultant scalar potential, in terms of the equivalent volume and surface pole densities, is given by:

$$(12.13) \quad \phi_m = \frac{1}{4\pi}\left[\iint \frac{\vec{M}\cdot d\vec{S}}{r} - \iiint \frac{\vec{\nabla}\cdot\vec{M}}{r}\, dv\right]$$

The field of a permanent magnet of a given magnetization can also be described by a vector potential $\vec{A}$ which is derived from the equivalent surface currents and volume currents within the magnetized body.

We saw that the surface current at a boundary equivalent to the magnetization is given by the tangential component of the magnetization, $-\vec{n}_1 \times \vec{M}$, and the equivalent volume current is given by the curl of the magnetization, as was shown in Equation (10.14). From (11.19) the vector potential of a magnetized body is thus given by:

$$(12.14) \quad \vec{A} = \frac{\mu_o}{4\pi}\left[\iiint \frac{\vec{\nabla}\times\vec{M}}{r}\, dv - \iint \frac{\vec{n}_1\times\vec{M}}{r}\cdot d\vec{S}\right]$$

from which the magnetic induction field can be derived by the use of (11.16).

In the case of a uniformly magnetized medium, all internal currents cancel and hence the equivalent surface currents are the only ones present. A cylindrical magnet magnetized in a direction parallel to the axis of the cylinder therefore has a magnetic field equivalent to the field of a solenoidal coil, carrying current

on the cylindrical face of the magnet with the current flow lines lying in planes normal to the axis of the cylinder. This situation can be described qualitatively by noting that for a permanent magnet, $\vec{H}$ can be thought of as arising from a layer of equivalent pole charges located on the magnet pole faces in the same manner as an electrostatic field would be formed by charges so placed. On the other hand, $\vec{B}$ arises from an equivalent solenoid which can be thought of as being wound on the cylindrical surface of the magnet in the same manner as a vacuum current field arises. $\vec{B}/\mu_0$ and $\vec{H}$ are identical outside of the region where $\vec{M}$ has a finite value, but they differ by $\vec{M}$ inside of the magnet. Note that $\vec{B}$ and $\vec{H}$ in Figure (12.1) are actually in opposite directions inside the magnet, as is obvious from the fact that the line integral of $\vec{H}$ must be zero around any closed path.

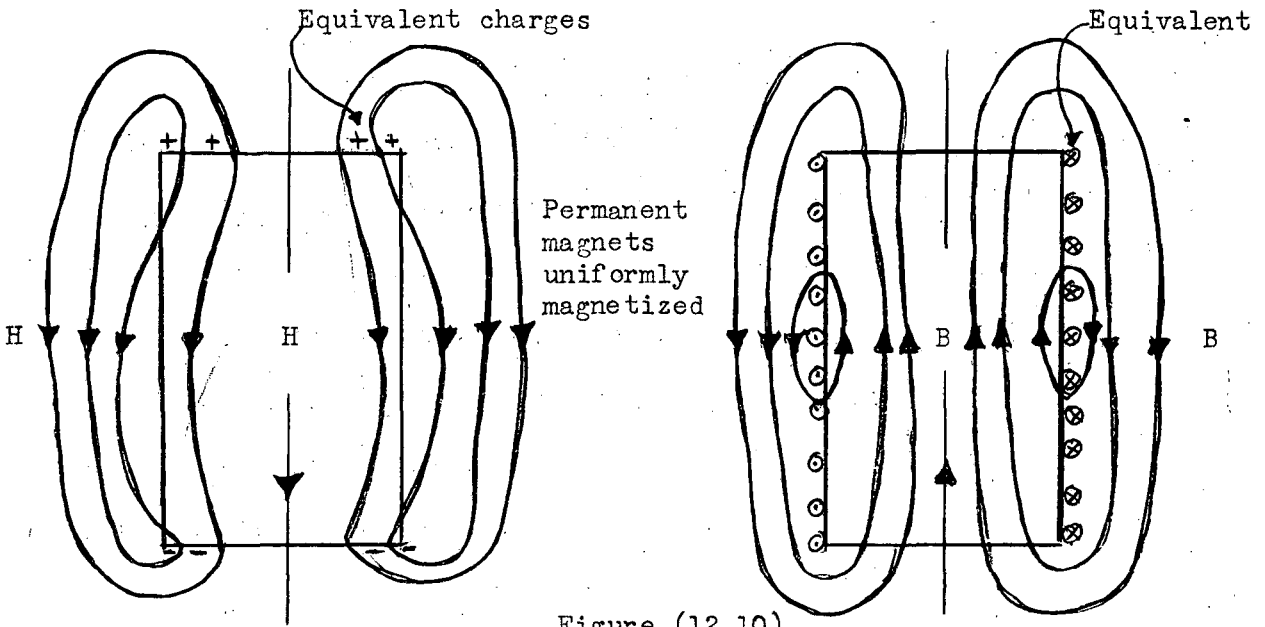


Figure (12.10)

($\vec{B}$ and $\vec{H}$ fields of a permanent magnet)

Equations (12.13) and (12.14) describe scalar and vector magnetic potentials in terms of the equivalent pole or current distributions. The potentials can, of course, be described in terms of the integral over the potentials of the individual magnetic moments themselves. If we start from the vector and

scalar potentials of a current loop, we will then obtain the expression:

$$(12.15) \qquad \vec{A} = \frac{\mu_0}{4\pi} \iiint \vec{M} \times \vec{\nabla}\left(\frac{1}{r}\right) dv'$$

and

$$(12.16) \qquad \phi_m = -\frac{1}{4\pi} \iiint \vec{M} \cdot \vec{\nabla}\left(\frac{1}{r}\right) dv'$$

By means of an integration by parts these expressions can be shown to be equivalent to (12.13) and (12.14). The fields derived from either expression must of course be the same. We can show, be vector identities, that the field due to a magnetic dipole of moment $\vec{m} = \iiint \vec{M}\, dv'$ is:

$$(12.17) \qquad \vec{B} = \vec{\nabla} \times \vec{A} = \frac{\mu_0}{4\pi} \left[\vec{\nabla} \times \left\{\vec{m} \times \vec{\nabla}\left(\frac{1}{r}\right)\right\}\right] = -\frac{\mu_0}{4\pi} (\vec{m} \cdot \vec{\nabla})\vec{\nabla}\left(\frac{1}{r}\right) - \vec{m}\nabla^2\left(\frac{1}{r}\right)$$

The last term vanishes except at r = 0. From (12.16):

$$(12.18) \qquad \vec{H} = -\vec{\nabla}\phi_m = \frac{1}{4\pi} \left[\vec{\nabla} \left\{\vec{m} \cdot \vec{\nabla}\left(\frac{1}{r}\right)\right\}\right] = -(\vec{m} \cdot \vec{\nabla})\vec{\nabla}\left(\frac{1}{r}\right) (4\pi)^{-1}$$

The two fields differ only at r = 0.

Thus we have seen that permanent magnets may be described equivalently either in terms of "equivalent currents" or "equivalent poles." Since the entire description of magnetic fields has been based on the premise that they are produced by moving charges,* we are led to believe that the interpretation of the field

*There exists no basic objection to the existence of magnetic poles; their fields are simply not considered here since there is no experimental evidence as to their existence. If single magnetic poles did exist, all the above equations would have to be supplemented. It can be shown quantum mechanically that if magnetic poles did exist, the magnitude of the "elementary" unit pole would have to be related to the elementary charge by a constant factor.

of a permanent magnet in terms of the circulation of atomic currents is a more fundamental one than the concept of magnetic charges, and that therefore $\vec{B}$ which arises from currents is a more fundamental field than $\vec{H}$ which arises from "magnetic charges." However, the description in terms of $\vec{H}$ is more attractive from a practical point of view, since it reduces problems that involve permanent magnets or problems involving magnetized pieces of iron whose magnetization can be determined by other means, to problems in electrostatics.

The question mentioned above as to whether $\vec{B}$ or $\vec{H}$ is basically the more fundamental field can be formulated in a different way. If we ask the question, "consider a charge q moving with a velocity $\vec{v}$ in a magnetized medium and let us suppose that the force acting on it is of the form $\vec{F} = q(\vec{v} \times \vec{X})$. Then should we use $\vec{B}$ or $\vec{H}$ for $\vec{X}$, or use a combination of them?" This question has been tested experimentally by Rassetti by measuring the deflection of cosmic rays in magnetized iron, and has been studied theoretically by Wannier by analyzing in detail the motion of charged particles in magnetized media. The answer is essentially this: If the motion of the charged particles is truly random relative to the magnetized material, that is, it is not affected by the presence of the magnetized medium, to a first approximation, then the force that is exerted on a charged particle corresponds to the use of $\vec{B}$ as the magnetic field in the force equation. If, on the other hand, the particle is moving slowly and its motion is substantially affected by the magnetized medium, then the particle is effectively prevented from passing through the insides of the equivalent atomic current loops and in this case, since the individual current loops act like impenetrable dipoles, the averaging process favors a deflection that corresponds to the use of $\vec{H}$ in the force equation. Rassetti's experimental results actually indicate that the deflection for very high speed particles corresponded approximately to the use of $\vec{B}$ in the force equation. In order for the results to correspond to the use of $\vec{H}$ it would have been necessary for the deflection to have been in the opposite direction. The answer to the above question can be given precisely in the limits: $\vec{X}_{(v \to c)} = \vec{B}$.

Thus far we have considered the case of magnetic media where the magnetization $\vec{M}$ is a given function of position as in a permanent magnet. We must now investigate the case in which we have an ideally permeable medium, that is a medium which has no magnetic moment in the absence of external true currents and which gives rise

to a magnetic moment which is proportional to the field that is produced by the external true currents. The field equations are then:

(12.2) $\vec{\nabla} \cdot \vec{B} = 0$

(12.8) $\oint \vec{H} \cdot d\vec{\ell} = J_{true}$

and if we assume that $\vec{M} = \chi_m \vec{H}$ we get a relation that corresponds to Equation (3.11) in the discussion of electrostatics:

(12.19) $\vec{B} = \mu \mu_o \vec{H}$ $\qquad \mu = \chi_m + 1$ = the permeability

χ_m = the magnetic susceptibility

By means of a derivation which is completely analogous to the one used in Chapter 3, to derive the boundary conditions for $\vec{E}$ and $\vec{D}$, the boundary conditions for $\vec{B}$ and $\vec{H}$ for linear media may be shown to be:

(12.20) (Normal Components)

$$\vec{n}_1 \cdot (\vec{B}_2 - \vec{B}_1) = \vec{n}_1 \cdot \mu_o(\mu_2 \vec{H}_2 - \mu_1 \vec{H}_1) = 0$$

(12.21) (Tangential Components) $\vec{K}$ = the true surface current on the boundary between the two media.

$$\vec{n}_1 \times (\vec{H}_2 - \vec{H}_1) = \vec{n}_1 \times (\vec{\nabla}\phi_{m_1} - \vec{\nabla}\phi_{m_2}) = \frac{\vec{n}_1}{\mu_o} \times \left(\frac{\vec{B}_2}{\mu_2} - \frac{\vec{B}_1}{\mu_1} \right) = \vec{K}$$

Note that the equations (12.2), (12.8) and (12.19) are mathematically identical to the equations governing stationary current flow in a continuous medium in the presence of a non-conservative electromotive force:

(9.2) $\vec{\nabla} \cdot \vec{j} = 0$

(9.3) $\vec{j} = \sigma \vec{E}$

(9.5) $\oint \vec{E} \cdot d\vec{\ell} = \mathcal{E}$

which led to the expression:

(9.17) $R = \frac{\ell}{\sigma S} = \sum_i \frac{\ell_i}{\sigma S_i}$

for the "resistance" of linear conductors in series. This analogy gives rise to the concept of the magnetic circuit, namely the solution of linear magnetic media

problems given by:

$$(12.23)\quad \Phi_m = \iint \vec{B} \cdot d\vec{S} = \frac{J}{R_m} \quad , \qquad R_m = \text{the magnetic "reluctance" (Amp/Weber) of the circuit.}$$

$$(12.24)\quad R_m = \sum_i \frac{\ell_i}{\mu \mu_o S_i}$$

Note that this solution is based only on the correspondence of the differential equations for linear magnetization problems to the differential equations for steady current problems, and that the solutions themselves will actually correspond only in case the boundary conditions for the magnetic and current problems are identical. This cannot be true in general. In fact, it can never be completely accurate since the conductivity of free space is zero, while the permeability of free space is unity. This means that the magnetic circuit solution will only be valid if the permeability of the media being considered is quite high compared to 1, or at least if the regions of space that are accessible to the magnetic field in which the permeability is comparable to 1, are small compared to those regions in which the permeability is much larger than 1. The set of solutions (12.23) and (12.24) do, however, form the basis of industrial magnetic machinery design, since they permit an approximate treatment in cases where direct boundary value solutions are impractical.

CHAPTER 13 MAGNETIC BOUNDARY VALUE PROBLEMS

In general, boundary value problems in the presence of magnetic media can be attacked either by the use of the magnetic scalar or the vector potential. Problems involving magnetic media, located in external fields, where true currents do not enter the region of interest, are best treated by the use of magnetic scalar potentials. In this case, the magnetic boundary conditions (Equations (12.20) and (12.21)) expressed in terms of the magnetic scalar potential are analogous to the electrostatic boundary conditions given in Chapter 3 if the relative permeability replaces the specific inductive capacity, and the absence of any quantity corresponding to true surface charge and the possible presence of a true surface current are taken into account. We have, for the boundary conditions in linear media:

(Normal Components)

$$(13.1)\quad \vec{n}_1 \cdot (\vec{D}_2 - \vec{D}_1) = \vec{n}_1 \cdot (k_2 k_o \vec{E}_2 - k_1 k_o \vec{E}_1) = \sigma \qquad (3.15)$$

$$(13.2)\quad \vec{n}_1 \cdot \mu_o (\mu_2 \vec{H}_2 - \mu_1 \vec{H}_1) = \vec{n}_1 \cdot (\vec{B}_2 - \vec{B}_1) = 0 \qquad (12.20)$$

(Tangential Components)

$$(13.3)\quad \vec{n}_1 \times \left(\frac{\vec{D}_2}{k_2 k_o} - \frac{\vec{D}_1}{k_1 k_o} \right) = \vec{n}_1 \times (\vec{E}_2 - \vec{E}_1) = 0 \qquad (3.19)$$

$$(13.4)\quad \vec{n}_1 \times (\vec{H}_2 - \vec{H}_1) = \frac{\vec{n}_1}{\mu_o} \times \left(\frac{\vec{B}_2}{\mu_2} - \frac{\vec{B}_1}{\mu_1} \right) = \vec{K} \qquad (12.21)$$

where $\vec{K}$ is a surface boundary current.

For example problems involving magnetic shields can be treated by electrostatic boundary value problem methods. The only additional complication which enters in the magnetic problem is the fact that in practical cases the permeability μ for a particular material is not constant. This is especially true in cases when the flux density within the permeable medium becomes high. In such cases of high flux density or saturation the method of successive approximations may be used. In this method a solution based on the assumption that μ is constant is

first obtained. Then the resultant flux density is computed and the permeability is secured from it. The problem is then repeated using the new permeability. This method will give an accurate solution to problems where the field inside the permeable medium turns out to be uniform. The problem of an ellipsoid of magnetic material situated in a uniform magnetic field is such a problem. For this reason, the torque acting on an ellipsoid suspended in a uniform field can be used as a measurement of the permeability of the material of the ellipsoid as a function of the external field. Problems of the behavior of permeable media in high fields where the resultant magnetization is appreciably non-uniform are essentially impossible to treat by purely analytical methods.

Problemsin which currents are present must be treated by the use of the vector potential unless it is possible to introduce an equivalent dipole sheet in place of the current. The vector potential obeys uniqueness conditions that are similar to those obeyed by the scalar potential. We can show this by a proof that is very similar to the proof used in Chapter 4 for the uniqueness of the scalar potential. We will use the vector form of Green's theorem or Gauss' divergence theorem in our proof. If we substitute:

(13.5) $\vec{V} = \vec{U} \times (\vec{\nabla} \times \vec{W})$ $\vec{U}$, $\vec{V}$ and $\vec{W}$ are arbitrary vector functions.

into Gauss' divergence theorem (1.7):

$$(13.6) \quad \iiint \vec{\nabla} \cdot \vec{V} \, dv = \iint \vec{V} \cdot d\vec{S}$$

we obtain:

$$(13.7) \quad \iiint \left[(\vec{\nabla} \times \vec{W}) \cdot (\vec{\nabla} \times \vec{U}) - \vec{U} \cdot \vec{\nabla} \times (\vec{\nabla} \times \vec{W}) \right] dv = \iint \vec{U} \times (\vec{\nabla} \times \vec{W}) \cdot d\vec{S}$$

Let us now consider a region in space that is bounded by the surface S and has a volume v, within which there is no current flow. In order to accomplish this it may be necessary to choose subsurfaces which will exclude the regions of current flow. Thus, for all points within v, if we let:

$$(13.8) \quad \vec{j}_{total} = 0 \quad , \vec{W} = \vec{A} \quad , \vec{U} = \vec{A}$$

and using the field equations:

$$(13.9)\quad \vec{\nabla} \times \vec{B} = \mu_o \vec{j}_{total} \quad , \quad \vec{B} = \vec{\nabla} \times \vec{A} \quad , \quad \vec{\nabla} \times \vec{\nabla} \times \vec{A} = \mu_o \vec{j}_{total}$$

Substituting into Equation (13.7) we obtain:

$$(13.10)\quad \iiint (\vec{\nabla} \times \vec{A})^2 dv = \iint (\vec{A} \times \vec{B}) \cdot \vec{dS} = \iint (\vec{A} \times \vec{B} \cdot \vec{n_1})\, dS$$

which can be put into the form:

$$(13.11)\quad \iiint (\vec{\nabla} \times \vec{A})^2 dv = \iint (\vec{B} \times \vec{n_1}) \cdot \vec{A} dS = \iint \vec{B}_t \cdot \vec{A}\, dS,$$

$\vec{B}_t$ = the tangential component of $\vec{B}$ parallel to the surface and perpendicular to $\vec{S}$.

Now let us assume that $\vec{A}$ in (13.10) and (13.11) represents the difference between alternative solutions corresponding to the same boundary values of either the tangential component of the magnetic field of induction $\vec{B}$, or the vector potential $\vec{A}$. The righthand side of (13.11) then vanishes since it is evaluated on the boundary where the alternate solutions are equal. On the other hand, the left hand side of (13.11) is positive definite and hence its integrand must vanish. So $\vec{\nabla} \times \vec{A}$ must be zero throughout v and hence the field $\vec{B} = \vec{\nabla} \times \vec{A}$ is unique.

This means that the tangential component of the magnetic field or the value of the vector potential on the surface S uniquely defines the magnetic field within the volume bounded by this surface. This is equivalent to the analogous electrostatic consideration in which the value of the scalar potential on the bounding surface or the value of the normal electrostatic field defines the electrostatic field in the volume bounded by this surface. It is possible of course to carry the analogy of this procedure still further by writing the vector potential within v explicitly in terms of the currents $\vec{j}$ within v and the boundary values of the field over S which are chosen such as to make the field outside of S equal to zero. The surface terms will then correspond to the complementary solution of the differential equation while the volume integral of the currents will correspond to the particular integral of the differential equation. We shall

not carry out the details of this process here.* We shall, however, obtain the particular integral which can be made equal to the general solution in case we let our boundary expand to infinity in a manner similar to the electrostatic case, that is, the solution which corresponds to knowing the sources over all space. The differential equation whose particular integral is to be obtained must first be derived from the field equations (13.9) by the use of a vector identity for the double curl. This gives us:

$$(13.12)\quad \vec{\nabla} \times \left(\vec{\nabla} \times \vec{A}\right) = \vec{\nabla}\left(\vec{\nabla} \cdot \vec{A}\right) - \nabla^2 \vec{A} = \mu_o \vec{j}$$

Some care must be used in the interpretation of the operation of the symbol ∇^2 when it is applied to a vector. In a Cartesian coordinate system $\nabla^2 \vec{A}$ means a vector whose $i^{\underline{th}}$ component is $\nabla^2 \vec{A}_i$. In a non-Cartesian coordinate system it means a vector whose $i^{\underline{th}}$ component must be evaluated by the use of the identity:

$$(13.13)\quad \nabla^2 \vec{A} = -\vec{\nabla} \times \left(\vec{\nabla} \times \vec{A}\right) + \vec{\nabla}\left(\vec{\nabla} \cdot \vec{A}\right)$$

The choice of $\vec{\nabla} \cdot \vec{A}$ thus far has been left arbitrary since $\vec{A}$ was only defined in terms of the equation $\vec{B} = \vec{\nabla} \times \vec{A}$. It is here convenient to take:

$$(13.14)\quad \vec{\nabla} \cdot \vec{A} = 0$$

This does not involve any new physical assumptions. We shall find later when we are considering non-stationary currents that a more complicated expression must be substituted for (13.14) in order to preserve symmetry between the electric and the magnetic case, and in the more general application to obtain relativistic covariance of the resulting equations. Assumption (13.14) reduces (13.13) to:

$$(13.15)\quad \nabla^2 \vec{A} = -\mu_o \vec{j}$$

This is the vector form of Poisson's equation. The particular integral of the scalar Poisson equation is the Coulomb potential, and since (13.15) is the superposition of three scalar Poisson equations, the particular integral of (13.15) is:

$$(13.16)\quad \vec{A} = \frac{\mu_o}{4\pi} \iiint \frac{\vec{j}\, dv'}{r}$$

---

* Stratton, "Electromagnetic Theory", McGraw-Hill, 1941.

which becomes the general solution if the integral extends over all of the currents that contribute to the field.

The solution of (13.15), subject to arbitrary boundary conditions, is usually considerably more complicated than that of the corresponding scalar potential equation (1.16). The reason for this is that $\vec{A}$ actually does not have three independent components, but only two independent components, which arises from the restriction of (13.14). This means that we cannot expect to expand the components of $\vec{A}$ in normal orthogonal functions and then have a sufficient number of boundary conditions to determine all of the coefficients since there are too many functions to be fitted.

There are several methods which can be used to get around this difficulty and solve the equations:

$$(13.17)\quad \vec{\nabla} \times (\vec{\nabla} \times \vec{A}) = 0 \quad , \quad \vec{\nabla} \cdot \vec{A} = 0$$

We shall mention one, which is more fully discussed by Smythe.*

For example, if we express $\vec{A}$ in terms of two scalar functions U and W, each of which obeys Laplace's equation, by the relation:

$$(13.18)\quad \vec{A} = \vec{\nabla} U + \vec{\nabla} \times (\vec{a}_1 W) \qquad (\vec{a}_1 \text{ is any unit vector})$$

then equation (13.14) is automatically satisfied by (13.18), as can be verified by inspection. The number of necessary boundary conditions is thus reduced from 6 to 4 and orthogonal expansions can be used for U and W.

In two dimensional problems, a simple use can be made of the vector potential. In a two-dimensional problem we assume that the fields are not functions of the z coordinate. If also, all current flow is parallel to the z axis, it follows from (13.16) that the vector potential has only a z component. If this is so, then the magnetic fields are derived from the vector potential by the equations:

$$(13.19)\quad B_x = \frac{\partial A_z}{\partial y} \quad , \quad B_y = -\frac{\partial A_z}{\partial x}$$

---

*Smythe, "Static and Dynamic Electricity," McGraw-Hill, 1939.

and Laplace's equation becomes:

$$(13.20)\quad \nabla^2 A_z = 0$$

Equation (13.14) is then satisfied since $\vec{A}$ has only the z component A_z which is not a function of z. Since the z component of the vector potential obeys the two dimensional Laplace's equation (13.20) we can make it either the real or the imaginary part of a complex potential. The only difference between this case and the analogous two-dimensional electrostatic case that was discussed in Chapter 5 is that the z component of the vector potential A_z corresponds to the stream function Ψ in the electrostatic case. That is, Equations (13.19) and (13.20) are the same mathematically as the equations* which related the stream function to the

---

*These equations are (5.3) and the relation:

$$\left(\frac{dy}{dx}\right)_{\psi} = -\frac{\partial \Psi}{\partial x}\Big/\frac{\partial \Psi}{\partial y}$$

found in the footnote on page 52. It may be seen by dividing B_y by B_x in (13.19) to get the slope of a $\vec{B}$ line, that the $\vec{B}$ line corresponds to a Ψ = constant line or a stream line.

---

corresponding electrostatic field. Since the form of the Coulomb potential for each rectangular component of the vector potential is the same as the scalar Coulomb potential, the vector potential of the line current, in two dimensions, will take the logarithmic form:

$$(13.21)\quad A_z = \frac{\mu_o}{2\pi} J \ln\left(\frac{r_2}{r_1}\right)$$

and the corresponding complex potential is given by:

$$(13.22)\quad W = \phi + i\Psi = \frac{i\mu_o}{2\pi} J \ln\left(\frac{z_1}{z_2}\right)$$

The imaginary part of W, the stream function Ψ, is the vector potential. Therefore we can use all of the methods developed for finding the stream function in the solution of electrostatic problems in two dimensions for the solution of two-dimensional magnetic boundary value problems.

An important method of solution is the use of complex transformations, such

as the Schwarz transformation, to transform the solution of simple problems that involve currents in the neighborhood of permeable media having rectlinear boundaries into more complicated configuations. For example, theproblem of a line current located at a given distance from the surface of a semi-infinite permeable medium is soluble by the method of images* and therefore the solution of various problems concerning slots or gaps in permeable materials under the influence of magnetizing windings can be derived from this simple image solution by means of a suitable Schwarz transformation.

A three dimensional case capable of analytical treatment and also of practical importance is the case of cylindrical symmetry, i.e. current flow in coaxial circles only. In cylindrical coordinates the differential equation obeyed by A_ϕ, the only component of $\vec{A}$, is:

$$(13.23)\quad -\left[\vec{\nabla} \times (\vec{\nabla} \times \vec{A})\right]_\phi = \frac{\partial^2 A_\phi}{\partial \rho^2} + \frac{1}{\rho}\frac{\partial A_\phi}{\partial \rho} - \frac{A_\phi}{\rho^2} + \frac{\partial^2 A_\phi}{\partial z^2} = -\mu_o j_\phi$$

This separates into solutions of the form:

$$(13.24)\quad A_\phi = \cos(k z + \alpha_k)\left[A_k I_1(k\rho) + B_k K_1(k\rho)\right] \qquad \text{or:}$$

$$(13.25)\quad A_\phi = e^{\pm kz}\left[A_k^{\pm} J_1(k\rho) + B_k^{\pm} Y_1(k\rho)\right]$$

Solutions can then be obtained using the boundary conditions on an interface of constant ρ between regions (1) and (2):

$$(13.26)\quad A_\phi^{(1)} = A_\phi^{(2)} \qquad \text{and:}$$

$$(13.27)\quad \frac{1}{\mu^{(1)}}\frac{\partial}{\partial \rho}(\rho A_\phi^{(1)}) - \frac{1}{\mu^{(2)}}\frac{\partial}{\partial \rho}(\rho A_\phi^{(2)}) = j_{s\phi}$$

corresponding to $\vec{\nabla} \cdot \vec{B} = 0$ and $\vec{\nabla} \times \vec{H} = \vec{j}$, respectively. Since only one component of $\vec{A}$ is to be calculated, the solutions are similar to the electrostatic case.

---

*The image current here is of the same magnitude and sign as the current in the original conductor.

CHAPTER 14 MAXWELL'S EQUATIONS

Stationary Media

In electrostatics, the electric field was conservative, that is, we had:

$$(1.10)\quad \vec{\nabla} \times \vec{E} = 0$$

We have seen in Chapter 9 that there must be electric fields, such as $\vec{E}'$ in Equation (9.8) which violate this condition in order to produce stationary currents. It is found experimentally that a non-conservative electric field is actually observed in the presence of varying magnetic fields. The law which describes this situation is usually known as the Faraday Law of Induction and can be formulated as follows: Consider a circuit of resistance R carrying a current J and containing an electromotive force $\mathcal{E}$. If the magnetic flux Φ_m, which links this current, defined by:

$$(14.1)\quad \Phi_m = \iint \vec{B} \cdot \vec{dS}$$

changes at a given rate, it is found experimentally that:

$$(14.2)\quad JR - \mathcal{E} = - \frac{d\Phi_m}{dt}$$

This means that the current which flows in the circuit differs from the current predicted by Ohm's law (9.6) by an amount which is equivalent to an additional electromotive force equal to the negative time rate of change of flux through the circuit. Note that (14.2) is an independent experimental law and is in no way derivable from any of the relations that have been previously used. In particular, contrary to the statement that is sometimes made, Faraday's law of induction is not the consequence of the law of conservation of energy applied to the overall energy balance of currents in magnetic fields. Equation (14.2) is formulated in terms of the total flux passing through the given circuit. This flux can change for several reasons. It can change because of changes in the external field with time. It can change because of motion of the circuit itself or parts of the circuit. We shall consider (14.2) to be an experimental law which holds for all such cases. That is, we shall expect that (14.2) will also hold

for currents in moving media. It was recognized by Maxwell that the Faraday Law of Induction had a very much more general significance than the case actually described by Equation (14.2) would indicate. Equation (14.2) can be written in the equivalent form:

$$(14.3)\quad \oint \vec{E}' \cdot d\vec{\ell} = -\frac{d\Phi_m}{dt} = JR - \mathcal{E}$$

which indicates that in order for (14.2) to be valid, there must be an electric field along the wire which is non-electrostatic. However, from the boundary condition (3.19) that requires the tangential components of the electric field across the boundary of a wire to be continuous, we can conclude that (14.3) is also valid in the region that is immediately adjacent to the wire. Since the characteristics of the wire, namely its resistance and its electromotive forces, are not contained in the left hand side of (14.3) it appears to be likely that this relation is in fact independent of the presence of a current-carrying conductor and is a general physical law relating an electric field <u>in vacuo</u> to the rate of change of a magnetic field. If we make this induction, then (14.3) can be transformed into a differential form either in the case of free space or in the case of a stationary medium. In either case, after substituting (14.1) into (14.3) the total derivative of the flux integral can be written as an integral of the partial time derivative of the magnetic field, giving:

$$(14.4)\quad \oint \vec{E} \cdot d\vec{\ell} = -\frac{d}{dt} \iint \vec{B} \cdot d\vec{S} = -\iint \frac{\partial \vec{B}}{\partial t} \cdot d\vec{S}$$

and, using Stokes' theorem (1.14), we have:

$$(14.5)\quad \vec{\nabla} \times \vec{E} = -\frac{\partial \vec{B}}{\partial t}$$

Equation (14.5) expresses the modification which Faraday's Law of Induction introduces to the condition of irrotationality which was valid in electrostatic fields. Therefore we now have expressions for both the source densities (divergence) and the circulation densities (curl) of both of the basic field vectors $\vec{E}$ and $\vec{B}$. From (3.3), (3.4), (11.15), (12.3) and (14.5) we have:

$$(14.6)\quad (1)\ \vec{\nabla}\cdot\vec{E} = k_o^{-1}\rho_{total} = k_o^{-1}(\rho_{true} - \vec{\nabla}\cdot\vec{P})$$

$$(2)\ \vec{\nabla}\cdot\vec{B} = 0$$

$$(3)\ \vec{\nabla}\times\vec{E} = -\frac{\partial\vec{B}}{\partial t}$$

$$(4)\ \vec{\nabla}\times\vec{B} = \mu_o(\vec{j}_{true} + \frac{\partial\vec{P}}{\partial t} + \vec{\nabla}\times\vec{M} + k_o\frac{\partial\vec{E}}{\partial t})$$

Equations (14.6) are the formulation of Maxwell's Electro Dynamic Field Equations, valid for media at rest. The restriction to material media at rest arises from the omission of any convective current terms in Eq. (14.6) (4) and in the transformation from Equation (14.3) to (14.4) when flux changes due to motion of the medium were ignored. Equations (14.6) are written in terms of the equivalent vacuum charges or currents which give rise to the fields and contain the expressions for the equivalent current and charge densities explicitly. If the additional field vectors $\vec{D}$ and $\vec{H}$ are introduced by the defining equations:

$$(3.3)\ \vec{D} = k_o\vec{E} + \vec{P}$$

$$(12.5)\ \vec{H} = \frac{\vec{B}}{\mu_o} - \vec{M}$$

then Maxwell's field equations become:

$$(14.7)\quad (1)\ \vec{\nabla}\cdot\vec{D} = \rho_{true}$$

$$(2)\ \vec{\nabla}\cdot\vec{B} = 0$$

$$(3)\ \vec{\nabla}\times\vec{E} = -\frac{\partial\vec{B}}{\partial t}$$

$$(4)\ \vec{\nabla}\times\vec{H} = \vec{j}_{true} + \frac{\partial\vec{D}}{\partial t}$$

The form of the field equations in (14.7) appears to be simpler than (14.6), but it is actually more involved physically. The solution of these field equations is only possible if additional constitutive equations are available connecting $\vec{D}$ to $\vec{E}$, $\vec{j}$ to $\vec{E}$, and $\vec{H}$ to $\vec{B}$, such as:

$$(14.8)\ \vec{j}_{true} = \sigma\vec{E},\quad \vec{D} = kk_o\vec{E},\quad \vec{H} = \frac{\vec{B}}{\mu\mu_o}$$

for a linear medium, or whatever forms apply for a non-linear medium.

Moving Media

It is necessary to use considerable care in extending the field equations to the more general case when the media in which the fields are being measured are considered to be in motion. We must first derive the subsidiary theorem which expresses the total time rate of change of the flux across a given surface in terms of a surface integral of the vector function $\vec{B}$, even in case the surface itself across which the flux is being evaluated is also in motion. Let Φ_m be the flux of the vector field $\vec{B}$, across the surface S. We are looking for the function $\frac{D\vec{B}}{Dt}$ defined by:

$$(14.9)\quad \frac{d}{dt}\Phi_m = \frac{d}{dt}\iint \vec{B}\cdot d\vec{S} = \iint \left(\frac{D\vec{B}}{Dt}\right)\cdot d\vec{S}$$

In order to evaluate $\frac{D\vec{B}}{Dt}$ consider the surface in Figure (14.1) in a position 1, at a time t_1, and in a position 2 at a time t_2. By the rules for differentiation

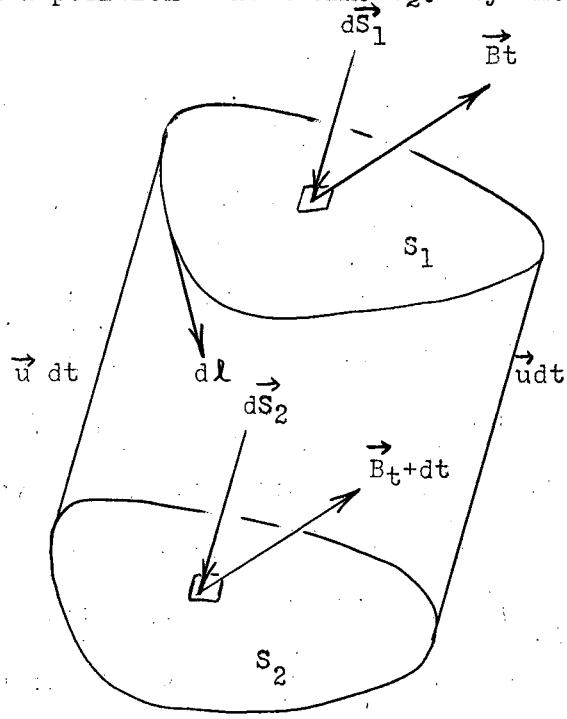


Figure (14.1)

we have:

$$(14.10)\quad \frac{\Delta}{\Delta t}\iint \vec{B}\cdot d\vec{S} = \frac{1}{\Delta t}\iint (\vec{B}_{t+dt}\cdot d\vec{S}_2 - \vec{B}_t\cdot d\vec{S}_1)$$

If we apply Gauss' theorem at the time t to the volume enclosed by S_1 and S_2 and the traces of the edges of S, we have:

$$(14.11)\quad \iiint \vec{\nabla} \cdot \vec{B}\, dv = \iint (\vec{B}_t \cdot d\vec{S}_2 - \vec{B}_t \cdot d\vec{S}_1) - \oint \vec{B}_t \cdot (\vec{u}\, dt \times d\vec{\ell})$$

Note that the flux change across the side surface generated by the motion of the boundary of S is included in the last term of (14.11). Also note that the flux across the surfaces S_1 and S_2 in Equation (14.11) is considered instantaneously at a given time t, since Gauss' theorem only applies to instantaneous values of the vector field $\vec{B}$. If we expand the value of $\vec{B}$ on S_2 by Taylor's theorem to get the value of $\vec{B}$ at the time t + dt in terms of its value at the time t, we get:

$$(14.12)\quad \vec{B}_{t+dt} = \vec{B}_t + \frac{\partial \vec{B}}{\partial t}\, dt + - - - -$$

Substituting (14.11) and (14.12) into (14.10) and passing to the limit, we obtain:

$$(14.13)\quad \frac{d}{dt}\iint \vec{B} \cdot d\vec{S} = \iint \frac{\partial \vec{B}}{\partial t} \cdot d\vec{S} + \oint (\vec{B} \times \vec{u}) \cdot d\vec{\ell} + \iiint \frac{\vec{\nabla} \cdot \vec{B}\, dv}{dt}$$

and using Stokes' theorem and:

$$(14.14)\quad dv = \vec{u} \cdot d\vec{S}\, dt$$

we obtain the desired relation:

$$(14.15)\quad \frac{D\vec{B}}{Dt} = \frac{\partial \vec{B}}{\partial t} + \vec{\nabla} \times (\vec{B} \times \vec{u}) + (\vec{\nabla} \cdot \vec{B})\, \vec{u}$$

The first term of this expression represents the change in the flux through S that is caused by the time variation in the vector field. The second term represents a change in the flux that is caused by flux loss across the boundary of the moving surface. The third term represents the change in the flux that is caused by the passage of the surface S through an inhomogeneous vector field.

Let us now use Equation (14.15) to express Faraday's law (14.2) in differential form in a moving medium. Since $\vec{B}$ is always solenoidal, we have from (14.3) and (14.15):

$$(14.16)\quad \oint \vec{E}' \cdot d\vec{\ell} = -\frac{d\Phi_m}{dt} = -\iint \left[\frac{\partial \vec{B}}{\partial t} + \vec{\nabla} \times (\vec{B} \times \vec{u})\right] \cdot d\vec{S}$$

We have designated the field by $\vec{E}'$ around the circuit since $\vec{E}'$ is to be measured in the moving frame of reference because Faraday's law applied specifically to the current measured in the wire through which the flux was changing, no matter what might be the cause of the flux change. By Stokes' theorem, (14.16), when the integrand is taken out of the integral, becomes:

$$(14.17) \qquad \vec{\nabla} \times \vec{E}' = - \frac{\partial \vec{B}}{\partial t} - \vec{\nabla} \times (\vec{B} \times \vec{u})$$

where $\vec{E}'$ still represents the field measured in the moving medium. Equation (14.17) can be written in the form:

$$(14.18) \qquad \vec{\nabla} \times (\vec{E}' - \vec{u} \times \vec{B}) = - \frac{\partial \vec{B}}{\partial t}$$

We shall now show that the argument of the curl in (14.18), $\vec{E}' - \vec{u} \times \vec{B}$, actually represents the field which is measured by a stationary observer. The reason for this is that an observer carrying a charge q through a magnetic field $\vec{B}$ with a velocity $\vec{u}$ will experience a force, $-q(\vec{u} \times \vec{B})$, in addition to the force of the electric field $\vec{E}$ which may also be present. Hence the electric field observed by a stationary observer is equal to the electric field $\vec{E}'$ observed by the moving observer, minus the effective field $\vec{u} \times \vec{B}$ and hence, in terms of the field $\vec{E}$ observed by a stationary observer, Equation (14.18) becomes:

$$(14.5) \qquad \vec{\nabla} \times \vec{E} = - \frac{\partial \vec{B}}{\partial t}$$

This means that the differential formulation of Faraday's Law of Induction is independent of the motion of the medium inside of the field. This is as it should be, since (14.5) is purely a field relation in terms of the equivalent vacuum fields $\vec{B}$ and $\vec{E}$, and should therefore be independent of the characteristics of the medium, including its motion.

However, the electric field observed by the moving observer does contain two terms, namely the "induced field" produced by the time rate of change of the external magnetic field, and the "motional field" $\vec{u} \times \vec{B}$, produced by the motion of the observer in the magnetic field. Note that in this discussion it has

been assumed that the electric field proper is not affected by the state of motion of the observer and that only an additional equivalent electric field is introduced as a result of the motion of the observer in the magnetic field. This assumption is actually justified only in cases when the motion of the observer is small compared to the velocity of light, as will be shown later by a relativistic analysis, and therefore all of our conclusions which we shall draw concerning Maxwell's equations in moving media will only apply when the velocities of such media are small compared to the velocity of light.

We have thus concluded that the third of Maxwell's equations in the form of (14.6) is not affected by the motion of the medium in which the fields are measured. The first and second equations are not affected by this motion either, since non-relativistically the charge density of the medium is not affected by the state of motion of the observer. The only modification which must be introduced is the addition of terms to the currents that appear in the fourth of Maxwell's equations in (14.6). Two additional terms are necessary, a convective term and a correction term to the polarization current. The convection current, due to the motion of the charge density and equivalent polarization charge is given by $\vec{u}\,(\rho_{true} - \vec{\nabla}\cdot\vec{P})$. The correction term which replaces the term $\frac{\partial \vec{P}}{\partial t}$ is $\frac{D\vec{P}}{Dt}$* which takes into account the charges lost due to the change of the polarization flux across the moving surface. To enumerate the various currents appearing in the total current which gives rise to magnetic fields:

(1) True currents $\vec{j}$.

(2) Convective currents of true and polarization charges given by $\vec{u}\,(\rho_{true} - \vec{\nabla}\cdot\vec{P})$.

(3) Currents caused by the rate of change of the polarization and the motion of polarized media, in analogy with (14.15), are given by:

$$\frac{D\vec{P}}{Dt} = \frac{\partial \vec{P}}{\partial t} + \vec{\nabla} \times (\vec{P} \times \vec{u}) + (\vec{\nabla}\cdot\vec{P})\,\vec{u}$$

---

* See Equation (14.15).

(4) Vacuum displacement current $k_o \frac{\partial \vec{E}}{\partial t}$. The other part of the displacement current is contained in (3).

Maxwell's equations in a non-magnetized medium moving with a velocity $\vec{u}$, slow compared to the velocity of light, are therefore given by:

(14.19) (1) $\vec{\nabla} \cdot \vec{D} = \rho_{true}$

(2) $\vec{\nabla} \cdot \vec{B} = 0$

(3) $\vec{\nabla} \times \vec{E} = -\frac{\partial \vec{B}}{\partial t}$

(4) $\vec{\nabla} \times \vec{B} = \mu_o \left[\vec{j}_{true} + \vec{u} \rho_{true} - \vec{\nabla} \cdot \vec{P}) + \frac{\partial \vec{P}}{\partial t} + \vec{\nabla} \times (\vec{P} \times \vec{u}) + (\vec{\nabla} \cdot \vec{P}) \vec{u} + k_o \frac{\partial \vec{E}}{\partial t} \right]$

$$\vec{\nabla} \times \vec{B} = \mu_o \left[\vec{j}_{true} + \rho_{true} \vec{u} + \frac{\partial \vec{D}}{\partial t} + \vec{\nabla} \times (\vec{P} \times \vec{u}) \right]$$

The constitutive equations which give the true currents in the moving medium and the polarization of the moving medium are derived from the fields measured in the moving medium, and are given by:

(14.20) $\vec{j} = \sigma (\vec{E} + \vec{u} \times \vec{B}) \pm \sigma \vec{E}'$

$\vec{P} = k_o (k-1)(\vec{E} + \vec{u} \times \vec{B})$

If we consider a non-charged dielectric, Maxwell's equations can be written in the form:

(14.21) (1) $\vec{\nabla} \cdot \vec{D} = 0$

(2) $\vec{\nabla} \cdot \vec{B} = 0$

(3) $\vec{\nabla} \times \vec{E} = -\frac{\partial \vec{B}}{\partial t}$

(4) $\vec{\nabla} \times \left[\vec{B} - \mu_o \vec{P} \times \vec{u} \right] = \mu_o \frac{\partial \vec{D}}{\partial t}$

This shows from the macroscopic point of view that a moving polarized dielectric is equivalent to a magnetized material of magnetic moment

(14.22) $\vec{M}_{eq} = \vec{P} \times \vec{u}$

This can be easily understood by considering a polarized slab of material moving at right angles to the direction of polarization. Under these conditions there is an equivalent positive current moving in one direction parallel to the direction

of motion, and another positive current displaced from the first current, due to the motion of the negative charges, moving in the opposite direction. These currents give rise to a net current loop and thus a magnetic moment. Hence the moving polarized dielectric will give rise to a magnetic field which is indistinguishable from that of a magnetized material. This has been demonstrated by the experiments of Roentgen and Eichenwald and others.

As an example of these considerations let us consider a conducting bar, as seen in Figure (14.2), which is infinitely long and has a rectangular cross section, and is moving with a velocity u relative to a constant magnetic field $\vec{B}$, which is directed at right angles to the direction of motion of the bar and which is constant in time. Two sliding contacts touch the conducting bar at

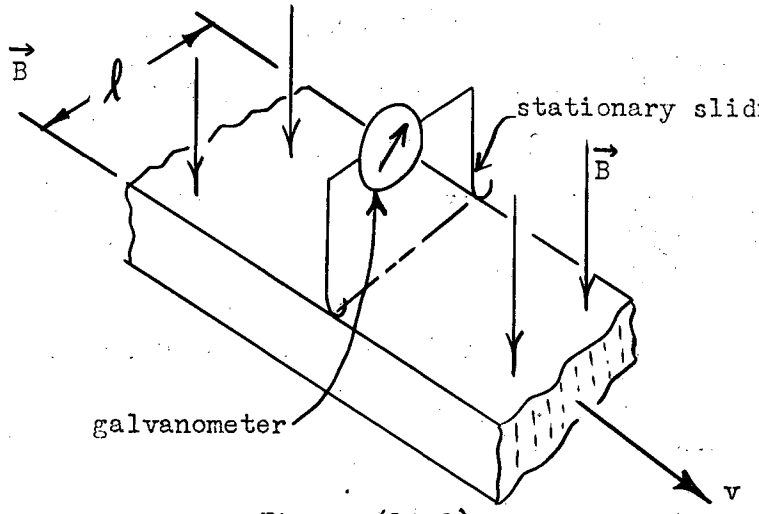


Figure (14.2)

points which are on opposite sides of the bar across a line perpendicular to the direction of the magnetic field and to the velocity as shown. Let us investigate how we can reconcile the phenomena observed with our previous considerations. Physically speaking, one would expect a current to flow in the external stationary loop. If an electron moving with the bar is considered, we find that an effective field given by $\vec{E'} = \vec{u} \times \vec{B}$ will act on the electron as a result of its motion in the magnetic field $\vec{B}$, and that a current through the contacts will be produced which will be measured in the external circuit by a stationary observer. $\vec{\nabla} \times \vec{E}$ must vanish, since $\vec{B}$ is not changing in time. Whether the source of $\vec{B}$ is stationary or, is in

motion is entirely irrelevant, since any observed phenomena which depend on a field description must be describable in terms of the behavior of the field quantities alone, independent of the nature of the mechanism which produces the field quantities. Hence a stationary observer must observe a field that has a vanishing curl; if, as appears logical by the above electron argument, there is an electric field, then such a field must be irrotational, i.e., electrostatic.

The effective electric fields within the moving bar will cause a current to flow within the bar causing charges on the bar to move to the faces and these charges will produce the observed external electrostatic field. On the other hand, the same charge displacement will exactly cancel the effective electric field $\vec{u} \times \vec{B}$ produced within the bar and therefore if we consider an integration path partially contained in the bar and partially outside of the bar, connected by the sliding contacts, then we shall have a circuit over which the line integral of the field will not vanish since we have a contribution to the line integral in the stationary part of the path and none in the bar. This result is therefore in agreement with the physically observed result that an electromotive force of magnitude $uB\ell$ is measured across the bar. Note, however, that if the electric field were measured entirely by a stationary circuit, then this field would actually be irrotational, that is, purely electrostatic.

If we imagine a small hole drilled through the bar transverse to $\vec{u}$ and $\vec{B}$ and consider a charge describing a loop threading this hole, no work will be done on the charge since the $\vec{u} \times \vec{B}$ term produced by the transverse force on the charge exerted by its neighbors is now missing.

If the galvanometer link is moved relative to the bar in the field $\vec{B}$ again the EMF $Bu\ell$ is observed, since the role of the link and the bar are simply interchanged in the above integration. The following table summarizes some of the cases of relative motion:

Motion of: / Case	Bar	Source of $\vec{B}$	Observer	Electromotive force measured by the observer
Case 1	u	0	0	$uB\ell$
Case 2	0	u	0	0
Case 3	0	u	u	$uB\ell$
Case 4	u	u	0	$uB\ell$
Case 5	0	0	u	$uB\ell$
Case 6	u	0	u	0

The results in the above table are characterized by two salient facts:

(1) The state of motion of the source $\vec{B}$ is irrelevant as long as $\vec{B}$ is uniform.

(2) Absolute motion cannot be detected in this arrangement.

The latter fact is an indication that Maxwell's equations, if carefully interpreted, are in agreement with relativistic principles. This will be shown later in greater detail and generality.

The situation is more complicated in case there is in addition to an external magnetic field $\vec{B}$, a field caused by the magnetic moment $\vec{M}$ of the slab, either induced or permanent. Our conclusion that the electric field observed in a stationary loop will be a purely electrostatic one still remains valid. However, the source of the electrostatic field will not become fully clear until permeable media have been introduced into the equations for moving media. This, unfortunately, cannot be done in a reasonable way without introducing relativistic considerations. However, the result is physically clear, since the source of the magnetic field, provided it is constant in time, does not effect the considerations. The force which acts on a moving electron within a moving bar will be independent of whether the magnetic field that produces that force is produced by an external magnetic field or by the magnetic moment of the bar itself. Therefore we should expect to obtain an electromotive force, given by $uB\ell$ as before, where $\vec{B}$ is the

magnetic field in the moving magnetized bar. We would also again expect that the total electric field measured by a stationary observer is electrostatic. The only part which appears still to be paradoxical is that the effect of a moving material should be describable in terms of the sum of the effects of its atomic components, and since a moving magnet is essentially an assembly of current loops, we would therefore conclude that the motion of the loops carrying a steady current gives rise to an electrostatic field. This, as will be shown later, by relativistic considerations, is in fact true. We shall show that in general if the medium of magnetization $\vec{M}$ is observed by an observer who is moving with a velocity $\vec{u}$ relative to the medium, then the observer will observe an equivalent electric moment, given by:

$$(14.23)\quad \vec{P}_{eq} = \frac{1}{c^2}\,\vec{u} \times \vec{M}$$

and therefore:

$$(14.24)\quad \vec{\nabla} \cdot \vec{E} = -\frac{\vec{\nabla} \cdot \vec{P}_{eq}}{k_o}$$

will define the sources of the field. Note that this effect, although it appears deceptively similar to the classical effect:

$$(14.22)\quad \vec{M}_{eq} = \vec{P} \times \vec{u}$$

is actually only explainable in terms of the special theory of relativity. It is caused by the fact that a charge moving in a circuit appears to spend a different amount of time, traveling in the direction parallel to the relative motion between the circuit and the observer, from the time that it spends moving in the anti-parallel direction. This gives rise to an effective polarization that is perpendicular to the direction of motion and lies in the plane of the current loop. We shall discuss this effect in detail later.

If the length of the magnetized slab is finite, then the field $\vec{E}$ is no longer irrotational since $\frac{\partial \vec{B}}{\partial t} \neq 0$ in the rest frame.

In fact:

$$\vec{\nabla} \times \vec{E} = -\frac{\partial \vec{B}}{\partial t} = +(\vec{u} \cdot \vec{\nabla})\vec{B} = \vec{u}(\vec{\nabla} \cdot \vec{B}) - \vec{\nabla} \times (\vec{u} \times \vec{B})$$

$$= -\vec{\nabla} \times (\vec{u} \times \vec{B})$$

(since $\vec{\nabla} \cdot \vec{B} = 0$ and since temporal and spatial variations are connected by $\frac{\partial}{\partial t} = -\vec{u} \cdot \vec{\nabla}$ for uniform motion $\vec{u}$).

Hence if $\vec{B}$ is no longer uniform, then $\vec{E}$ is no longer irrotational; its curl is however identical with the curl of $-(\vec{u} \times \vec{B})$ which is also the effective electric field acting in the moving medium.

CHAPTER 15 ENERGY RELATIONS IN THE ELECTROMAGNETIC FIELD AND FORCES ON CURRENT SYSTEMS

We shall now consider some energy relations that are associated with electromagnetic fields. In the discussion of the energy relations in electrostatic fields in Chapter 7, we succeeded in associating an energy density with the electric field by considering a specific process; namely, the assembly of charges, in which the work that was done and the changes in the fields could be calculated. It was possible to obtain a free energy density of the electric field, in the thermodynamic sense, by balancing the work and the energy terms. In order to treat the energy of a magnetic field let us start by considering a process in which a battery which produces a non-electrostatic field $\vec{E}'$, is feeding energy both into heat losses and into a magnetic field. If we take the scalar product of $\vec{j}$ and the equation:

$$\vec{j} = \sigma(\vec{E} + \vec{E}') \tag{9.4}$$

we obtain:

$$\vec{E}' \cdot \vec{j} = \frac{j^2}{\sigma} - \vec{E} \cdot \vec{j} \tag{15.1}$$

The lefthand side of (15.1) represents the time rate at which the battery does work, the first term on the righthand side represents the Joule heat loss in the current-carrying medium, and the last term on the righthand side we tentatively identify as the rate at which energy is fed into the magnetic field produced by the currents. If we consider the fields to be slowly varying, that is, quasi-stationary so that the displacement current terms need not be taken into account, we obtain from Maxwell's Equation (14.7)(4):

$$\vec{\nabla} \times \vec{H} = \vec{j} \tag{15.2}$$

If we substitute (15.2) into (15.1) and integrate over all space, we have:

$$\iiint \vec{E}' \cdot (\vec{\nabla} \times \vec{H}) dv = \iiint \frac{(\vec{\nabla} \times \vec{H})^2}{\sigma} dv - \iiint \vec{E} \cdot (\vec{\nabla} \times \vec{H}) dv \tag{15.3}$$

Using the relation:

$$(15.4)\quad \vec{\nabla}\cdot(\vec{E}\times\vec{H}) = \vec{H}\cdot\vec{\nabla}\times\vec{E} - \vec{E}\cdot\vec{\nabla}\times\vec{H}$$

to integrate the last term by parts we obtain:

$$(15.5)\quad \iiint \vec{E}\cdot\vec{\nabla}\times\vec{H}\,dv = \iiint \vec{H}\cdot\vec{\nabla}\times\vec{E}\,dv - \iint \vec{\nabla}\cdot(\vec{E}\times\vec{H})dv$$

$$\iiint \vec{E}\cdot\vec{\nabla}\times\vec{H}\,dv = -\iiint \vec{H}\cdot\frac{\partial\vec{B}}{\partial t}\,dv - \int \vec{E}\times\vec{H}\cdot d\vec{S}$$

and dropping the surface term and substituting back into (15.3), we get:

$$(15.6)^{*}\quad \iiint \vec{E}'\cdot\vec{\nabla}\times\vec{H}\,dv = \iiint \frac{(\vec{\nabla}\times\vec{H})^2}{\sigma}\,dv + \iiint \vec{H}\cdot\frac{\partial\vec{B}}{\partial t}\,dv$$

(Power expanded by the battery) — (Rate of Joule heat loss) — (Rate at which energy is fed into the field)

We can neglect the surface term generated in the integration by parts, since $\vec{E}\times\vec{H}$ varies at least as $1/r^5$ in electrostatic and quasi-stationary magnetic fields, and therefore the resultant integral vanishes at least to the order $1/r^3$. Note however, that this will not be true in case $\vec{E}$ and $\vec{H}$ represent radiation fields which fall off as $1/r$. The surface terms will then represent a constant radiation energy loss. Since we neglected the displacement current term in Equation (15.2), we are justified in neglecting any radiation contributions at this point subject to further study of the energy balance when radiation terms will be taken into account. The same thing applies to our electrostatic field energy studies also. In this analysis we are using, separately, energy relations in electrostatic fields on the one hand, and quasi stationary current magnetic fields on the other hand. We shall see later how these concepts can be modified in a consistent way to obtain the general energy expressions. From Equation (15.6) the variation in the magnetic field energy δU_m can therefore be given by:

$$(15.7)\quad \delta U_m = \iiint \vec{H}\cdot\delta\vec{B}\,dv$$

This is analogous to the electrostatic expression:

$$(7.20)\quad \delta U = \iiint \vec{E}\cdot\delta\vec{D}\,dv$$

---

* If the term $\frac{\partial\vec{D}}{\partial t}$ had been retained in Eq. (15.2), a term $\iiint \vec{E}\cdot\frac{\partial\vec{D}}{\partial t}\,dv$ would have appeared in Eq. (15.6), corresponding to the rate of increase of energy of the associated electric field.

U_m represents only a free energy term as was the case with U. In order to put (15.7) into the form of an integral over the magnetic field energy density, the expression must be made integrable. That is we must assume a functional relationship between $\vec{H}$ and $\vec{B}$. In case the medium magnetizes linearly, (15.7) can be integrated in the same manner as in the case of (7.14), giving:

$$(15.8) \qquad U_m = \frac{1}{2}\iiint \vec{H} \cdot \vec{B}\, dv$$

In non-linear materials such as ferro-magnets, (15.7) can only be integrated between definite states and the answer will in general depend on the past history of the sample of iron that is being considered. In the case of ferro-magnets the cyclic integral of (15.7) is in general not zero, but has a finite value when $\vec{B}$ is evaluated around a complete cycle, as in a field produced by an alternating current. The cyclic energy loss is given by:

$$(15.9) \qquad U_m = \iiint \oint \vec{H} \cdot d\vec{B}\, dv$$

Equation (15.9) says that the energy expended per unit volume when a magnetic material is carried through a magnetization cycle is equal to the area of its hysteresis loop, when plotted in the H - B plane.

Equation (15.8) gives directly the energy density in terms of a volume integral over the fields. If instead we wish an equation for the energy expressed as a volume integral over the current sources of the field, we need only to express $\vec{B}$ in terms of the vector potential $\vec{A}$, and $\vec{H}$ in terms of the stationary current field equation (15.2). This gives:

$$(15.10) \qquad U_m = \frac{1}{2}\iiint \vec{H} \cdot (\vec{\nabla} \times \vec{A})\, dv$$

and integrating by parts and dropping a surface term as was done in (15.5) we obtain:

$$(15.11) \qquad U_m = \frac{1}{2}\iiint \vec{j} \cdot \vec{A}\, dv$$

This expression is analogous to the expression for the electrostatic energy in terms of volume charge density and the scalar potential.

This expression and Equations (15.7) and (15.8), although derived by a particular "virtual process", can be taken to represent the general expressions for the energy of the magnetic field expressed as a field integral or a matter integral respectively, since the expressions depend only on the final fields, and not on the nature of the process.

The factor 1/2 in (15.11) is similar to the factor 1/2 in Eq. (7.1) and is due to the fact that the vector potential $\vec{A}$ includes the fields of the currents $\vec{j}$ themselves. The interaction energy of a system of currents and charges in an external field of potentials ϕ and $\vec{A}$ respectively is given by:

$$(15.12) \qquad U_{interaction} = \iiint \left\{ \vec{j} \cdot \vec{A}_{external} + \rho \phi_{external} \right\} dv$$

We will now use the energy expressions for two purposes; first, to derive expressions for the forces between currents in terms of the currents themselves and suitable geometrical parameters which depend on the location of the currents; and second, to express the variation of the magnetic field energy in terms of variation of the currents that produce the magnetic fields and the variations in the geometrical coordinates.

Let us solve this problem by analyzing a system of n geometrically linear circuits carrying currents J_k. For these line circuits, the energy expression (15.11) reduces to:

$$(15.13) \qquad U_m = \frac{1}{2} \sum_{k=1}^{n} J_k \oint \vec{A} \cdot d\vec{\ell}_k = \frac{1}{2} \sum_{k=1}^{n} J_k \iint (\vec{\nabla} \times \vec{A}) \cdot d\vec{S}_k$$

$$U_m = \frac{1}{2} \sum_{k=1}^{n} J_k \iint \vec{B} \cdot d\vec{S}_k = \frac{1}{2} \sum_{k=1}^{n} J_k \Phi_k$$

where the transformation from the vector potential line integral to the fluxes Φ_k linking the k^{th} circuit has been made by the use of Stokes' theorem, (11.16) and (14.1). To derive the forces, let us consider that the i^{th} circuit is subjected to a virtual, infinitesimally slow velocity $\vec{v}_i$. Then the rate at which an external force $\vec{F}_i$ acting on the i^{th} circuit is doing work is equal

to $\vec{F}_i \cdot \vec{u_i}$. The rate of energy change must be balanced between the following four quantities:

(1) Rate at which mechanical work is being done by the external forces.

(2) Rate of change of magnetic field energy.

(3) Rate of Joule heat losses.

(4) Rate at which work is being done by the electromotive forces within the circuits.

Equating the overall rate of energy change to zero, we obtain:

$$(15.14)\quad \vec{F}_i \cdot \vec{u_i} + \frac{dU_m}{dt} + \sum_{k=1}^{n} J_k^2 R_k - \sum_{k=1}^{n} J_k \mathcal{E}_k' = 0$$

(Rate of mechanical work) (Rate of change of magnetic field energy) (Rate of Joule heat loss) (Rate at which battery does work)

We are assuming that the magnetic field energy U_m is explicitly expressed as a function of the coordinates x_k of the k^{th} current loop and of the current J_k flowing in the k^{th} loop as independent variables. Note that because of the different terms in Eq. (15.14), it is not justified to simply equate the force on the i^{th} circuit $\vec{F}_i$ to the negative gradient of the field energy U_m at constant current, a conclusion which would be justified only if no other energy terms than U_m were present.

Let us now consider a special type of constant current process, namely let the external electromotive forces be adjusted as a function of the virtual velocity $\vec{u_i}$ corresponding to the rate of a single parameter $\vec{x_i}$, so that the currents within the system remain constant. In this case, if we substitute (15.13) into (15.14) and use Faraday's law (14.2), we obtain:

$$\vec{F}_i \cdot \vec{u_i} + \frac{1}{2} \sum_{k=1}^{n} J_k \frac{d\Phi_k}{dt} - \sum_{k=1}^{n} J_k \frac{d\Phi_k}{dt} = 0$$

or:

$$(15.15)\qquad \vec{F}_i \cdot \vec{u}_i = \frac{\partial U_m}{\partial t} = \frac{\partial U_m}{\partial x_i} u_i \Bigg|_{\substack{J \text{ constant} \\ x_k \text{ constant; } k \neq i}}$$

Hence:

$$(15.16)^{\star}\qquad F_i = \frac{\partial U_m}{\partial x_i}\Bigg|_{J \text{ constant}}$$

is the force exerted by the field on the i^{th} conductor. Note that the opposite sign has been obtained from the sign which would be expected from elementary consideration if the other energy terms were neglected. This means that in order to maintain a constant current in the circuits, as the geometry changes, the external batteries must do exactly twice the amount of work that is done by the external forces, in addition to supplying the Joule heat losses. Equation (15.16) is very useful when it is desired to calculate the forces acting on current carrying circuits if the magnetic field energy is expressible in terms of the current producing the field.

To express the magnetic field energy

$$(15.13)\qquad U_m = \frac{1}{2}\sum_{k=1}^{n} J_k \Phi_k$$

as a function of current and geometry, it is useful to introduce the concept of inductance. The flux through the k^{th} circuit is given by

$$(15.17)\qquad \Phi_k = \iint \vec{B} \cdot d\vec{S}_k = \iint (\vec{\nabla} \times \vec{A}) \cdot d\vec{S}_k = \oint_k \vec{A} \cdot d\vec{\ell}_k$$

$\vec{A}$ in turn can be evaluated by writing the integral of the vector form of Poisson's equation (13.15) in the form, similar to (13.16) but expressed as a line integral

$$(15.18)\qquad \vec{A}(x_k) = \frac{\mu_o}{4\pi}\sum_i \oint \frac{J_i \, d\vec{\ell}_i}{r_{ik}}$$

Substituting into Eqs. (15.13) and (15.17), we obtain

$$(15.19)\qquad \Phi_k = \sum_i L_{ik} J_i \,, \quad \text{and,}$$

$$(15.20)\qquad U_m = \frac{1}{2}\sum_i \sum_k L_{ik} J_i J_k$$

$\star$ This equation is written in terms of force components; x_i here signifies anyone of the possible geometrical parameters of the i^{th} circuit.

where

$$(15.21) \qquad L_{ik} = \frac{\mu_o}{4\pi} \oint_i \oint_k \frac{d\vec{\ell}_i \cdot d\vec{\ell}_k}{r_{ik}} = L_{ki}$$

is a purely geometrical quantity, called the mutual inductance between the i^{th} and k^{th} circuit. The force acting on the i^{th} circuit is thus, from Eq. (15.16):

$$(15.22) \qquad F_i = + \frac{\partial}{\partial x_i} U_m \Big|_{J=\text{const.}} = \frac{1}{2} \sum_j \sum_k J_j J_k \frac{\partial L_{jk}}{\partial x_i}$$

Note that in the sums of Eqs. (15.20) and (15.22) each term for which the two indices are different occur twice, while for equal indices, the term occurs only once. The mutual energy of two circuits is thus:

$$(15.23) \qquad U_m = J_1 J_2 L_{12}$$

while the self energy of the circuits is:

$$(15.24) \qquad U_m = \frac{1}{2} \left\{ J_1^2 L_{11} + J_2^2 L_{22} \right\}$$

where the L_{ii} are called the self inductances.

The force expression (15.22) is in agreement with the original magnetic interaction expression Eq. (11.1). If we substitute "Neumann's formula" (15.21) into (15.22) (applying to two circuits) we obtain:

$$(15.25) \qquad F_1 = \frac{\mu_o}{4\pi} J_1 J_2 \oint_1 \oint_2 (d\vec{\ell}_1 \cdot d\vec{\ell}_2) \vec{\nabla} \left(\frac{1}{r_{12}} \right)$$

$$(15.26) \qquad = - \frac{\mu_o}{4\pi} J_1 J_2 \oint_1 \oint_2 \vec{r}_{12} \left(\frac{d\vec{\ell}_1 \cdot d\vec{\ell}_2}{r_{12}^3} \right)$$

which is identical with Eq. (11.4).

The force equation (15.16) can also be written in the simple form:

$$(15.27) \qquad F_i = \frac{\partial U_m}{\partial x_i} \Big|_{J \text{ const.}} = 1/2 \sum_{jk} J_j J_k \frac{\partial L_{jk}}{\partial x_i} = J_i \sum_j J_j \frac{\partial L_{ij}}{\partial x_i}$$

the last step follows since only the term for which either $j = i$ or $k = i$ depends on x_i and gives non-zero derivatives, and since $L_{ij} = L_{ji}$, (15.27) can be written as:

$$(15.28) \qquad F_i = J_i \left.\frac{\partial \Phi_i}{\partial x_i}\right|_{J \text{ const.}}$$

where Φ_i is the flux linking the i^{th} circuit. This expression is in evident agreement with the elementary force relation:

$$(15.29) \qquad \vec{dF} = J\, \vec{d\ell} \times \vec{B}$$

These considerations enable us to express the general variation field energy as a function, independently, of the geometrical parameters x_i and of the currents. Since:

$$(15.30) \qquad U_m = 1/2 \sum_i \sum_j L_{ij} J_i J_j$$

we have (note each term occurring twice):

$$U_m = \sum_i \left\{ \left(\sum_j L_{ij} J_j \right) \delta J_i + \sum_j \frac{\partial L_{ij}}{\partial x_i} \delta x_i J_i J_j \right\}$$

Hence, from (15.19) and (15.22):

$$(15.31) \qquad U_m = \sum_i \left\{ \Phi_i \delta J_i + F_i \delta x_i \right\}$$

where Φ_i is the total flux linking the i^{th} circuit. Note that positions and currents play the roles of extensive variables in the termodynamic sense, while the forces and the fluxes play the roles of intensive variables.

Note also, direct from (15.14) that the "Back E.M.F." terms can be ignored in force calculations if the flux linkages are held constant; it follows then directly that:

$$(15.32) \qquad F_i = - \left.\frac{\partial U_m}{\partial x_i}\right|_{\Phi \text{ constant}} \qquad \text{in contrast to Eq. (15.16)}$$

The self inductances and mutual inductances can be calculated by several means other than Neumann's formula. One method is to use the defining equation (15.19). The flux linking the j^{th} circuit due to the current in the i^{th} circuit, can be evaluated directly from the known field or vector potential of the i^{th} circuit. A second method which is particularly useful in case continuous current distributions and therefore partial flux linkages are involved is a computation

which uses the magnetic field energy directly, namely by use of Equations (15.8) and (15.20):

$$(15.33) \quad U_m = 1/2 \sum_{i=j}^{n} \sum_{j=1}^{n} L_{ij} J_i J_j = 1/2 \iiint \vec{H} \cdot \vec{B} \, dv$$

$\vec{H}$ and $\vec{B}$ may be computed by methods that have already been discussed. In this case, the calculation of inductances is then carried out by the evaluation of certain integrals of the solutions of boundary value problems.

In the calculations of the inductances of current-carrying conductors, it is usually advantageous to separate the problem into two parts; first, the calculation of the external inductance, that is the inductance associated with the field outside of the wire; and second, the calculation of the contribution to the inductance by the field energy of the field inside the wire. It is necessary to make this separation since the inductance due to the external field cannot be computed by using the assumption that the conductor has a zero radius, since this will generally lead to a logarithmic divergence of the integral involved. A lower limit must be put on the coordinate in the integral representing the flux or the energy in the external field and then the contribution to the inductance by the field within the wire must be computed separately. This latter calculation generally gives a term which will depend on the permeability of the wire. At high frequencies this term becomes negligible, since the currents do not penetrate into the wire. At lower frequencies, and particularly in case the current-carrying conductors are ferro-magnetic, this internal term may give an appreciable contribution.

CHAPTER 16 FIELD FORMULATION OF THE ENERGY AND MOMENTA IN THE GENERAL ELECTROMAGNETIC FIELD

In Chapter 15 we have calculated the forces between current systems in terms of the currents within the current systems, and the necessary geometrical quantities. These forces are, of course, as would be expected, re-expressions of the original Ampere interaction law given in Equation (11.1). We can, in analogy to the electrostatic case, derive an expression for the magnetic body force per unit volume in terms of the field, the permeability and the current at a given point. In the electrostatic case such a body force was defined by:

$$(8.2)\quad \frac{dU}{dt} = -\iiint \vec{F}_v \cdot \vec{u}\, dv$$

The expression for $\vec{F}_v$ the body force was:

$$(8.19)\quad \vec{F}_v = \rho \vec{E} - \frac{k_o}{2} E^2 \vec{\nabla} k + \frac{k_o}{2} \vec{\nabla}\left(E^2 \frac{dk}{dg} g\right)$$

Under the following restrictions:

(1) The medium is linear, that is, its permeability μ is not a function of the field.

(2) There is no permanent magnetic moment $\vec{M}$ present.

(3) There is no magnetostriction, that is $\frac{d\mu}{dg} = 0$.

We can derive for the magnetic body force the expression:

$$(16.1)\quad \vec{F}_v = \vec{j} \times \vec{B} - \frac{1}{2\mu_o} B^2 \vec{\nabla} \mu$$

It is again possible, as should be the case in a satisfactory field theory, to derive the total force on a volume element, bounded by a given surface, in terms of the value of the field on the boundary of this volume. That is, it is also possible in the magnetic case to define a stress tensor from which the volume force is derivable by the tensor divergence relation (8.21). The form of the Maxwell tensor in the magnetic case, in the absence of a magnetostriction term, can be shown to be:

$$(16.2)\quad T_{ij} = H_i B_j - \frac{\delta_{ij}}{2} H_k B_k$$

The Einstein summation convention mentioned in Chapter 8, has been assumed here. The geometrical interpretation of this tensor leads to the same conclusions about magnetic forces as were reached in Chapter 8 about electrical forces. These are that the direction of the magnetic field bisects the angle between the normal to a surface and the direction of the resultant magnetic stress that acts on this surface. The magnitude of the magnetic stress normal to the magnetic field or parallel to the field is given by HB/2.

Thus far we have considered the electrostatic field and the magnetostatic field, or quasi-stationary current fields separately. We considered energy and force problems separately for the two cases, deriving expressions for the electric and magnetic energy densities separately. If we now consider the general case in which no restrictions as to the time rate of variation of the field quantities are imposed, then we must inquire as to which of the energy, force or momentum expressions will need modification, and which ones can be taken over in the more general theory without modification. We shall restrict our considerations to vacuum fields or non-permeable conductors, since no additional information of particular interest will result if this restriction were not imposed.

Maxwell's equations in vacuo were shown to be:

$$(14.7) \qquad (1)\ \vec{\nabla} \cdot \vec{D} = \rho_{true}$$

$$(2)\ \vec{\nabla} \cdot \vec{B} = 0$$

$$(3)\ \vec{\nabla} \times \vec{E} = -\frac{\partial \vec{B}}{\partial t}$$

$$(4)\ \vec{\nabla} \times \vec{H} = \vec{j}_{true} + \frac{\partial \vec{D}}{\partial t}$$

These equations completely represent the behavior of electromagnetic field when they are considered in combination with suitable constitutive equations and boundary conditions. This is true even for rapidly varying fields, or at least no internal contradiction is present if an arbitrary rate of change is assumed. Care must be taken in case constitutive equations are to be used, since generally

the material constants are dependent on the frequency of the fields. We will frequently restrict ouselves to vacuum conditions in order to avoid unnecessary complications due to such special properties of the constitutive equations.

An energy integral of Maxwell's equations can be obtained by taking the dot product of the third and fourth equations of (14.7) with $\vec{H}$ and $\vec{E}$ respectively. Subtracting the two equations that result, and using the well known vector identity:

$$(15.4) \qquad \vec{\nabla} \cdot (\vec{E} \times \vec{H}) = \vec{H} \cdot (\vec{\nabla} \times \vec{E}) - \vec{E} \cdot (\vec{\nabla} \times \vec{H})$$

we obtain:

$$(16.3) \qquad \vec{\nabla} \cdot (\vec{E} \times \vec{H}) = - \frac{\partial \vec{B}}{\partial t} \cdot \vec{H} - \vec{E} \cdot \vec{j} - \vec{E} \cdot \frac{\partial \vec{D}}{\partial t}$$

Taking the volume integral of (16.3) and using the linear relations $\vec{B} = \mu\mu_0 \vec{H}$, $\vec{D} = kk_0 \vec{E}$ and the divergence theorem, we have:

$$(16.4) \qquad - \frac{\partial}{\partial t} \iiint \frac{1}{2} (\vec{H} \cdot \vec{B} + \vec{E} \cdot \vec{D}) dv = \iiint \vec{E} \cdot \vec{j} \, dv + \iint (\vec{E} \times \vec{H}) \cdot d\vec{S}$$

The left-hand side of (16.4) represents the rate of decrease of the sum of the electric and magnetic field energies (7.19) and (17.3) that were derived in the static cases. Using Equation (15.1), the first term on the right-hand side of (16.4) can be written as:

$$(16.5) \qquad \iiint \vec{E} \cdot \vec{j} \, dv = \iiint \left(\frac{j^2}{\sigma} - \vec{E}' \cdot \vec{j} \right) dv$$

Equation (16.5) therefore represents the sum of the Joule heat loss and the negative rate at which the electromotive forces are doing work. ($\vec{E}'$ is the electromotive force field.) Thus (16.4) becomes:

$$(16.6) \qquad - \frac{\partial}{\partial t} \iiint 1/2(\vec{H} \cdot \vec{B} + \vec{E} \cdot \vec{D}) dv = \iiint \frac{j^2}{\sigma} \, dv - \iiint \vec{E}' \cdot \vec{j} \, dv + \iint (\vec{E} \times \vec{H}) \cdot d\vec{S}$$

(Rate of change of electric and magnetic field energy) (Rate of Joule heat loss) (Rate of work by sources of electromotive force) (Surface rate of energy transfer)

The third term on the right-hand side of Equation (16.6) is a radiation energy term which has previously been neglected, since for static and quasi-static fields it can be made to vanish if an arbitrarily large enclosing surface is used for the integration. As we shall see later, the electric and magnetic radiation

fields of charge motions and currents fall off in general only as $1/r$ at large distances and therefore the integral $\iint (\vec{E} \times \vec{H}) \cdot d\vec{S}$ will approach a constant value when evaluated for an arbitrarily large surface and thus may contribute to the energy balance.

The vector

$$\vec{N} = \vec{E} \times \vec{H} \tag{16.7}$$

is known as the Poynting vector and in terms of Equation (16.6) it can be considered to represent the electromagnetic field energy flow per unit area per unit time across a given surface. It must be noted, however, that only the entire surface integral of $\vec{N}$ contributes to the energy balance and since questions of localization of energy cannot be decided uniquely by experiment, paradoxical results will often be obtained if one tries to identify the Poynting vector with the energy flow per unit area at any particular point. Among other things, since only the surface integral of the Poynting vector contributes to the overall energy balance, the net energy flow in the electromagnetic field will always vanish if the divergence of the Poynting vector is zero. If, for example, we have static superposed electric and magnetic fields, we may have non-zero values of the Poynting vector at various points in space, but the divergence of the Poynting vector will vanish everywhere, implying that radiation does not contribute to the energy balance in this special case.

Equation (16.6) can therefore be considered to represent the overall energy balance between the electric and the magnetic energy of the field, the loss due to resistive heating, the work done by sources of electromotive force, and the radiation loss. It appears therefore even in the case of time-varying fields, that we can retain the expressions derived for the energy densities of the electrostatic and the magnetostatic fields. The only additional consideration which has been introduced in order to conserve energy is to assume that radiation fields may carry energy in or out of the volume of integration at a rate that is given by the surface integral of the Poynting vector.

Let us now apply a similar consideration to the momentum balance in the electromagnetic field.* Let us inquire as to whether the tensor divergence of the tensor formed from the sum of (8.32) and (16.2), the complete Maxwell stress tensor:

$$(16.8)\quad T_{ij} = E_i D_j - 1/2\ \delta_{ij} E_k D_k + H_i B_j - 1/2\ \delta_{ij} H_k B_k$$

will still give a volume force which is in accordance with experience. If the tensor divergence of (16.8) is taken, we obtain:

$$(16.9)\quad F_{vi} = \frac{\partial T_{ij}}{\partial X_j} = E_i \frac{\partial D_j}{\partial X_j} + D_j \frac{\partial E_i}{\partial X_j} - 1/2\ E^2 k_o \frac{\partial k}{\partial X_i} - D_j \frac{\partial E_j}{\partial X_i}$$

$$+ H_i \frac{\partial B_j}{\partial X_j} + B_j \frac{\partial H_i}{\partial X_j} - 1/2\ H^2 \mu_o \frac{\partial \mu}{\partial X_i} - B_j \frac{\partial H_j}{\partial X_i}$$

Equation (16.9) can be expressed in vector form by using Maxwell's first equation (14.7) on the first term, noting that the third and seventh terms involve gradients of k and μ respectively, noting that the second and fourth terms become, $-\vec{D} \times (\vec{\nabla} \times \vec{E})$ which by Maxwell's third equation (14.7) is equal to $\vec{D} \times \frac{\partial \vec{B}}{\partial t}$, noting that the sixth and eighth terms become, $-\vec{B} \times (\vec{\nabla} \times \vec{H})$ which by Maxwell's fourth equation (14.7) is equal to $-\vec{B} \times (\vec{j}_{true} + \frac{\partial \vec{D}}{\partial t})$, and noting that the 5th term vanishes since $\vec{\nabla} \cdot \vec{B} = 0$.

$$(16.10)\quad \vec{F}_v = \vec{E} \rho_{true} - \frac{k_o}{2} E^2 \vec{\nabla} k - \frac{\mu_o}{2} H^2 \vec{\nabla} \mu - \vec{B} \times \vec{j}_{true} + \frac{\partial}{\partial t} (\vec{D} \times \vec{B})$$

The entire volume force of Equation (16.10) may be expressed as the sum of two terms:

$$(16.11)\quad \vec{F}_v = \vec{F}_{ev} + \frac{\partial}{\partial t} (\vec{D} \times \vec{B})$$

where the first term:

$$(16.12)\quad \vec{F}_{ev} = \vec{E} \rho_{true} - \frac{k_o}{2} E^2 \vec{\nabla} k - \frac{\mu_o}{2} H^2 \vec{\nabla} \mu - \vec{B} \times \vec{j}_{true}$$

is the ordinary volume force acting on material bodies in a quasi-stationary electromagnetic field. The volume force resulting from the presence of true charges

---

* We are neglecting the electro- and magnetostriction terms.

or inhomogeneous dielectrics in an electric field or of true currents or inhomogeneous permeable material in a magnetic field is fully accounted for by (16.12). The second term in (16.11) is new and is proportional to the time rate of change of the Poynting vector. The volume force (16.11) may be expressed as:

$$(16.13)\quad \vec{F}_v = \vec{F}_{ev} + kk_o\mu\mu_o \frac{\partial(\vec{E} \times \vec{H})}{\partial t} = \vec{F}_{ev} + kk_o\mu\mu_o \frac{\partial \vec{N}}{\partial t} = \vec{F}_{ev} + \frac{k\mu}{c^2}\frac{\partial \vec{N}}{\partial t}$$

where:

$$(16.14)\quad \mu_o k_o = \frac{1}{c^2}$$

$\vec{F}_{ev}$ is the volume force whose existence is dependent upon the presence of material bodies carrying charges or endowed with dielectric or permeable properties. On the other hand, the second term in (16.13) does not vanish even <u>in vacuo</u> and therefore it would superficially suggest the paradoxical idea of a volume force on the vacuum. This term has evoked a great deal of speculation and does fit into an ether theory in which vacuum is supposed to be endowed with various mechanical properties which, among other things, enable it to transmit elastic waves and which also enable it to sustain body forces. The body forces can be transmitted from the ether to matter across the boundaries between the matter and the ether. The only way that such an ether force could be measured would be by means of the action of the ether on matter.

According to Lorentz' electron theory the only force which has physical existence is a resultant force which arises from the space-time average forces acting on material charges and currents, namely, forces obtained by averaging:

$$(16.15)\quad \vec{F} = \rho(\vec{E} + \vec{u} \times \vec{B})$$

Also, according to the special theory of relativity, no measurement can be devised which can determine the velocity or other properties of the ether and therefore its existence cannot be established experimentally, and no physical law can be contingent upon its existence. Therefore if we adopt the point of view that the only volume force which has a place in a physical theory is a force which is derivable from the Lorentz force (16.15), then the second term of (16.13) must

be subtracted out. We then have for the volume force, when $k = \mu = 1$, which is equal to the Lorentz force:

$$(16.16) \qquad F_{vi} = \frac{\partial T_{ij}}{\partial X_j} - \frac{1}{c^2}\frac{\partial N_i}{\partial t}$$

If we apply this equation to a volume containing both matter and radiation and bounded by a finite surface, (16.16) can be written in the form:

$$(16.18)^* \qquad F_i = \iint T_{ij}\, dS_j - \frac{1}{c^2}\frac{\partial}{\partial t}\iiint N_i\, dv$$

Since the body integrated force F_i represents the total rate of change of mechanical momentum P_i, of the volume, (16.18) can be written as:

$$(16.19) \qquad \frac{d}{dt}\left[P_i + \frac{1}{c^2}\iiint N_i\, dv \right] = \iint T_{ij}\, dS_j$$

This equation states that the sum of the rate of change of the mechanical momentum, of a particular volume, plus a term equal to the volume integral over this volume, of the Poynting vector divided by c^2, is equal to the surface integral of the total Maxwell stress transmitted across the surface, surrounding this volume. If it were possible to choose a surface that was large enough so that it was in field-free space, then the sum of the mechanical momentum and the Poynting vector volume integral term would be constant in time for there would be no stress transmitted across the integration surface to change the momentum of the system. This implies that the correction term, whose introduction into (16.16) was demanded by the physical reality of only the volume force on matter, makes a change in our concept of momentum necessary.

In the absence of measurable physical properties for the ether we are forced to modify the law of the conservation of momentum by having it not only apply to the momentum of matter alone but also to include a momentum density of the electromagnetic radiation field which is equal to the Poynting vector divided

---

* We have omitted a term which, in matter, is given by $\frac{(k\mu - 1)}{c^2}\frac{\partial}{\partial t}\iiint N_i\, dv$ which is actually a matter term present when an electro-magnetic wave travels through matter. Its net impulse due to a finite wave-train always vanishes.

by the square of the velocity of light. The Poynting vector therefore appears in a dual role, both as carrying energy and also as carrying momentum. It will turn out in the special theory of relativity that the property, which transfers the energy, also corresponds to a transfer of momentum in the proportions that have been derived here. Actually, this is a more general property which must be true of all forms of energy flow.

A new consequence of the introduction of the surface term into the conservation laws is the possibility of balancing energy and momentum over part of a system only. In the case of balancing energy the surface integral over the Poynting vector permits obtaining conservation even over parts of a system, whether radiative processes are present or not. To illustrate this point, consider the simple process of a battery ($\vec{E}'$) feeding a current ($\vec{j}$) to a resistor (σ), Fig. (16.1). Consider the energy balance

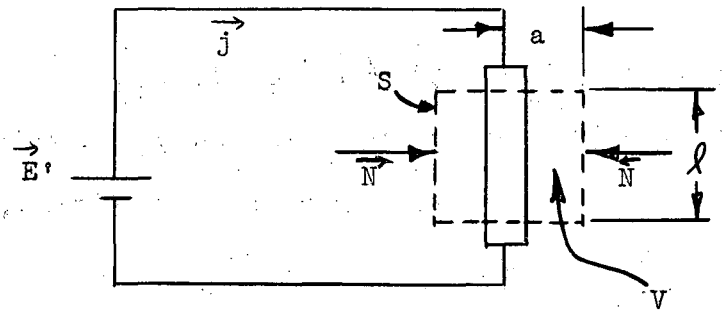


FIGURE (16.1)

over a volume V bounded by a cylindrical surface S of lenth ℓ and radius a as shown. By elementary considerations:

$$(16.20)\quad \vec{E} = \vec{j}/\sigma \;;\quad H = \frac{1}{2}\, j\, a \quad ;\quad N = \frac{j^2}{2\sigma}\, a$$

where $\vec{N}$ is directed outward. Hence:

$$(16.21)\quad \iint \vec{N}\cdot d\vec{S} = -\frac{j^2(\pi a^2 \ell)}{\sigma} = -\iiint \frac{j^2}{\sigma}\, dv = -\text{ Joule Heat.}$$

Energy is thus balanced considering the field as "feeding" the resistor via the Poynting vector, without explicitly introducing the source of the energy, i.e., the battery ($\vec{E}'$).

CHAPTER 17 RADIATION PRESSURE

We have seen that the ether theory and the Lorentz electron theory or special theory of relativity differ in their points of view. In the ether theory, forces are acting on the ether and the ether is acting mechanically on matter, while in relativity theory no forces are acting on the ether. In fact, there is no ether. But in the relativity theory electromagnetic radiation carries momentum which in addition to the mechanical momentum of the matter involved is conserved. Let us investigate how these considerations effect some explicit cases of the interaction of radiation with matter.

Let us first consider a plane polarized plane wave incident normally upon a slab of material which absorbs the momentum of the electromagnetic wave fully without reflection. Let the y coordinate be parallel to the electric field of the incident wave, and the z coordinate be parallel to the magnetic field of the incident wave with the wave traveling along the x direction. The Maxwell stress tensor is:

$$(16.8)\quad T_{ij} = E_i D_j + H_i B_j - \frac{\delta_{ij}}{2}\left[E_k D_k + H_k B_k\right]$$

The stress tensor has only three non-vanishing components which are all located on the major diagonal of the matrix. These terms are given by:

$$(17.1)\quad T_{xx} = -\frac{1}{2}\left[E_y D_y + H_z B_z\right] = -g$$

$$T_{yy} = E_y D_y - \frac{1}{2}\left[E_y D_y + H_z B_z\right] = E_y D_y - g$$

$$T_{zz} = H_z B_z - \frac{1}{2}\left[E_y D_y + H_z B_z\right] = H_z B_z - g$$

This results in the matrix:

$$T = \begin{pmatrix} -g & 0 & 0 \\ 0 & (E_y D_y - g) & 0 \\ 0 & 0 & (H_z B_z - g) \end{pmatrix}$$

The force derived from this stress matrix, using the correction term in (16.13) which makes the volume force agree with the Lorentz force in a material medium,

using (16.16):

$$(17.3)\quad F_{vi} = \frac{\partial T_{ij}}{\partial X_j} - \frac{\mu k}{c^2}\frac{\partial [(\vec{E} \times \vec{H})i]}{\partial t}$$

The three components of the force are:

$$(17.4)\quad F_{vx} = -1/2\,\frac{\partial}{\partial x}\left[E_y D_y + H_z B_z\right] - \frac{\mu k}{c^2}\frac{\partial}{\partial t}\left[E_y H_z\right]$$

$$F_{vy} = 1/2\,\frac{\partial}{\partial y}\left[E_y D_y - H_z B_z\right] = 1/2\,\frac{\partial}{\partial y}\left[kk_o E_y^2 - \mu\mu_o H_z^2\right]$$

$$F_{vz} = 1/2\,\frac{\partial}{\partial z}\left[H_z B_z - E_y D_y\right] = 1/2\,\frac{\partial}{\partial z}\left[\mu\mu_o H_z^2 - kk_o E_y^2\right]$$

The y and z components of the force will vanish over the slab of material upon which the $\vec{E}$ and $\vec{H}$ fields are impinging, since the fields are not functions of y or z. The Poynting vector has no y or z components. The x component, however, does not vanish, and in fact can be integrated, giving a total time average pressure on the slab of material:

$$(17.5)\quad \bar{P} = \int_o^\infty F_{vx}\,dx = 1/2\left[kk_o E^2 + \mu\mu_o H^2\right]$$

which is exactly equal to the negative of the energy density of the incoming radiation field.

The term depending on the time has been omitted for the reason that, if E_y and H_z of the incoming wave vary sinusoidally, then when the time average of the pressure is taken, the time derivative term would only contribute a transient variation, while the energy density term, will give a secular variation. This means for a sufficiently large averaging period for a continuous wave train, that the time-dependent term does not contribute to $\bar{P}$. If we are considering the net impulse that is transmitted by a wave train of finite length, then the pressure must be integrated from time equal to minus infinity to time equal to plus infinity, and in this case the time-derivative term will also integrate out. The only case in which the time-dependent term will contribute is to the instantaneous value of the pressure during the absorption of a wave train. The time-dependent term will give rise to a fluctuation term which represents the fluctuation of

the momentum of the radiation field. The addition of this term therefore does not result in any experimental differences in the radiation pressure from the value that it would have according to the ether theory where this term is absent.

We can summarize these results by stating that the phenomena of radiation pressure is in accordance with the concept of momentum of electromagnetic waves and the more general concept of momentum carried by any energy transmitting process. However, the radiation pressure considerations do <u>not</u> disagree with the results of the classical prerelativity ether theory. Since the radiation pressure that has been computed here by general considerations must be identical with the Lorentz force, it must be possible to compute it directly by calculating the induced currents, in the medium that absorbs an electromagnetic field, if a particular model of absorption is used, such as, for instance, finite conductivity. This calculation can be carried out easily and in all cases does give a value for the radiation pressure that is equal to the energy density of the incident radiation.

As a second example let us consider the radiation pressure produced by radiation that is non-polarized and traveling in a random direction when it strikes a material medium. By symmetry the only non-vanishing component of the volume force (17.3) is the component of the force normal to the surface which we shall designate by the subscript $_1$. It is given by:

$$(17.6)\quad F_1 = \frac{\partial T_{1j}}{\partial X_j} = \frac{\partial}{\partial X_j}\left[E_1 D_j + H_1 B_j - \frac{\delta_{ij}}{2} (E_k D_k + H_k B_k)\right]$$

which can be written as:

$$(17.7)\quad F_1 = kk_o \left\{ \frac{\partial (E_1^2)}{\partial X_1} + \frac{\partial}{\partial X_2} (E_1 E_2) + \frac{\partial}{\partial X_3} (E_1 E_3)\right\} + \mu\mu_o\left\{\frac{\partial H_1^2}{\partial X_1} + \frac{\partial}{\partial X_2} (H_1 H_2) + \frac{\partial}{\partial X_3} (H_1 H_3)\right\} - \frac{kk_o}{2} \frac{\partial E^2}{\partial X_1} - \frac{\mu\mu_o}{2} \frac{\partial H^2}{\partial X_1}$$

Due to the fact that the i^{th} and the j^{th} components of the electric field are uncorrelated, the time average value of all of the cross terms vanishes,

$\overline{E_1 E_2} = \overline{E_1 E_3} = \overline{H_1 H_2} = \overline{H_1 H_3} = 0$, we obtain:

$$(17.8)\quad \overline{F_1} = kk_0 \overline{\frac{\partial E_1^2}{\partial X_1}} = \frac{kk_0}{2} \overline{\frac{\partial E^2}{\partial X_1}} + \mu\mu_0 \overline{\frac{\partial H_1^2}{\partial X_1}} - \frac{\mu\mu_0}{2} \overline{\frac{\partial H^2}{\partial X_1}}$$

Also since in the random orientation of the fields the square of each component has the same a priori probability, we have:

$$(17.9)\quad \overline{E_1^2} = 1/3\ \overline{E^2}$$

Therefore the value of the normal force becomes:

$$(17.10)\quad \overline{F_1} = \frac{\partial}{\partial X_1}\left\{ kk_0 \overline{E^2}\left[\frac{1}{3} - \frac{1}{2}\right] + \mu\mu_0\ \overline{H^2}\left[\frac{1}{3} - \frac{1}{2}\right]\right\} = -\frac{1}{6}\frac{\partial}{\partial X_1}\left[kk_0 E^2 + \mu\mu_0 H^2\right] = -\frac{\partial}{\partial X_1}\left(\frac{U}{3}\right)$$

where U is the energy density of the incident radiation. Integrating from the absorbing medium into field-free space, we obtain the result that the total radiation pressure is equal to 1/3 of the energy density of the incident wave:

$$(17.11)\quad \overline{P} = 1/3\ U$$

This theorem gives what might be called an equation of state for radiation in a space bounded by absorbing walls. This equation of state forms the basis of thermodynamic derivations of the Stefan-Boltzmann law and the Wien displacement laws for black body radiation.

CHAPTER 18 ELECTROMAGNETIC WAVE EQUATIONS

We shall now consider the form taken by Maxwell's equations if they are reduced from a system of four first order linear partial differential equations to a system of two second order linear partial differential equations. Maxwell's equations are:

(14.7)
$$(1)\quad \vec{\nabla} \cdot \vec{D} = \rho_{true}$$
$$(2)\quad \vec{\nabla} \cdot \vec{B} = 0$$
$$(3)\quad \vec{\nabla} \times \vec{E} = -\frac{\partial \vec{B}}{\partial t}$$
$$(4)\quad \vec{\nabla} \times \vec{H} = \vec{j}_{true} + \frac{\partial \vec{D}}{\partial t}$$

Consider a region where there are no true charges and no sources of EMF so that $\rho_{true} = 0$, and $\vec{E}' = 0$, and where k and μ are not functions of the coordinates. Take the curl of Equation (14.7)(3) and substitute $\mu\mu_o\vec{H}$ for $\vec{B}$:

$$(18.1)\quad \vec{\nabla} \times (\vec{\nabla} \times \vec{E}) = -\frac{\partial}{\partial t}\left[\vec{\nabla} \times (\mu\mu_o \vec{H})\right]$$

Substituting from (14.7)(4), we have:

$$(18.2)\quad \vec{\nabla} \times (\vec{\nabla} \times \vec{E}) = -\mu\mu_o \frac{\partial}{\partial t}\left(\vec{j}_{true} + kk_o\frac{\partial \vec{E}}{\partial t}\right)$$

Using the vector relation:

$$(18.3)\quad \vec{\nabla} \times (\vec{\nabla} \times \vec{E}) = \vec{\nabla}(\vec{\nabla} \cdot \vec{E}) - \nabla^2 \vec{E}$$

Since $\vec{\nabla} \cdot \vec{E} = 0$ in the charge-free field, and using (9.3) we obtain from (18.2):

$$(18.4)\quad \nabla^2\vec{E} - \frac{\mu k}{c^2}\frac{\partial^2 \vec{E}}{\partial t^2} - \mu\mu_o \sigma \frac{\partial \vec{E}}{\partial t} = 0$$

where:

$$(16.14)\quad \mu_o k_o = \frac{1}{c^2}$$

Equation (18.4) is known as the general wave equation. Usually only the second or the third term individually in connection with the first term is used in the solution of a particular application of this equation. In a non-conducting medium the third term vanishes, giving rise to a wave equation for waves that travel with the velocity $v = 1/\sqrt{\mu\mu_o kk_o}$. In a conducting medium, the second

term is usually negligible, and we are left with the differential equation for heat conduction or diffusion. The relative magnitude of the two terms can be easily estimated by assuming that the field varies sinusoidally:

(18.5) $\vec{E} = \vec{E}_o e^{i\omega t}$

When this is substituted into the wave equation (18.4) we get:

(18.6) $\nabla^2 \vec{E}_o + \frac{\mu k}{c^2} \vec{E}_o \omega^2 - i\mu\mu_o\sigma\omega\vec{E}_o = 0$

We can re-express the coefficients in various ways. If, for example, we introduce the wavelength, divided by 2π, in free space $\lambda\!\!\!^{-}_o$, and $\lambda\!\!\!^{-}$ in the medium, then we have the relations:

(18.7) (1) In free space, $\mu = k = 1$, $\frac{1}{\lambda\!\!\!^{-}_o} = \frac{\omega}{c}$

(2) In a medium, $\mu \neq 1$, $k \neq 1$, $\frac{1}{\lambda\!\!\!^{-}} = \frac{\omega}{u}$ where u = the velocity of the electromagnetic wave in the medium.

(3) $c = \sqrt{\mu k}\ u$

(4) $\frac{1}{\lambda\!\!\!^{-2}} = \frac{\omega^2}{u^2} = \frac{\omega^2}{c^2}\mu k$

The relaxation time of the dielectric, which we discussed in Chapter 9, is:

(9.20) $\tau = \frac{kk_o}{\sigma}$

Substituting (18.7)(4) and (9.20) into (18.6) we have:

(18.8) $\nabla^2 \vec{E}_o + \left[1 - \frac{i}{\omega\tau}\right]\frac{\vec{E}_o}{\lambda\!\!\!^{-2}} = 0$

Hence, if the relaxation time τ is long compared to the period $2\pi/\omega$ of the sinusoidal vibration, then we have essentially a propagation equation, since the imaginary term drops out. On the other hand, if the relaxation time is short compared to the period, then the imaginary term is large compared to unity and we have essentially a diffusion condition.

For all pure metals the relaxation time is of the order of 10^{-14} seconds, so that the diffusion type of equation is valid from zero frequency all the way into the optical frequency region. This means that in metallic conductors even in the ultra high frequency radio region the propagation term $\frac{\mu k}{c^2}\frac{\partial^2 \vec{E}}{\partial t^2}$ can be omitted from the wave equation (18.4). To put it in different terms, the

displacement current is negligible relative to the conduction current in metals at the highest frequencies that are theoretically attainable with macroscopic oscillators. In the optical frequency region, both terms may become important, due to the fact that the displacement current term and the conduction current term become of the same order of magnitude and therefore indistinguishable. The binding and the inertia of the electrons in the metal introduce phase lags between the electronic motions and the incident electromagnetic fields. The two terms do not physically represent different phenomena in the optical frequency region, but simply represent a current with one component in phase with the field and another component that is in quadrature with the field. The quadrature current component is physically not a displacement current term, but it effects the propagation equations in the same way as a displacement current would. In the optical region therefore, the distinction between the dielectric constant and the conductivity is a purely formal one. They represent respectively the real and the imaginary parts of a complex dielectric constant.

The ratio between the magnitude of the conduction current and the magnitude of the displacement current can be rewritten from the relations of (18.7) as:

$$(18.9) \qquad \frac{1}{\omega\tau} = \frac{\sigma}{\omega k k_o} = \frac{\sigma}{k k_o}\frac{\lambda}{u} = \sigma\lambda\sqrt{\frac{\mu\mu_o}{k k_o}} = \sigma\lambda R_o\sqrt{\frac{\mu}{k}} \quad .$$

where:

$$(18.10) \qquad R_o = \sqrt{\frac{\mu_o}{k_o}}$$

is a resistance whose numerical value is 376.7 ohms. This number is sometimes called the characteristic impedance of free space. It can be shown that a conducting sheet whose resistance is R_o ohms per square★ has exactly the same reflecting conditions on the incident wave as has free space itself, provided a reflector is placed a quarter wavelength behind such a sheet so as to effectively make the impedance behind the sheet infinite. Such concepts are, of course, useful in practical considerations, but we shall avoid extensive discussions of

★The resistance of a square sheet is independent of the dimensions of the sheet.

such analogies in favor of considering the characteristics of the electromagnetic fields themselves.

The significance of the ratio (18.9) might be stated by saying that if the resistance of a cube of a medium, whose edge is λ in length, is larger than R_o, then in such a medium the displacement current is dominant, while if the reverse is true, the conduction current governs the behavior of electromagnetic fields in this medium.

Let us introduce the sympolic operator $\Box$, known as the D'Alembertian, defined by:

$$(18.11)\quad \Box = \nabla^2 - \frac{\mu k}{c^2}\frac{\partial^2}{\partial t^2}$$

The wave equation (18.4) is then reduced to:

$$(18.12)\quad \Box \vec{E} - \mu\mu_o\sigma\frac{\partial \vec{E}}{\partial t} = 0$$

An identical equation can be shown to hold for $\vec{B}$ by taking the curl of (14.7)(4) and using a process that is analogous to the one that was used to derive (18.4). Let us now re-express these equations in terms of the potentials of the electromagnetic field. Due to the relation expressed by the induction equation (14.7)(3), it is no longer possible to derive the electric field solely from a scalar potential ϕ, but the magnetic field as a result of (14.7)(2) is still derivable from a vector potential by the equation:

$$(11.16)\quad \vec{B} = \vec{\nabla} \times \vec{A}$$

If Equation (11.16) is assumed to hold, then the electric field can be derived from the sum of the gradient of a scalar potential and a supplementary non-conservative contribution from the rate of change of the vector potential. That is, we may derive $\vec{E}$ from the scalar and the vector potential by:

$$(18.13)\quad \vec{E} = -\vec{\nabla}\phi - \frac{\partial \vec{A}}{\partial t}$$

which makes $\vec{E}$ conform to the relation (14.7)(3). As before, the divergence of $\vec{A}$ remains undefined, or at least remains undefined within an additive arbitrary

function of position. Let us now define the divergence of $\vec{A}$ by the so-called Lorentz condition:

$$(18.14)\quad \vec{\nabla}\cdot\vec{A} + \frac{\mu k}{c^2}\frac{\partial \phi}{\partial t} + \mu\mu_0\sigma\phi = 0$$

which in free space becomes:

$$(18.15)\quad \vec{\nabla}\cdot\vec{A} + \frac{1}{c^2}\frac{\partial \phi}{\partial t} = 0$$

The Lorentz condition appears to be a very arbitrary introduction of a subsidiary relation. As we shall see, it has the advantage of introducing complete symmetry between the scalar and the vector potentials, namely, it makes the scalar and the vector potentials obey the same differential equation (18.12) as those obeyed by the fields. It turns out that the Lorentz condition assures a relativistic covariant relation between the scalar and the vector potential. If we introduce the defining equations (11.16) and (18.13) and the Lorentz condition into the wave eqaution (18.12) and simplify, we obtain the symmetrical set of equations:

$$(18.16)\quad \nabla^2\vec{A} - \frac{\mu k}{c^2}\frac{\partial^2 \vec{A}}{\partial t^2} - \mu\mu_0\sigma\frac{\partial \vec{A}}{\partial t} = -\mu\mu_0\vec{j'}$$

$$(18.17)\quad \nabla^2\phi - \frac{\mu k}{c^2}\frac{\partial^2\phi}{\partial t^2} - \mu\mu_0\sigma\frac{\partial\phi}{\partial t} = -\frac{\rho}{kk_0}$$

Here $\vec{j'}$ represents a current given by $\vec{j'} = \sigma\vec{E'}$, that is, only that part of the current density which is produced by the electromotive forces and does not contain any part of the current that is induced by the electric fields in the conducting medium itself. In free space these equations become simply:

$$(18.18)\quad \Box\vec{A} = -\mu_0\vec{j'}$$

$$(18.19)\quad \Box\phi = -\frac{\rho}{k_0}$$

where $\vec{j'}$ and ρ are the sources of the electric field and are produced by external agents. These equations are known as the inhomogeneous wave equations. Their complementary solutions, namely the solutions of the equations:

$$(18.20)\quad \Box\vec{A} = 0$$

$$(18.21)\quad \Box\phi = 0$$

will be shown in general to be wave solutions. The particular solutions of

Equations (18.18) and (18.19) are expressible in terms of integrals over the charge and current distributions. We are therefore interested in obtaining both the homogeneous, complementary solutions, and also the particular source solutions, of Equations (18.18) and (18.19).

By inspection of Equations (11.16) and (18.13) it can be seen that the resultant electric and magnetic fields are unchanged by transformations of the type:

$$(18.22)\quad \vec{A}' = \vec{A} - \vec{\nabla}\Psi$$

$$(18.23)\quad \phi' = \phi + \frac{\partial \Psi}{\partial t}$$

where Ψ is a function of the coordinates and the time. This means that if any physical law involving electromagnetic interaction is to be expressed in terms of the general electrodynamic potentials $\vec{A}$ and ϕ then such a physical law must also be unaffected by a transformation of the type (18.22) or (18.23). These transformations are usually known as gauge transformations and a physical law that is invariant under such a transformation is said to be gauge invariant. The property of gauge invariance, if possessed by a physical law, insures that this physical law will not lead to consequences that cannot be formulated in terms of the interactions of charges and currents in terms of electromagnetic fields.

Note that the velocity $c = 1/\sqrt{\mu_o k_o}$ of electromagnetic waves in vacuo enters into the equations as a characteristic constant of the theory <u>per se</u>, not as a constant describing a particular physical parameter entering the theory. This is a feature of Maxwell's equations at variance with the laws of classical mechanics which contains no characteristic constants. Classical mechanics can be scaled freely with respect to all physical quantities; Maxwell's equations can be scaled in relation to length and time individually but not as to velocity. It is this fact which indicates that the laws of classical electrodynamics are actually relativistic in the sense of introducing a characteristic velocity.

CHAPTER 19 SOLUTIONS OF THE ELECTROMAGNETIC WAVE EQUATIONS

Plane Wave Solutions in Stationary Media

Let us consider first the solution of the homogeneous equations:

(18.20) $\Box \vec{A} = 0$

(18.21) $\Box \phi = 0$

in the case in which all fields are functions only of the distance of a given plane from the origin as in Figure (19.1). If this distance is ξ and if $\vec{r}$ is the radius vector from the origin to a point on the plane, and if $\vec{n}$ is a unit

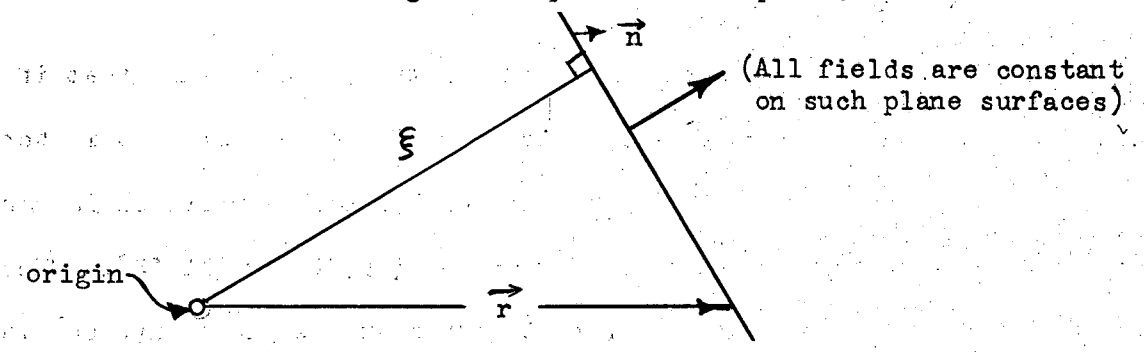


Figure (19.1)

vector normal to the plane, then all operational derivatives are functions of ξ only and the $\vec{\nabla}$ operator becomes:

(19.1) $\vec{\nabla} \longrightarrow \vec{n} \frac{\partial}{\partial \xi}$

and Maxwell's equations (14.7) in the absence of charges, become:

(19.2) (1) $\vec{n} \cdot \frac{\partial \vec{D}}{\partial \xi} = 0$

(2) $\vec{n} \cdot \frac{\partial \vec{B}}{\partial \xi} = 0$

(3) $\vec{n} \times \frac{\partial \vec{E}}{\partial \xi} = - \frac{\partial \vec{B}}{\partial t}$

(4) $\vec{n} \times \frac{\partial \vec{H}}{\partial \xi} = \vec{j} + \frac{\partial \vec{D}}{\partial t}$

If we take the scalar product of $\vec{n}$ and the fourth equation of (19.2) and use the relations (3.11) and (9.3) we have:

(19.3) $\vec{n} \cdot \left[\frac{\sigma}{k k_0} + \frac{\partial}{\partial t} \right] \vec{D} = 0$

Equations (19.3) and (19.2)(1) imply that the longitudinal components of $\vec{D}$ and $\vec{E}$, that is those components that are perpendicular to the plane surface in Figure (19.1), are independent of ξ and that their time dependence follows an exponential decay law in accordance with the characteristic relaxation time of the medium. Thus:

$$(19.4) \quad E_n = E_{no}\, e^{-t/\tau} = E_{no}\, e^{-\frac{\sigma t}{kk_o}}$$

This means that the only longitudinal solution of the field equations is an electrostatic solution and that, in the presence of finite conductivity, the electrostatic solutions will vanish exponetially in time.

Performing the same scalar product operation on (19.2)(3) we have:

$$(19.5) \quad \vec{n} \cdot \frac{\partial \vec{B}}{\partial t} = 0$$

Equations (19.5) and (19.2)(2) show that the only solution compatible with the field equations, for the magnetic field component normal to the plane surface in Figure (19.1), is a stationary uniform magnetic field.

If there is a non-static part of the wave, it must be composed of transverse fields or fields whose vectors lie parallel to the plane. The velocity with which this transverse wave is propagated in the direction $\vec{n}$, if $\sigma = 0$, is given by:

$$(19.6) \quad u = \frac{c}{\sqrt{\mu k}}$$

Each transverse wave component obeys the one-dimensional wave equation:

$$(19.7) \quad \frac{\partial^2 \vec{E}}{\partial \xi^2} - \frac{1}{u^2} \frac{\partial^2 \vec{E}}{\partial t^2} = 0$$

which may be derived by eliminating $\vec{H}$ between (19.2)(3) and (19.2)(4), securing the "Telegraphers' Equation," and then setting $\sigma = 0$. The general solution of (19.7) is:

$$(19.8) \quad E = g(\xi - ut) + f(\xi + ut)$$

where g and f are arbitrary functions. If $\vec{E}$ is assumed to have a sinusoidal time variation, the solution of (19.7) is:

$$(19.9) \quad E = E_o\, e^{i(K\xi \pm \omega t)}$$

where K is given by:

$$(19.10)\quad \vec{K} = \frac{\omega}{u^2}\vec{u} = \frac{1}{\lambda}\frac{\vec{u}}{u}\,, \quad K = \frac{\omega}{u} = \frac{1}{\lambda}$$

By general theorems in Fourier analysis, the general solution of the homogeneous wave equation can be expanded in terms of Fourier integrals over the three components of the wave propagation vector $\vec{K}$, which is related to the frequency by the relation (19.10). The Fourier integrals are taken over all of the positive and all of the negative values of each component of $\vec{K}$. In addition, the resultant fields are summed over all possible polarizations. This gives the general expansion of the solution of the homogeneous wave equation in terms of a super-position of plane waves. This expansion has the form:

$$(19.11)\quad \vec{E} = \sum_{j=1}^{3}\left[\vec{x}_j \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} a^j_K\, e^{i(\vec{K}\cdot\vec{r} - \omega t)}\, dK_1\, dK_2\, dK_3\right]$$

The $\vec{x}_j$ are unit vectors in the three coordinate directions. The a^j_K are amplitude functions of the frequency for the three coordinate directions, which may be complex to give arbitrary starting phases.

In a similar way, it is also possible to generate a solution of the homogeneous wave equations in terms of an expansion in spherical waves, or in terms of an expansion in cylindrical waves, that is fields whose magnitude is dependent only on distance and azimuth relative to a given axis. Which of these expansions is the more convenient one to use depends on the symmetry properties of the problem being considered.

Plane Wave Solutions in Moving Media

Let us now consider a plane wave solution in one dimension in case the medium in which the electromagnetic wave is being propagated is in motion with a velocity $\vec{v}$ relative to the observer. The phase velocity of propagation that results under these conditions was measured in the Fresnel-Fizeau convection coefficient experiment. The experimental result was that the phase velocity was given by

an equation of the type:

$$(19.12) \qquad u = u_o + \left(1 - \frac{1}{n^2}\right) \vec{v} \cdot \vec{n}$$

$n = \sqrt{\mu k}$ = the index of refraction of optics

$u_o = \dfrac{c}{\sqrt{\mu k}}$ = the velocity of the wave in the medium if the medium is at rest, $\vec{v} = 0$.

$\vec{n}$ = unit vector in direction $\vec{u}$.

We shall now show that Equation (19.12), which will later be derived relativistically, is also in agreement with Maxwell's equations in moving media, provided that we interpret the velocity $\vec{v}$ as the relative velocity of the medium to the frame in which Maxwell's equations are valid, i.e. to the frame in which the free space velocity would be c. Maxwell's equations, in the absence of permeable materials and in the absence of true charges and currents, in moving media from (14.19) are:

$$(19.13) \qquad (1) \ \vec{\nabla} \cdot \vec{D} = 0$$

$$(2) \ \vec{\nabla} \cdot \vec{B} = 0$$

$$(3) \ \vec{\nabla} \times \vec{E} = -\frac{\partial B}{\partial t}$$

$$(4) \ \vec{\nabla} \times \vec{B} = \mu_o \left[\frac{\partial \vec{P}}{\partial t} + k_o \frac{\partial \vec{E}}{\partial t} + \vec{\nabla} \times (\vec{P} \times \vec{v})\right]$$

The polarization is given by the effective field in the moving frame by:

$$(14.20) \qquad \vec{P} = k_o(k-1)(\vec{E} + \vec{v} \times \vec{B})$$

Note that all of the fields in (19.13) and (14.20) are the fields which would be measured in the stationary frame of the observer. Note also that any results derived here will be carried out only to the first order in v/c. Let us substitute $\vec{P}$ the polarization field (14.20) into the fourth equation of (19.13) and using (19.13)(3) we have:

$$(19.14) \qquad \vec{\nabla} \times \vec{B} = \mu_o \left[k_o \frac{\partial \vec{E}}{\partial t} + k_o(k-1) \frac{\partial \vec{E}}{\partial t} + k_o(k-1) \left\{\vec{v} \times \frac{\partial \vec{B}}{\partial t} + \vec{\nabla} \times (\vec{E} \times \vec{v})\right\}\right]$$

$$\vec{\nabla} \times \vec{B} = \frac{k}{c^2}\left[\frac{\partial \vec{E}}{\partial t} + \left(1 - \frac{1}{k}\right)\left\{-\vec{v} \times (\vec{\nabla} \times \vec{E}) + \vec{\nabla} \times (\vec{E} \times \vec{v})\right\}\right]$$

For a plane wave solution using (19.1) this becomes:

$$(19.15)\quad \vec{n} \times \frac{\partial}{\partial \xi} \vec{B} = \frac{k}{c^2}\left[\frac{\partial \vec{E}}{\partial t} + \left(1 - \frac{1}{k}\right)\left\{-\vec{v} \times \left(\vec{n} \times \frac{\partial \vec{E}}{\partial \xi}\right) + \vec{n} \times \left(\frac{\partial \vec{E}}{\partial \xi} \times \vec{v}\right)\right\}\right]$$

Note that $\vec{v}$ is independent of the coordinates. Expanding the triple cross product, and, since the wave is transverse, omitting the term $\vec{n} \cdot \frac{\partial \vec{E}}{\partial \xi}$*, we obtain, by taking the curl of both sides of (19.15):

$$(19.16)\quad -\frac{\partial^2 \vec{B}}{\partial \xi^2} = \frac{k}{c^2}\,\vec{n} \times \frac{\partial}{\partial \xi}\left[\frac{\partial \vec{E}}{\partial t} - \left(1 - \frac{1}{k}\right) 2(\vec{v} \cdot \vec{n})\frac{\partial \vec{E}}{\partial \xi}\right]$$

Let us make the additional approximation that in the last term:

$$(19.17)\quad \frac{\partial}{\partial \xi} = -\frac{1}{u_o}\frac{\partial}{\partial t}$$

where the velocity of the wave would be u_o in the medium if the medium were at rest. This approximation is justified since we are only interested in the correction term to the velocity. Using (19.17) we have:

$$(19.18)\quad -\frac{\partial^2 \vec{B}}{\partial \xi^2} = -\frac{1}{u_o^2}\frac{\partial^2 \vec{B}}{\partial t^2}\left[1 - \left(1 - \frac{1}{k}\right)\frac{2\vec{v} \cdot \vec{n}}{u_o}\right]$$

which is the wave equation that corresponds to the propagation velocity:

$$(19.19)\quad u = u_o\left[1 - \left(1 - \frac{1}{k}\right)\frac{2\vec{v} \cdot \vec{n}}{u_o}\right]^{-1/2} \cong u_o + \left(1 - \frac{1}{k}\right)\vec{v} \cdot \vec{n}$$

which agrees to the order v/c with the Fizeau result given in (19.12). The physical interpretation of (19.19) is that the only part of the propagation velocity of the wave which is affected by the motion of the medium through which the wave is passing is that part which is proportional to:

$$(19.20)\quad 1 - \frac{1}{k} = \frac{k - 1}{k} \propto \frac{\partial \vec{P}}{\partial t} \Big/ \frac{\partial \vec{D}}{\partial t}$$

This fraction is proportional to the ratio of the polarization current to the displacement current. Since a polarization current does actually correspond to the motion of dipoles, it is quite reasonable to assume that that part of the wave which corresponds to these dipoles will be affected by the velocity of the medium. The effect of a medium on a plane wave is in general simply that

* We have also omitted a term $\vec{n}\left(\vec{v} \cdot \frac{\partial \vec{E}}{\partial \xi}\right)$; this term will vanish in the next step (Eq. (19.16)) due to the $\vec{n}$ x operation.

the medium is polarized by the incident wave and that the resulting dipoles retard the wave which is coherent with the incident wave such that the combination with the primary wave corresponds to the overall phase velocity. It is this coherent retarded component which is being radiated from a moving source in this case, and which gives rise to the Fresnel-Fizeau coefficient. Note that all these results are consequences of the classical non-relativistic electromagnetic theory. We shall later show that these results are also in accordance with relativistic principles. We shall not discuss other examples of problems involving plane wave propagation, such as boundary value problems, propagation in non-isotropic solids, etc. These problems are covered in optics courses.

Solution by Fourier Analysis

Let us concern ourselves with the integration of the inhomogeneous equations and let us investigate only the particular solutions. The inhomogeneous wave equations, for the vector and scalar potentials, in free space are:

$$(18.18)\quad \Box \vec{A} = -\mu_o \vec{j}'$$

$$(18.19)\quad \Box \phi = -\frac{\rho}{k_o}$$

subject to the free space Lorentz condition:

$$(18.15)\quad \vec{\nabla} \cdot \vec{A} + \frac{1}{c^2}\frac{\partial \phi}{\partial t} = 0$$

from which the fields are derived by:

$$(18.13)\quad \vec{E} = -\vec{\nabla}\phi - \frac{\partial \vec{A}}{\partial t}$$

$$(11.16)\quad \vec{B} = \vec{\nabla} \times \vec{A}$$

Let us review the solution of the analogous static problem. In the static case, (18.19) reduced to Poisson's equation:

$$(1.16)\quad \nabla^2\phi = -\frac{\rho}{k_o}$$

whose particular solution was:

$$(19.21)\quad \phi(p) = \frac{1}{4\pi k_o}\iiint \frac{\rho(p')\, dv'}{r(p,p')}$$

Note that the integrand is a function both of the point of observation p and the point of integration p', and that the time does not appear explicitly.

We are looking for a solution of (18.18) and (18.19) that is similar in form to (19.21). We wish to study the modification of the solution of (1.16) that is caused by the presence of the time-dependent term in (18.18) and (18.19). Both of these equations have the general form:

$$(19.22)\quad \Box \Psi(p,t) = -\, g(p,t)$$

Let us assume that the source function g(p,t) can be analyzed by the Fourier integral:

$$(19.23)\quad g(p,t) = \int_{-\infty}^{\infty} g_\omega(p) e^{-i\omega t}\, d\omega$$

which has the Fourier inversion:

$$(19.24)\quad g_\omega(p) = \frac{1}{2\pi}\int_{-\infty}^{\infty} g(p,t) e^{+i\omega t}\, dt$$

Let us assume similarly that the general potential $\Psi(p,t)$ can be analyzed into Fourier components by:

$$(19.25)\quad \Psi(p,t) = \int_{-\infty}^{\infty} \Psi_\omega(p) e^{-i\omega t}\, d\omega$$

with a corresponding inverse relation:

$$(19.26)\quad \Psi_\omega(p) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \Psi(p,t) e^{+i\omega t}\, dt$$

By substitution of (19.23) and (19.25) into (19.22) we see that the Fourier component $\Psi_\omega(p)$ obeys the differential relation:

$$(19.27)\quad \nabla^2 \Psi_\omega + \frac{\omega^2}{c^2}\Psi_\omega = -\, g_\omega$$

which is similar to Poisson's equation. Let us synthesize the solution of (19.27) out of the superposition of a set of unit point solutions corresponding to a source at the point p', with the source $g_\omega(p) = \delta(p-p')$ which obey the equation:

$$(19.28)\quad \nabla^2 G(p,p') + \frac{\omega^2}{c^2} G(p,p') = -\, \delta(p - p')$$

where $\delta(p - p')$ is the Dirac δ function defined in (5.48) and (5.49). The resultant unit source potential G(p,p') will be a function both of the points

p and p'. The partial solution corresponding to the frequency ω of the total source is then given by the superposition:

$$(19.29)\quad \Psi_\omega(p) = \iiint g_\omega(p')\, G(p,p')\, dv'$$

If we let r represent the distance between the point p and the point p', the resultant solution of (19.28) will be spherically symmetric in r, and hence the solution at every point other than r = 0 will be the solution of the differential equation:

$$(19.30)\quad \frac{1}{r}\frac{\partial^2}{\partial r^2}(rG) + K^2 G = 0 \qquad K = \frac{\omega}{c}$$

which integrates immediately into:

$$(19.31)\quad G = \frac{A}{r}\, e^{\pm iKr}$$

The solution is valid everywhere except at r = 0. To evaluate the constant A let us consider the volume integral of the differential equation (19.28); with the definition of the δ function (5.48) we obtain:

$$(19.32)\quad \iiint \nabla^2 G\, dv + K^2 \iiint G\, dv = -1$$

The second integral vanishes, as the integration volume is shrunk to zero, since the singularity of G at r = 0 is only of order 1/r. The first term can be transformed by means of Green's theorem to give:

$$(19.33)\quad \iint \vec{\nabla} G \cdot d\vec{S} = -1$$

Operating on the solution (19.31) with the $\vec{\nabla}$ operator, and then substituting into (19.33) we obtain:

$$(19.34)\quad \iint \left[-\frac{\vec{r}}{r^3} A \mp i \frac{\vec{r}}{r^2} A K \right] e^{\pm iKr} \cdot d\vec{S} = -1$$

which in the limit as $r \longrightarrow 0$ becomes:

$$(19.35)\quad -4\pi A = -1$$

So:

$$(19.36)\quad G(p,p') = \frac{1}{4\pi r}\, e^{\pm iKr} \quad \text{is the solution of (19.28).}$$

Substituting into (19.29) we obtain:

$$(19.37)\quad \Psi_\omega(p) = \frac{1}{4\pi}\iiint \frac{g_\omega(p')}{r(p,p')}\, e^{\pm iKr(p,p')}\, dv'$$

The effect of the second term in (19.27) is therefore the introduction of the exponential factor in the Coulomb type integral (19.21). If we use (19.25) to resynthesize the time-dependent potential function in terms of its Fourier components we have:

$$(19.38)\quad \Psi(p,t) = \int \Psi_\omega(p) e^{-i\omega t}\, d\omega = \frac{1}{4\pi}\iiint\!\!\iint \frac{g_\omega(p') e^{-i(\omega t \pm Kr)}}{r(p,p')}\, d\omega\, dv'$$

If we introduce a new time defined by:

$$(19.39)\quad t'(p,p') = t \pm \frac{r}{c} = t \pm \frac{Kr}{\omega}, \text{ where } c = \frac{\omega}{K}$$

The new time t' corresponds to shifting the origin of time by an amount equal to the time that it takes a light signal to be propagated from point p to point p'. The Fourier transformation can be evaluated by using (19.23) giving:

$$(19.40)\quad \Psi(p,t) = \frac{1}{4\pi}\iiint \frac{g(p', t \pm \frac{r}{c})}{r(p,p')}\, dv'$$

Note that now the point of observation p is contained explicitly both in the 1/r term and also by means of (19.39) in the time at which the time-varying currents or charges are introduced into the integration. Mathematically, both the plus and the minus sign in (19.40) are valid; however, only the minus sign appears to have physical significance. The minus sign corresponds to the cause producing the effect preceding the effect, which is presumably necessary in a physically meaningful theory. Eq. (19.40) with the minus sign only is known as the retarded potential solution of the inhomogeneous wave equation. The solution with the plus sign is known as the advanced potential and appears to have no physical significance, although at various times attempts have been made to use the advanced potential to explain certain difficulties in electrodynamics.

A retarded potential might be visualized as follows. Consider an observer located at the point p in space as seen in Figure (19.2) and let a sphere whose center is at p contract toward p with a radial velocity c such that it has just contracted onto the point p at the time of observation t. The time at which this information collecting sphere passes the source of the electric field at the

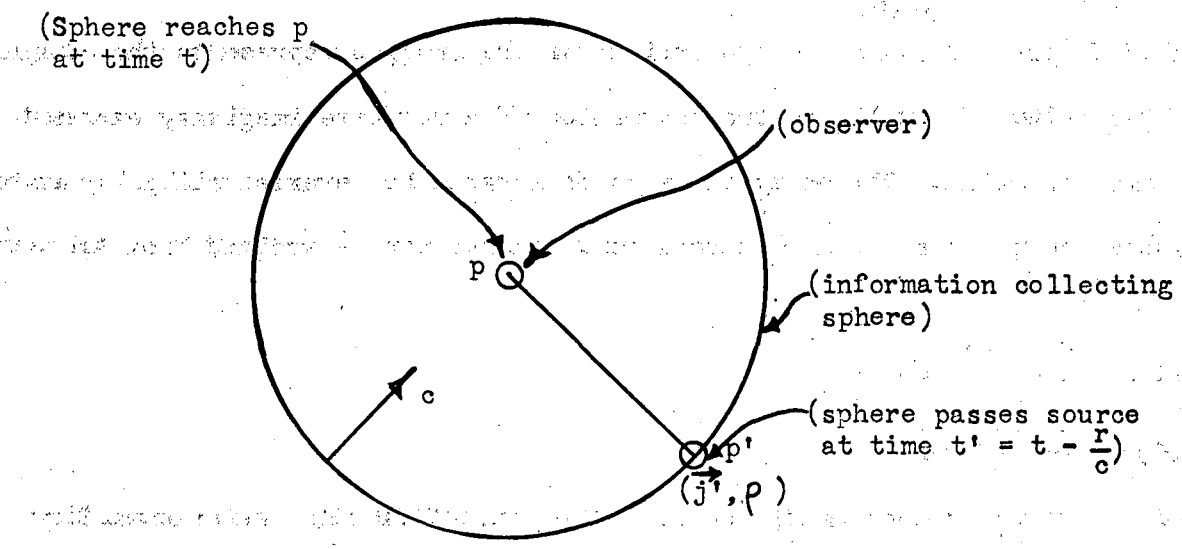


Figure (19.2)

point p' is then the time at which the source produces the effect which is felt at p at the time t. If we denote by the rectangular bracket symbol $[\quad]$ that the variables contained within the bracket are to be evaluated at the retarded time t', in this sense, then from (19.40) the integrals of the inhomogeneous wave equations (18.18) and (18.19) corresponding to a current distribution $\vec{j}(p')$ and a charge distribution $\rho(p')$ are:

$$(19.41)\qquad \vec{A}(p,t) = \frac{\mu_o}{4\pi}\iiint \frac{[\vec{j}(p')]}{r(p,p')}\,dv'$$

$$(19.42)\qquad \phi(p,t) = \frac{1}{4\pi k_o}\iiint \frac{[\rho(p')]}{r(p,p')}\,dv'$$

If the time variation of the scalar and the vector potentials is assumed to be:

$$(19.43)\qquad \vec{A}(p,t) = \vec{A}_o(p)e^{-i\omega t}, \qquad \phi(p,t) = \phi_o(p)e^{-i\omega t}$$

Then a single sinusoidal component of the solutions (19.41) and (19.42) may be written:

$$(19.44)\qquad \vec{A}_{o\omega}(p) = \frac{\mu_o}{4\pi}\iiint \frac{\vec{j}_\omega\, e^{iKr}}{r(p,p')}\,dv'$$

$$(19.45)\qquad \phi_{o\omega}(p) = \frac{1}{4\pi k_o}\iiint \frac{\rho_\omega\, e^{iKr}}{r(p,p')}\,dv'$$

Only the real part of either the potentials or the charges represents the actual physical quantity. We will use the convention of a negative imaginary exponent in the time variation. Obviously the sign of i could be reversed without changing the physical quantities. The electromagnetic fields can be derived from the potentials by:

$$(18.13)\qquad \vec{E} = -\vec{\nabla}\phi - \frac{\partial \vec{A}}{\partial t}$$

$$(11.16)\qquad \vec{B} = \vec{\nabla} \times \vec{A}$$

The gradient of the "monochromatic" scalar potential (19.45) is, after commuting the $\vec{\nabla}$ operator with ρ, changing it to $\vec{\nabla}'$ to operate on the coordinates of the source point p', (omitting ω from j and ρ):

$$(19.46)\qquad -\vec{\nabla}\phi_{o\omega}(p) = -\frac{\vec{\nabla}}{4\pi k_o}\iiint \frac{\rho_\omega e^{iKr(p,p')}}{r(p,p')}\,dv' = -\frac{1}{4\pi k_o}\iiint \rho_\omega \vec{\nabla}\left(\frac{e^{iKr}}{r}\right) dv'$$

$$-\vec{\nabla}\phi_{o\omega}(p) = \frac{1}{4\pi k_o}\iiint \rho_\omega \vec{\nabla}'\left(\frac{e^{iKr}}{r}\right) dv'$$

The electric field therefore becomes; (since $\frac{\partial}{\partial t} \rightarrow -i\omega$)

$$(19.47)\qquad \vec{E}_{o\omega}(p) = \frac{1}{4\pi k_o}\iiint \rho_\omega \vec{\nabla}'\left(\frac{e^{iKr}}{r}\right) dv' + \frac{i\mu_o\omega}{4\pi}\iiint \frac{j_\omega e^{iKr}}{r}\,dv'$$

And the magnetic field in a similar manner can be developed from:

$$(19.48)\qquad \vec{B}_{o\omega} = \frac{\mu_o}{4\pi}\vec{\nabla}\times\iiint \frac{j_\omega e^{iKr}}{r}\,dv' = \frac{\mu_o}{4\pi}\iiint \vec{j}_\omega \times \vec{\nabla}'\left(\frac{e^{iKr}}{r}\right) dv'$$

A change from $\vec{\nabla}$ to $\vec{\nabla}'$ was made after the vector transformation has been made in terms of the $\vec{\nabla}$ operator, noting that $\vec{\nabla}$ does not operate on $\vec{j}$.

We have seen under the modification of retardation, that the Coulomb field integrals do not contain the operator $\vec{\nabla}$ operating on $1/r$ only, as is the case in the non-retarded expression, but operating on e^{iKr}/r instead. When the differentiation indicated by the gradient operator is carried out on e^{iKr}/r, two terms will be secured. One term will vary as $1/r$ and the other term will vary as $1/r^2$. This will be true both in the expression for the electric field and the magnetic field. We have seen that the $1/r$ term might give a finite

contribution in terms of the surface integral of the Poynting vector, and that it therefore represents a net radiation energy loss. The zone in which the 1/r term is dominant is known as the wave zone, or radiation field zone of the system of radiating charges and currents, while the zones of higher order radial dependence are known as the quasi-stationary or inductance field zones.

The expression (19.48) for the magnetic fields, permits us to compute the total energy loss by radiation for an arbitrary current distribution. If only the 1/r dependent radiation field is taken into account, then (19.48) can be written as:

$$(19.49)\qquad \mu_o \vec{H}_{\omega\,rad} = \vec{B}_{\omega\,rad} = -\frac{i\mu_o K}{4\pi}\iiint (\vec{j} \times \vec{r}_1)\,\frac{e^{iKr}}{r}\,dv', \qquad \vec{r}_1 = \frac{\vec{r}}{r}$$

The rate of energy loss from the radiating system can then be calculated explicitly in terms of the current distribution. It is seen by comparing (19.47) to (19.49) that in the radiation field, $\vec{E}$ and $\vec{H}$ are at right angles to each other and that the ratio of the magnitudes is given by:

$$R_o = \sqrt{\frac{\mu_o}{k_o}}$$

The time average of the Poynting vector is therefore given by:

$$(19.50)\qquad \vec{\bar{N}} = \frac{\vec{E} \times \vec{H}}{2} = \frac{1}{2}\sqrt{\frac{\mu_o}{k_o}}\; H^2 \vec{r}_1$$

where the time averaging of the sinusoidal components, which were assumed in (19.43) has produced the factor of two in the denominator. Substituting (19.49) into (19.50), we obtain for the time average of the Poyting vector as a function of the position of the observation point p;

$$(19.51)\qquad \vec{N}_\omega = \frac{R_o K^2}{32\,\pi^2}\left|\iiint (\vec{j} \times \vec{r}_1)\,\frac{e^{iKr}}{r}\,dv'\right|^2 \vec{r}_1$$

In general this formula does permit the calculation of the rate of radiation from a given system of currents. If the point of observation is at a distance r large compared to the dimension R_1 of the radiating system, as is the case in the radiation from atomic charge systems, then the function 1/r in the integral is a slowly varying function compared to the remaining functions, and the rate

of total energy radiation can then be expressed as an integral over a sphere of radius r which reduces to an integral over the solid angle Ω subtended at the point of radiation. The final relation for the rate of total energy radiated by a radiating system is then:

$$(19.52) \qquad U_\omega = \iint \vec{N}\cdot d\vec{S} = \frac{R_0 K^2}{32\pi^2}\int \left| \iiint (\vec{j}_\omega \times \vec{r}_1) e^{iKr} \, dv' \right|^2 d\Omega$$

where $d\Omega = \frac{\vec{r}_1 \cdot d\vec{S}}{r^2}$ ★.

If the radiating system is small in extent, not only compared to the distance from the radiating system to the observer, $r >> R_1$, but also relative to the wavelength of the radiation $\lambda >> R_1$, then the factor e^{iKr} will also be a slowly varying function and therefore it can be expanded in a power series about the center of the radiating system. The various terms in such a power series expansion will have energy losses that depend on successively higher powers of the frequency. We shall investigate this expansion equation (21.16) in more detail later. The various terms of this expansion represent the so-called multipole expansion of the radiation. This discussion has not included the induction field contribution at all, and in fact we have not proved explicitly that the induction field does not contribute to the radiated energy, but have only proved that the induction field cannot contribute to the energy that is radiated over a surface of very large radius, since its higher order of inverse radial dependence causes the surface integrals of the induction field to vanish. The evaluation of the Poynting vector over a given surface which does take into account the induction fields would actually give a zero time average contribution but a non-zero instantaneous value to the energy flow, which implies that the induction fields give rise to energy fluctuations in the radiation field. In particular, if we have a radiating system, whose oscillations are non-sinusoidal, which undergoes a net change in its configuration, then net energy can be transferred into the induction field.

★ Eqs. (19.51) and (19.52) give the radiation rates corresponding to a given Fourier components. It will be shown (see Eq. (23.22)) that the total rate of radiation is given by $4\pi\int_0^\infty U_\omega d\omega$ or: (19.52)' $\frac{dW}{dt} = \frac{R_0 K^2}{8\pi^2}\iint \left| \iiint (\vec{j}_\omega \times \vec{r}) e^{iKr} \, dv' \right|^2 d\Omega d\omega$.

CHAPTER 20 THE WIECHERT-LIÉNARD POTENTIALS

Let us now consider the application of the retarded potential expression (19.42) to the computation of the radiation of an electron. In this application, a certain difficulty immediately arises. In classical electrodynamics the only thing that is known about the electron is the fact that it has a certain total charge and any calculation of its radiation field cannot involve any details of how this charge is distributed geometrically in the electron. On the other hand, it is impossible to assume that this charge has a zero physical extent, since various divergences will result. Also certain features of the radiation field are actually independent of the radius of the electron, provided only that it is small compared to the other dimensions in the radiation field. In our discussion of the electron and its behavior, we shall assume that it has a finite radius, but we shall ascribe physical significance only to those properties of the electron which are independent of the magnitude of the radius.

One immediate difficulty arises in the application of the retarded potential concept to the radiation of a system whose total charge is known. If $[\rho]$ is the retarded charge density within the charge system, which must be substituted into Equation (19.42) in order to secure the correct potential $\phi(p,t)$, then it does not follow that $\iiint [\rho] dv$ represents the correct total charge of the charge system. The reason for this apparent paradox is that the various contributions to the integrand of $\iiint [\rho]\, dv$ are evaluated at different times and that during the time that elapses between the measurement of the charge at the various parts of the system as the information collecting sphere of Figure (20.1) sweeps over the charge, the charges may move and appear more or less dense than they should to give a correct value for the total charge. We can illustrate this system best by a detailed consideration of the information collecting sphere. Consider the sphere of Figure (20.1) converging onto the point of observation p with a

velocity c, and let it gather information as to the charge density within a

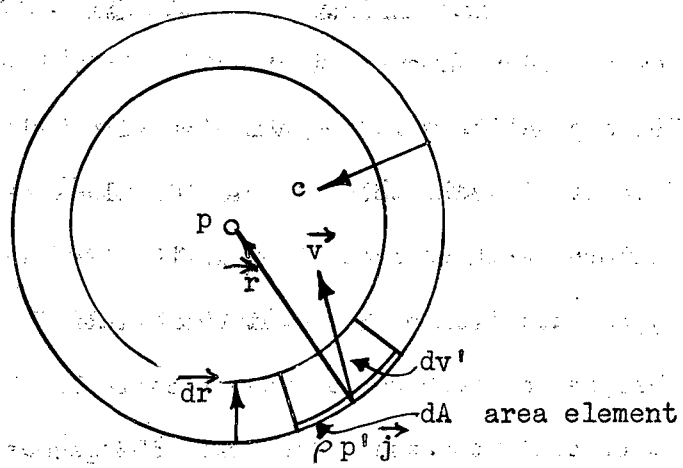


Figure (20.1)

certain charge system as it sweeps across the charge system. The sphere arrives at p at the time t. If the charge system has an average velocity component in the same direction as the motion of the converging sphere, then the volume integral of the retarded charge density will give a result that is in excess of the total charge. If the charge distribution has an average velocity component in opposition to the velocity of the contracting sphere, then the integral will give a result that is less than the total charge of the system.

This situation is analogous to the problem of taking a census of the population of a country. Let us assume that a group of census takers converge upon the "information center" with a certain speed; let us assume that they measure the population density at each point as they travel. The correct population will differ from the total of the census takers' information depending on whether the population had a net migration trend with the census takers (in which the true population is less than the sum of the reported densities) or against the census takers. The retarded potential of an approaching charge will be larger than that of a receding charge at the same distance from the observer since the approaching charge stays longer with the information collecting sphere. Let us consider the radiation field of an electron whose velocity is comparable to c. We shall assume that our electron is a system about which we know:

(1) The total charge of the electron is e.

(2) Within a certain volume V, all parts of the electron's charge are systematically moving with a velocity $\vec{v}$.

Let us now consider the sphere sweeping across this electron in Figure (20.1). Let $\vec{r}$ be the radius vector from the charge to the point of observation to which the sphere is converging. If the charge system is at rest, the amount of charge which this sphere will cross during the time dt as the sphere shrinks in radius dr is given by $[\rho]\, dAdr$. On the other hand, if the charge system is moving with a velocity $\vec{v}$, a quantity of charge which is less than $[\rho]\, dAdr$ by the amount $[\rho]\, dA\, \frac{\vec{v}\cdot\vec{r}}{r}\, dt$ will be crossed by the sphere. The total charge crossed by the sphere in terms of the retarded charge density $[\rho]$ which is observed by the information collecting sphere, is therefore given by:

$$(20.1) \qquad de = [\rho]\, dAdr - [\rho]\, dA\, \frac{\vec{v}\cdot\vec{r}}{r}\, dt$$

But dt and dr are related by:

$$(20.2) \qquad dt = \frac{dr}{c} \quad \text{and}$$

$$(20.3) \qquad dAdr = dv'$$

giving:

$$(20.4) \qquad de = [\rho]\, dv' - [\rho]\, \frac{\vec{v}\cdot\vec{r}}{rc}\, dv'$$

Solving for the retarded charge density, we obtain:

$$(20.5) \qquad [\rho]\, dv' = \frac{de}{1 - \frac{\vec{v}\cdot\vec{r}}{cr}} \,,$$

$$(20.6) \qquad \frac{[\rho]\, dv'}{r} = \frac{de}{r - \frac{\vec{v}\cdot\vec{r}}{c}}$$

Hence an approaching charge appears to have a larger effect, measured in terms of its effective retarded charge:

$$(20.7) \qquad \iiint [\rho]\, dv$$

than the true total charge:

$$(20.8) \qquad \iiint de$$

has. If we substitute this retarded value of the charge density and volume

element expression into the retarded potential expressions, (19.41) and (19.42), we obtain:

$$(20.9) \qquad \phi = \frac{1}{4\pi k_o} \iiint \left[\frac{de}{r - \frac{\vec{r}\cdot\vec{v}}{c}} \right]$$

$$(20.10) \qquad \vec{A} = \frac{\mu_o}{4\pi} \iiint \left[\frac{\vec{v}de}{r - \frac{\vec{r}\cdot\vec{v}}{c}} \right]$$

which are known as the Wiechert-Liénard potential of a single electron. At the limit of a point charge, the distance dependent terms are slowly varying and can be taken outside of the integral sign and, since $\iiint de = e$, the known electronic charge, we obtain for the potentials of a point charge:

$$(20.11) \qquad \phi = \frac{1}{4\pi k_o} \left[\frac{e}{r - \frac{\vec{r}\cdot\vec{v}}{c}} \right]$$

$$\vec{A} = \frac{\mu_o}{4\pi} \left[\frac{e\vec{v}}{r - \frac{\vec{r}\cdot\vec{v}}{c}} \right]$$

which should be valid for all values of the velocity. Note that these expressions are independent of the extent of the electron and are therefore independent of the detailed model used. The fields of the single electron can then be derived from these potentials in the usual way. The details of the field calculation will be taken up later. In general, since the relation of the "retarded" position to the "present" position is not always known, if the charges are accelerated, the Wiechert-Liénard potentials permit an evaluation of the potentials and fields only in terms of the retarded positions and velocities of the charges. If the charge is in uniform motion, it is possible to also express the potentials and the fields in terms of the "present" position of the charge since in this case, computation of the relation between the retarded and present positions is possible.

Consider an electron as in Figure (20.2) that is moving with a uniform velocity $\vec{v}$ in the x direction. The potentials of this electron at the retarded position $\left[\vec{r}\right]$ are given by the Wiechert-Liénard potentials. Let θ be the angle between $\vec{v}$ and $\vec{r}_o$, where $\vec{r}_o$ denotes the present position of the electron. Let

us evaluate the Wiechert-Lienard denominator $s = \left[r - \vec{r}\cdot\vec{v}/c\right]$ in terms of the present position of the electron. The square of s can be written as:

$$(20.13) \quad s^2 = \left(\left[r\right] - \frac{\left[\vec{r}\right]\cdot\vec{v}}{c}\right)^2$$

and since, by the geometry of Figure (20.2), $\vec{r}_o \times \vec{v} = \left[\vec{r}\right] \times \vec{v}$, s^2 can be

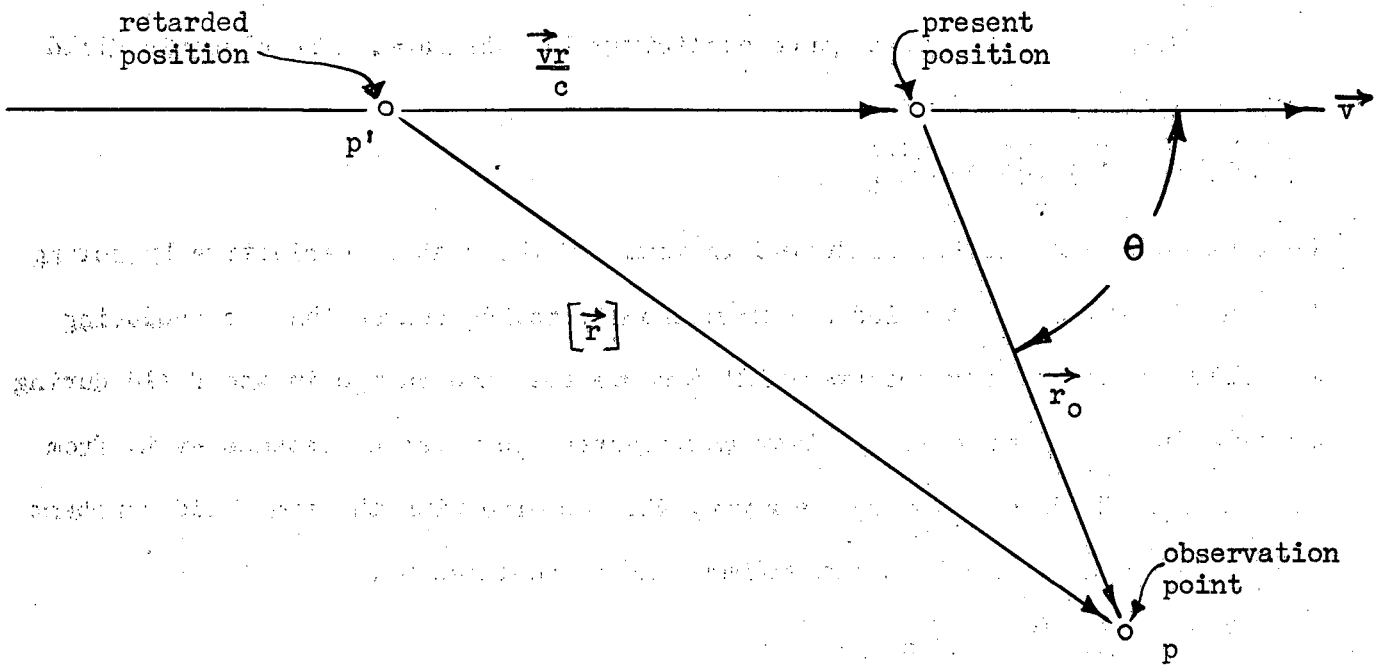


Figure (20.2)

transformed into:

$$(20.14) \quad s^2 = r_o^2 - \left(\frac{\vec{r}_o \times \vec{v}}{c}\right)^2$$

Hence, expressing s explicitly in terms of the present position coordinates, x_o, y_o, z_o, we have:

$$(20.15) \quad s = \sqrt{x_o^2 + y_o^2 + z_o^2 - \frac{v}{c}^2 (y_o^2 + z_o^2)}$$

$$s = \sqrt{x_o^2 + (1 - \beta^2)(y_o^2 + z_o^2)}$$

$$s = r_o\sqrt{1 - \beta^2 \sin^2\theta}$$

where:

$$(20.16) \quad \vec{\beta} = \frac{\vec{v}}{c}$$

The Wiechert-Liénard potentials for the uniformly moving electron in terms of s are given by:

$$(20.17) \qquad \phi = \frac{e}{4\pi k_o s}$$

$$(20.18) \qquad \vec{A} = \frac{\mu_o e \vec{v}}{4\pi s}$$

The fields are easy to compute explicitly in this case. The electric field is given by:

$$(18.13) \qquad \vec{E} = -\vec{\nabla}\phi - \frac{\partial \vec{A}}{\partial t}$$

The time derivative can be evaluated in terms of the spatial derivative by noting that the field must be carried by the uniformly moving charge that is producing the field. A stationary observer will observe the same change in the field during the time dt at a given position that an observer who moves a distance -vdt, from the position of the stationary observer, will observe with the time held constant during the move. Hence time derivatives can be replaced by:

$$(20.19) \qquad \frac{\partial}{\partial t} = - v \frac{\partial}{\partial x}$$

Note that the sign in (20.19) causes the field to be the same at a time dt later as it was a distance vdt behind at the start. Substituting (20.19) into (18.13) the components of the electric field become:

$A_x = A, \; A_y = A_z = 0$

$$(20.20) \qquad E_x = \frac{ex_o}{4\pi k_o s^3} \; (1 - \beta^2)$$

$$(20.21) \qquad E_y = \frac{ey_o}{4\pi k_o s^3} \; (1 - \beta^2)$$

$$(20.22) \qquad E_z = \frac{ez_o}{4\pi k_o s^3} \; (1 - \beta^2)$$

Although $\vec{A}$ has only an x component, the electric field is symmetrical in its three components. Note also that the electric field is directed toward the "present" position of the electron, using a negative electron, of course, and not toward the "retarded" position. Hence vectorially the electric field is

given by:

$$(20.23)\quad \vec{E} = \frac{e}{4\pi k_o s^3}\left[\vec{r} - \frac{\vec{v}r}{c}\right](1-\beta^2) = \frac{e\vec{r}_o(1-\beta^2)}{4\pi k_o s^3} = \frac{e\vec{r}_o}{4\pi k_o r_o^3}\,\frac{(1-\beta^2)}{(1-\beta^2\sin^2\theta)^{3/2}}$$

The magnetic field of the uniformly moving electron is given by:

$$(20.24)\quad \vec{H} = \sqrt{\frac{k_o}{\mu_o}}\;\vec{\beta}\times\vec{E} = \frac{e}{4\pi}\,\frac{\vec{v}\times\vec{r}_o}{r_o^3}\,\frac{(1-\beta^2)}{(1-\beta^2\sin^2\theta)^{3/2}}$$

Note that for low velocities, $\beta \longrightarrow 0$, $s \longrightarrow r$, that (20.23) and (20.24) reduce to the Coulomb (1.3) and Biot-Savart (11.8) fields of a single charge. For high velocities, $\beta \longrightarrow 1$, the magnitude of $\vec{E}$ depends on the angle between the direction of motion of the electron and the radius vector $\vec{r}_o$. As seen in (20.23) the field is increased in a direction at right angles to the direction of motion in the ratio of $1/\sqrt{1-\beta^2}$, while in the direction of motion the field is decreased in the ratio $\left(1-\beta^2\right)$. Therefore, at very high velocities the field resembles more and more the field in a plane wave. For a short time, as a high velocity electron passes an observer, he sees a purely transverse electric and magnetic field. Note that the field of a uniformly moving electron is a non-radiating field in the sense that it does not represent an energy loss. This can be shown by a direct evaluation of the Poynting vector corresponding to the fields given above.

The Wiechert-Liénard potentials (20.9) and (20.10), and the consequences derived from them, were obtained as integrals of the inhomogeneous wave equations (18.18) and (18.19) which in turn were derived from Maxwell's equations (14.7). In making this derivation, it was implied that even if the charged system was moving at a high velocity, no change in the basic equations would result. Whether this is actually so will depend on an examination of Maxwell's equations by relativistic principles. It turns out that provided certain interpretations in the meaning of the mechanical quantities are made, that Maxwell's equations do conform to relativity principles, and that all of the consequences that have been derived therefrom, including the Wiechert-Liénard potentials, remain valid, even at high relativistic velocities.

CHAPTER 21 THE HERTZ $\vec{\pi}$ VECTOR METHOD OF SOLUTION OF THE WAVE EQUATIONS

Since the scalar and the vector potentials are not independent, but in vacuo are connected by the Lorentz condition:

$$(18.14)\quad \vec{\nabla}\cdot\vec{A} + \frac{1}{c^2}\frac{\partial \phi}{\partial t} = 0$$

and, since the charges and the currents cannot be independently specified, but are connected by the equation of continuity:

$$(9.1)\quad \vec{\nabla}\cdot\vec{j} + \frac{\partial \rho}{\partial t} = 0$$

to calculate the field of an arbitrary current and charge distribution for all zones, including the induction field, it is usually advantageous not to derive the field from the scalar and the vector potentials. In fact, it can be easily shown that if the retarded potential forms of the vector potential and scalar potential solution:

$$(19.41)\quad \vec{A}(p,t) = \frac{\mu_o}{4\pi}\iiint \frac{\left[\vec{j}(p')\right]}{r(p,p')}\,dv'$$

$$(19.42)\quad \phi(p,t) = \frac{1}{4\pi k_o}\iiint \frac{\left[\rho(p')\right]}{r(p,p')}\,dv'$$

are assumed, then the Lorentz condition is a direct consequence of the equation of continuity and vice versa.

In order to derive the radiation field it is advantageous to substitute a single function to represent the current distribution and the charge distribution, which is chosen in such a way that the continuity equation is identically satisfied, and to represent the radiation field by a single potential which is chosen in such a way that the Lorentz condition is identically satisfied. The former condition is met by deriving the current density and the charge densities from a single vector function $\vec{p}(p')$, known as the polarization vector, which is a function of the source point p', by the relations:

$$(21.1)\quad \rho_{true} = -\vec{\nabla}\cdot\vec{p}$$

$$(21.2)\quad \vec{j}_{true} = \frac{\partial \vec{p}}{\partial t}$$

It is seen by inspection that the equation of continuity is satisfied by this choice, but that an otherwise arbitrary current distribution can be represented. The polarization vector $\vec{p}$ is related to the true charges and the true currents in the same way that the dielectric polarization $\vec{P}$ is related to the polarization charges and the polarization currents. However, this is only a mathematical parallel and it should be emphasized that $\vec{j}$ and ρ represent the true charges which constitute the external sources of the field, and that therefore $\vec{p}$ is different from the ordinary polarization vector $\vec{P}$. A vector which combines in a similar way the potentials $\vec{A}$ and ϕ, and which at the same time implies the Lorentz condition, is the vector $\vec{\pi}$ which is defined by the equations:

$$(21.3)\qquad \vec{A} = \frac{1}{c^2}\frac{\partial \vec{\pi}}{\partial t}\ , \qquad\qquad \phi = -\vec{\nabla}\cdot\vec{\pi}$$

$\vec{\pi}$ is known as the polarization potential, or Hertzian vector. Since the operation indicated by (21.3) is linear, $\vec{\pi}$ will obey the homogeneous wave equation in source-free space:

$$(21.4)\qquad \Box\vec{\pi} = 0$$

By combining the definitions (21.1), (21.2) and (21.3), we find that the Hertzian vector obeys the same inhomogeneous wave equation, with $\vec{p}$ as the source, that the ordinary potentials obey with the current and charge densities as sources. That is, we have:

$$(21.5)\qquad \Box\vec{\pi} = \nabla^2\vec{\pi} - \frac{1}{c^2}\frac{\partial^2\vec{\pi}}{\partial t^2} = -\frac{\vec{p}}{k_o}$$

The retarded potential solution of (21.5) is therefore given by:

$$(21.6)\qquad \vec{\pi}(p) = \frac{1}{4\pi k_o}\iiint \frac{[\vec{p}(p')]}{r(p,p')}\,dv'$$

for arbitrary time variation and by:

$$(21.7)\qquad \vec{\pi}_\omega(p) = \frac{1}{4\pi k_o}\iiint \frac{\vec{p}(p')e^{iKr(p,p')}}{r(p,p')}\,dv'$$

for sinusoidal time variation. The fields can be derived from the Hertzian vector by the use of the defining equations (21.3). If we let:

$$(21.8)\qquad \vec{Q} = \vec{\nabla}\times\vec{\pi}$$

then the magnetic field is given by:

$$(21.9)\quad \vec{B} = \frac{1}{c^2}\frac{\partial \vec{Q}}{\partial t}$$

and the electric field is given by:

$$(21.10)\quad \vec{E} = \vec{\nabla} \times \vec{Q}$$

The calculation of a radiation field therefore reduces to the calculation of the $\vec{\pi}$ vector from a given polarization vector $\vec{p}$ by the retarded potential integral (21.6) and then the derivation of the fields from equations (21.8), (21.9) and (21.10).

We shall apply this method to the calculation of the fields of a current and charge distribution in a case where we are observing these fields at distances from the sources that are large compared to the extent of the charge source distribution, and also where the extent of the charge distribution is considered to be reasonably small compared to the wave length of the outgoing radiation, as shown in Figure (21.1). To put these assumptions in other words, the

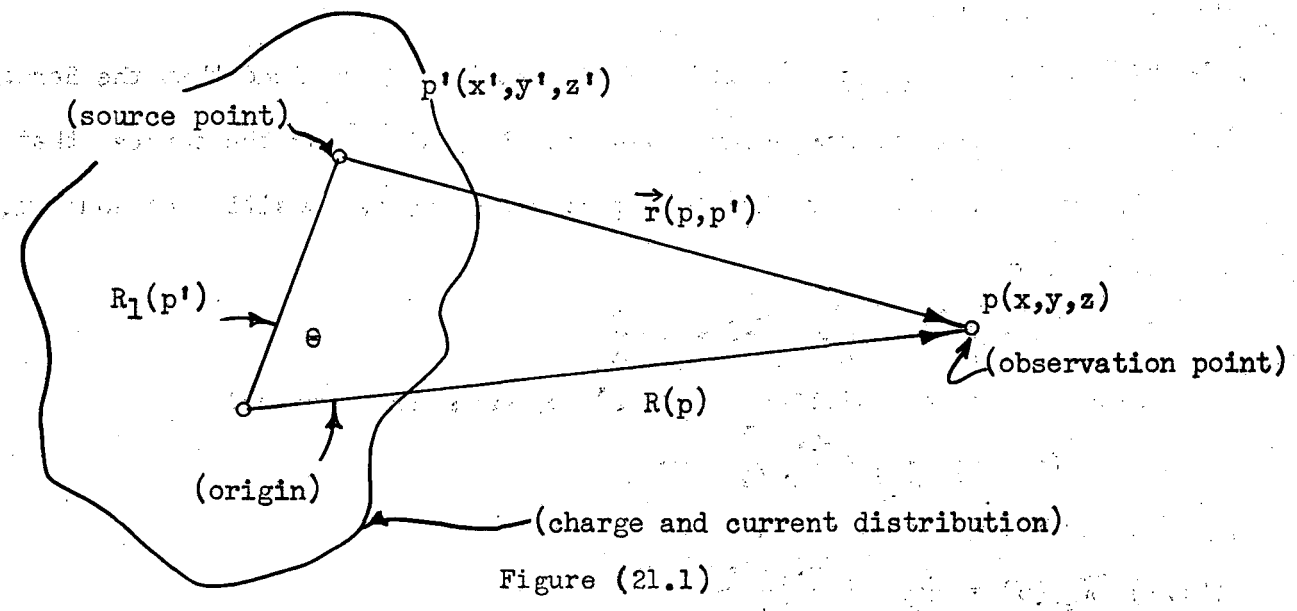


Figure (21.1)

retardation over the current and charge distribution is small compared to the retardation that arises as the signal is propagated from the charges to the observer. For example, if these restrictions are applied to a nuclear system

whose radius is of the order of the electron radius, this would imply that:

$$(21.11)^{*} \quad R_1(p') \ll r(p,p')$$
$$R_1(p') \ll \lambda\!\!\!^{-}$$

When the above approximation holds for a radiating system, then the function $\frac{e^{iKr}}{r}$ is a slowly varying function relative to the variation of $\vec{p}$ itself in the integral (21.7). It is therefore natural to expand this function in a power series about the origin of the charge distribution in terms of the distance $R_1(p')$ from this origin to p' the source point, and on the other hand, to consider the asymptotic behavior of this expansion for large distances R of the observation point p from the center of the charge distribution. The parameters of expansion will in a natrual way be taken as ratios to the wavelength $\lambda\!\!\!^{-} = \frac{1}{K}$. Such an expansion of the function $\frac{e^{iKr}}{r}$ can be derived by considering the expansion relative to a shifted origin:

$$(21.12) \quad \frac{e^{iKr}}{r} = \frac{\pi}{2\sqrt{RR_1}} \sum_{n=o}^{\infty} (2n+1) P_n(\cos\theta) J_{n+1/2}(KR_1) H_{n+1/2}(KR)$$

which expresses the function $\frac{e^{iKr}}{r}$ in terms of a parameter R which represents the distance from the point of observation to the origin, and a parameter R_1 which represents the distance of the source point from the origin as seen in Figure (21.1). In accordance with the approximations of (21.11), we are interested in the values of the function depending on R_1 for small values of R_1 and in the values of the function depending on R for large values of R. For values of $KR_1 \ll 1$, the R_1 dependent function is given by:

---

* Since $R_1 \sim \frac{e^2}{4\pi k_o mc^2}$ for a nuclear system, this restriction is a definite energy limit for the outgoing radiation, since:

$$\frac{e^2}{4\pi k_o mc^2} \ll \lambda\!\!\!^{-} = \frac{c}{\omega} = \frac{\hbar c}{E}$$

and hence:

$$E \ll \frac{4\pi k_o \hbar c}{e^2} \cdot mc^2 = 70 \text{ Mev}$$

where E is the energy of the outgoing radiation.

$$(21.13)\quad J_{n+1/2}(KR_1) \rightarrow \sqrt{\frac{2}{\pi}} \sum_{n=0}^{\infty} \frac{2^n\, n!\, (KR_1)^{n+1/2}}{(2n+1)!}$$

while the asymptotic value of the R dependent function for $RK >> 1$ is given by:

$$(21.14)\quad H_{n+1/2}(KR) \sim \sqrt{\frac{2}{\pi}}\, \frac{e^{iKr}}{\sqrt{KR}}$$

where the asymptotic phase shift terms are being neglected since they do not effect the magnitude of the field at large distances. The expression $\frac{e^{iKr}}{r}$ thus becomes:

$$(21.15)\quad \frac{e^{iKr}}{r} \rightarrow \frac{e^{iKR}}{R} \sum_{n=0}^{\infty} \frac{2^n\, n!}{(2n)!}\, P_n(\cos\theta)(KR_1)^n$$

Using this expansion the $\vec{\pi}_\omega$ vector of Equation (21.7) will be:

$$(21.16)\quad \vec{\pi}_\omega(p) = \frac{e^{iKR}}{4\pi k_o R} \iiint \sum_{n=0}^{\infty} \frac{2^n\, n!}{(2n)!}\, P_n(\cos\theta)\,(KR_1)^n\, \vec{p}_\omega(p')dv'$$

Note that this expansion expresses the radiation field in terms of a sum of moments of various orders of the polarization vector of the charge distribution and that the Legendre term automatically defines the angular distribution of the field corresponding to the various moments. The relative magnitude of the contribution of the various moments will depend on the symmetry properties of the charge distribution.

The zero$^{\underline{th}}$ term, $n = 0$ gives for the $\vec{\pi}_\omega$ vector.

$$(21.17)\quad \vec{\pi}_{o\omega}(p) = \frac{e^{iKR}}{4\pi k_o R} \iiint \vec{p}_\omega(p')dv' = \frac{e^{iKR}}{4\pi k_o R}\, \vec{p}_1$$

where $\vec{p}_1 = \iiint \vec{p}_\omega\, dv'$ is the total dipole moment of the distribution. $\vec{p}_1$ is the same as the dipole moment discussed in Chapter 2. The next higher order terms will only become important in case the dipole term vanishes, that is, if the distribution does not have a net oscillating dipole moment, but has oscillations of higher symmetry. Let us first consider the dipole radiation field in detail. In order to obtain the radiation field from the polarization potential (21.17) where we assume $\vec{p}_1$ to be parallel to the z axis, we must calculate the $\vec{Q}$ vector defined by (21.10). As seen in Figure (21.2) the

components of the $\vec{\pi}_{o\omega}$ vector in spherical polar coordinates are:

$$(21.17a)\quad \pi_R = \pi_{o\omega} \cos\theta = \frac{p_1 \cos\theta\, e^{iKR}}{4\pi k_o R}$$

$$\pi_\theta = -\pi_{o\omega} \sin\theta = -\frac{p_1 \sin\theta\, e^{iKR}}{4\pi k_o R}$$

$$\pi_\phi = 0$$

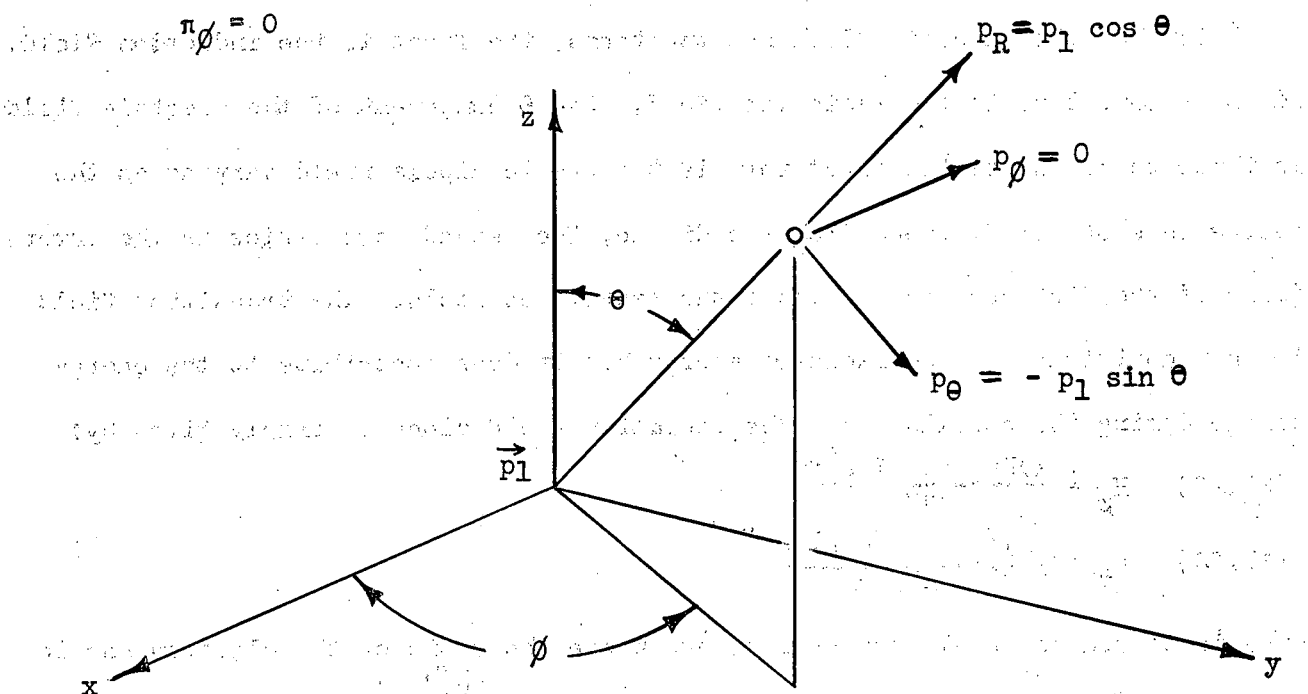


Figure (21.2)

The components of the $Q_{o\omega}$ vector are:

$$(21.18)\quad Q_R = (\vec{\nabla} \times \vec{\pi})_R = \frac{1}{R \sin\theta}\left[\frac{\partial}{\partial\theta}(\sin\theta\pi_\phi) - \frac{\partial \pi_\theta}{\partial\phi}\right]$$

$$Q_\theta = (\vec{\nabla} \times \vec{\pi})_\theta = \frac{1}{R}\left[\frac{1}{\sin\theta}\frac{\partial \pi R}{\partial\phi} - \frac{\partial}{\partial R}(R\pi_\phi)\right]$$

$$Q_\phi = (\vec{\nabla} \times \pi)_\phi = \frac{1}{R}\left[\frac{\partial}{\partial R}(R\pi_\theta) - \frac{\partial \pi R}{\partial\theta}\right]$$

The only nonvanishing component of the $Q_{o\omega}$ vector is therefore the azimuthal component whose magnitude, by (21.17a) is:

$$(21.19)\quad Q_\phi = \frac{p_1 \sin\theta}{4\pi k_o R}\left(\frac{1}{R} - iK\right)e^{iKR}$$

Therefore the component of the magnetic field is:

$$(21.20)\quad H_\phi = -\frac{i\omega}{4\pi} p_1 \sin\theta \left(\frac{1}{R^2} - \frac{iK}{R}\right)e^{iKR}$$

and the components of the electric field from (21.10) are:

$$(21.21)\quad E_R = \frac{1}{R \sin\theta}\frac{\partial}{\partial\theta}(\sin\theta\, Q_\phi) = \frac{p_1 \cos\theta}{2\pi k_o R2}\left(\frac{1}{R} - iK\right) e^{iKR}$$

$$E_\theta = -\frac{1}{R}\frac{\partial}{\partial R}(R\, Q_\phi) = \frac{p_1 \sin\theta}{4\pi k_o R}\left(\frac{1}{R^2} - \frac{iK}{R} - K^2\right) e^{iKR}$$

Note that the magnetic field has two terms, the first is the induction field, and the second term is the radiation field. The θ component of the electric field has three terms in it; the first term is the static dipole field varying as the inverse cube of the distance from the dipole; the second term varies as the inverse square of the distance and is called the transition field. The transition field will not contribute to the radiated energy but it does contribute to the energy storage during the oscillation. The radiation field alone is simply given by:

$$(21.22)\quad H_\phi = \frac{\omega K p_1 \sin\theta\, e^{iKR}}{4\pi R}$$

$$(21.23)\quad E_\theta = -\frac{K^2 p_1 \sin\theta\, e^{iKR}}{4\pi k_o R}$$

which is of course purely transverse, and where the ratio of the electromagnetic field vectors is the same as that in a plane wave, $\sqrt{\frac{\mu_o}{k_o}}$. The radiation field only can be put in the vector form.

$$(21.24)\quad \vec{E} = \frac{K^2\, e^{iKR}}{4\pi k_o R^3}\left[(\vec{p_1} \times \vec{R}) \times \vec{R}\right]$$

The $\vec{Q}$ vector can be interpreted simply in terms of the equations of the lines of force of the radiation. The equations of the lines of force are defined by the differential equation:

$$(21.25)\quad \frac{\partial}{\partial R}(RQ \sin\theta)\, dR + \frac{\partial}{\partial\theta}(RQ \sin\theta)\, d\theta = 0$$

The solution of (21.25) valid in all cases when the current flows along the z axis, is:

$$(21.26)\quad RQ \sin\theta = \text{constant}$$

Upon substituting the expression from Equation (21.19), we obtain:

$$(21.27)\quad R^{-1} \sin^2\theta \cos(KR - \omega t + \tan^{-1} KR) = \text{constant}$$

This is an exact equation applying both in the radiation and in the induction field. When we consider the radiation field alone and include the time variation that we have omitted so far, we have:

$$\sin^2\theta \cos(KR - \omega t) = \text{constant} \tag{21.28}$$

which is the desired equation of the lines of force. The lines $\theta = 0$ are radial lines. The field is periodical radially with the radial spacing corresponding to the wavelength λ*. The lines of force are sketched in Figure (21.3).

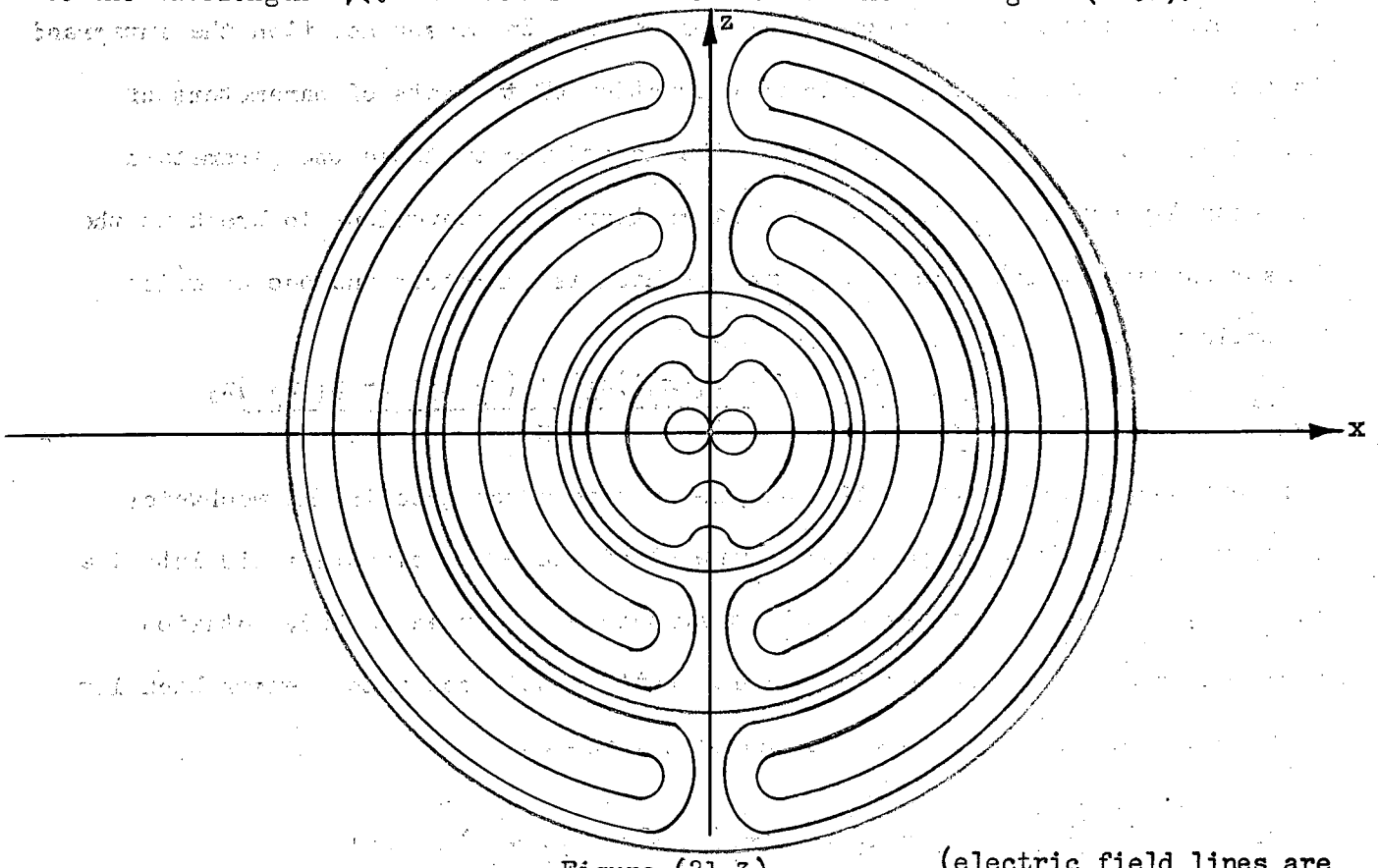


Figure (21.3) (electric field lines are shown)

Near the source a complete expression for $\vec{Q}$ must be used which indicates how the equations of the lines of force grow from the static case into the field of the wave zone pattern. Lines of force in successive states of the radiation are shown in the figure.

Let us now consider the significance of the higher moments that contribute to the $\vec{\pi}$ vector. Consider the term for n = 1, in the general expansion (21.16):

*True for large values of KR only. For small values of KR the apparent phase velocity of the wave front is $> c$.

$$(21.29)\quad \vec{\pi}_{1\omega} = \frac{e^{iKR}}{4\pi k_o R} \iiint K R_1 \vec{p} \cos\theta \, dv'$$

which can be written as:

$$(21.30)\quad \vec{\pi}_{1\omega} = \frac{e^{iKR}}{4\pi k_o R^2} K \iiint \vec{p}(\vec{R}_1 \cdot \vec{R}) dv'$$

Let us write the integrand in tensor notation. Let the components of the vector $\vec{R}_1$ be x_i' and let the components of the radius vector $\vec{R}$ be x_i and let the components of the polarization vector be p_i. In tensor notation the integrand is then given by $p_j(x_i' x_i)$. This is a function of two sets of parameters of the distribution, namely p_j and x_i'. The dependence on these two parameters is generally not a symmetric one and it is therefore convenient to break up the tensor expression into two terms, one of which is symmetric and one of which is antisymmetric. That is, let us put:

$$(21.31)\quad p_j(x_i' x_i) = \frac{(p_j x_i' + p_i x_j')x_i + (p_j x_i' - p_i x_j')x_i}{2}$$

This process is analogous to the process that is often used in the mechanics of continua, namely of separating a general strain of an elastic solid into the sum of a pure strain represented by a symmetric tensor and a body rotation represented by an antisymmetric tensor. Putting this separated tensor back into the vector notation, the integral can be written as:

$$(21.32)\quad \frac{1}{2} \iiint \left\{ \left[\vec{R} \times (\vec{p} \times \vec{R}_1) \right]_j + x_i(p_j x_i' + p_i x_j') \right\} dv'$$

Let us consider the significance of the antisymmetric term first. Since $\vec{R}$ is not a function of the primed variables of integration it can be taken outside of the integral sign, making the integral of the first term:

$$(21.33)\quad \frac{1}{2} \vec{R} \times \iiint (\vec{p} \times \vec{R}_1) dv'$$

The significance of the integrand can easily be recognized by replacing the polarization vector by the current. Since $\vec{p} = -\vec{j}/i\omega$ for sinusoidal time

variation, the integral becomes:

$$(21.34) \quad -\frac{\vec{R}\,x}{2i\omega}\iiint\left(\vec{j}\times\vec{R}_1\right)dv'$$

The integral is exactly equal to the magnetic moment $\vec{m}=\frac{1}{2}\iiint\vec{R}_1\times\vec{j}\,dv'$ of the current distribution and hence the integral becomes finally:

$$(21.35) \quad \frac{1}{i\omega}\left(\vec{R}\times\vec{m}\right)$$

and the corresponding $\vec{\Pi}$ vector for the antisymmetric part only, becomes:

$$(21.36) \quad \vec{\Pi}_{\substack{\text{anti-}\\\text{sym.}}}=\frac{e^{iKR}(\vec{R}\times\vec{m})}{4\pi k_o R^2 i\omega}\,K$$

The field can be computed directly. If we omit the induction terms, we obtain, by calculating the $\vec{Q}$ vector and differentiating with respect to time:

$$(21.37) \quad \vec{B}_{\substack{\text{anti-}\\\text{sym.}}}=\frac{1}{c^2}\frac{\partial\vec{Q}}{\partial t}=\frac{i\mu_o K^2\left[\vec{R}\times(\vec{m}\times\vec{R})\right]e^{iKR}}{4\pi R^3}$$

which can be reduced to:

$$(21.38) \quad \vec{H}_{\substack{\text{anti-}\\\text{sym.}}}=\frac{K^2 e^{iKR}}{4\pi R^3}\left[\vec{R}\times(\vec{m}\times\vec{R})\right] \quad \text{(ignoring phase)}$$

Note that this is exactly the same mathematical form we had in (21.24) for the electric field in the case of the electric dipole radiation. The radiation field described in (21.38) is known as a magnetic dipole radiation field and it is exactly the same as that of the electric dipole radiation except for the fact that the role of the electric and the magnetic vectors are interchanged. Magnetic dipole radiation corresponds to a current distribution which has no net electric oscillating dipole moment but which has a sinusoidally varying circulation of the charges.

Now let us consider the physical significance of the symmetrized term that was dropped from the discussion at Equation (21.32):

$$(21.39) \quad \Pi_{j\,\text{sym}}=\frac{Ke^{iKR}\,x_i}{4\pi k_o R^2}\iiint\frac{1}{2}\,(p_j\,x_i'+p_i\,x_j')dv'$$

The symmetry of (21.39) will become clearer if we write it, instead of in terms of the components of the polarization vector $\vec{p}$, in terms of the charge density $\rho(p')$ and the coordinates of the charge density. Consider a quantity defined by:

$$(21.40)\quad Q_{ij} = \iiint \rho\, x_i'\, x_j'\, dv' = -\iiint \frac{\partial p_k}{\partial x_k'}\, x_i'\, x_j'\, dv'$$

which is known as the electric quadrupole moment of the charge distribution. If we express this in terms of the polarization vector and integrate by parts, and omit the integrated out parts by the expedient of having the volume of integration extend outside of the charge distribution, then in terms of the quadrupole moment of the charge distribution, the $\vec{\Pi}$ vector of the radiation becomes:

$$(21.41)\quad \Pi_{j\,\mathrm{sym}} = \frac{Ke^{iKR}}{8\pi k_o R^2}\, x_i\, Q_{ij}$$

Note that this procedure is identical to the procedure which we could have used in the dipole case, namely the quantity p_1 is given by:

$$(21.42)\quad p_{1j} = \iiint \rho\, x_j'\, dv' = \iiint (-\vec{\nabla}\cdot\vec{p}) x_j\, dv' = \iiint -\frac{\partial p_i}{\partial x_i'}\, x_j'\, dv' = \iiint p_j\, dv'$$

which is the expression used in (20.17). Since the quadrupole moment is represented by a symmetric matrix:

$$(21.43)\quad Q_{ij} = \iiint \rho(p')\, x_i'\, x_j'\, dv'$$

it can be represented by a family of quadrics derived from the quadratic form:

$$(21.44)\quad x_i\, x_j\, Q_{ij} = C = \text{constant}$$

In terms of the parameter C of this equation, the $\vec{\Pi}$ vector can therefore be written as:

$$(21.45)\quad \vec{\Pi} = \frac{Ke^{iKR}}{8\pi k_o R^2}\, \vec{\nabla}\, C$$

indicating that the direction of the $\vec{\Pi}$ vector is everywhere normal to the quadric surfaces defined by the quadratic form of the quadrupole matrix. Let us calculate the components of the fields that correspond to a general quadrupole. Let us choose a system of axes x, y, z, that correspond to the principal axes of the quadrupole quadric. The components of the $\vec{\Pi}$ vector are therefore given by:

$$(21.46)\quad \pi_x = \frac{e^{iKR}}{8\pi k_o R} \sin\theta \cos\phi\, Q_{xx}$$

$$\pi_y = \frac{e^{iKR}}{8\pi k_o R} \sin\theta \sin\phi\, Q_{yy}$$

$$\pi_z = \frac{e^{iKR}}{8\pi k_o R} \cos\theta\, Q_{zz}$$

If we take the curl of these expressions, we will obtain the components of the $\vec{Q}$ vector:

$$(21.47)\quad Q_\theta = \frac{iK^2 e^{iKR}}{16\pi k_o R} \sin\theta \sin 2\phi\, (Q_{xx} - Q_{yy})$$

$$Q_\phi = \frac{iK^2 e^{iKR}}{32\pi k_o R} \sin 2\theta \left[(Q_{xx} + Q_{yy} - 2Q_{zz}) - (Q_{yy} - Q_{xx}) \cos 2\phi \right]$$

And from $\vec{Q}$ the electric field components are:

$$(21.48)\quad E_\theta = \frac{K^3 e^{iKR}}{32\pi k_o R} \sin 2\theta \left[(Q_{xx} + Q_{yy} - 2Q_{zz}) - (Q_{yy} - Q_{xx}) \cos 2\phi \right]$$

$$E_\phi = \frac{K^3 e^{iKR}}{16\pi k_o R} \sin\theta \sin 2\phi \left[Q_{yy} - Q_{xx} \right]$$

while the magnetic field components are:

$$(21.49)\quad B_\theta = -\frac{\omega\mu_o K^2 e^{iKR}}{16\pi R} \sin\theta \sin 2\phi \left[Q_{yy} - Q_{xx} \right]$$

$$B_\phi = \frac{\omega\mu_o K^2 e^{iKR}}{32\pi R} \sin 2\theta \left[(Q_{xx} + Q_{yy} - 2Q_{zz}) - (Q_{yy} - Q_{xx}) \cos 2\phi \right]$$

Note that the fields depend only on the differences in the quadrupole moments. There are no radial components, and only the radiation fields are given.

Two features of the quadrupole radiation fields can be noted by inspection. First, in case two of the moments are equal, that is, if the quadrupole is a spheroidal distribution, then if the polar axis of the spheroid is taken to be the z axis, then the only non-vanishing component of the $\vec{Q}$ vector and hence of the magnetic field, will be the azimuthal component given by:

$$(21.50)\quad Q_\phi = \frac{iK^2 e^{iKR}}{16\pi k_o R} \sin 2\theta\, (Q_{xx} - Q_{zz})$$

Second, in general, the distribution will have two nodal cones where there is zero field, compared to a single nodal line in the case of a dipole distribution. Physically the quadrupole distribution arises from a pulsating charge distribution of such symmetry that the dipole moment remains zero during the pulsation. The simplest example of a quadrupole is two dipoles oscillating in opposition, but displaced a slight distance from each other.

CHAPTER 22 THE CONVECTIVE POTENTIAL

In Chapter 20, we derived the Wiechert-Liénard potentials:

$$(20.11)\qquad \phi = \frac{1}{4\pi k_o}\left[\frac{e}{r - \frac{\vec{v}\cdot\vec{r}}{c}}\right] = \frac{1}{4\pi k_o}\left[\frac{e}{r - \vec{\beta}\cdot\vec{r}}\right]$$

$$(20.12)\qquad \vec{A} = \frac{1}{4\pi k_o}\,\frac{e\vec{v}}{c^2\left[r - \frac{\vec{v}\cdot\vec{r}}{c}\right]} = \frac{1}{4\pi k_o c}\,\frac{\vec{\beta}e}{\left[r - \vec{\beta}\cdot\vec{r}\right]}$$

which give the potentials due to a point charge moving with an arbitrary velocity $\vec{v} = c\vec{\beta}$. From these potentials we obtained the electromagnetic fields of a uniformly moving charge and in particular studied the angular distribution of the fields for charges moving with high velocity.

Considerable care must be exercised in applying these equations in such cases. The Wiechert-Liénard potentials were derived from the retarded potential solutions of the wave equation, (19.41) and (19.42) and the wave equation in turn was derived from Maxwell's equations (14.7). The wave equations (18.18) and (18.19) represent, as they stand, correct equations in one particular frame of reference, which classically was called the ether frame. Or to put it in other words, the velocity of propagation of electromagnetic radiation would only be c in that particular frame in which the wave equation is valid. As can be seen by substitution of the Galilean transformation relations:

$$(22.1)\qquad x' = x - vt$$
$$t' = t$$

into the wave equation, the wave equation in the x't' frame is not the same as in the xt frame and no longer represents propagation with the velocity c. If we take these conclusions at face value, we would therefore conclude that the velocity $\vec{v}$ which appears in the Wiechert-Liénard potentials represents the velocity of the electron relative to an observer at rest, where by rest we mean at rest relative to the classical ether frame in which the wave equation is valid. With this

definition of $\vec{v}$, no restriction that $\vec{\beta}$ be small is necessary. Actually, this restriction will be removed later by the Theory of Relativity in which the Lorentz transformation is substituted for the Galilean transformation and in which the wave equation remains valid for all frames. We shall show if the treatment is carried out relativistically that the Wiechert-Liénard potentials will remain intact and that all conclusions derived therefrom will remain valid provided that the relative velocity between an electron and an observer is used, rather than the velocity of the electron relative to a stationary ether. Even though the conclusions which we shall draw from the Wiechert-Liénard potentials and the wave equations appear to be limited, in fact so limited that their extensive use appears to be unjustified, we shall derive our conclusions with the understanding that we shall later show relativistically that all conclusions obtained here will remain valid provided the change in the interpretation of the velocity is made as indicated.

It is instructive to see how our conclusions regarding the field of a uniformly moving charge can be derived directly from the inhomogeneous wave equations. The inhomogeneous wave equations (18.18) and (18.19) are subject to the subsidiary condition (18.15). If we consider the field of an electron which is moving with a uniform velocity, then the time derivatives and spatial derivatives are no longer independent because of the fact that the field must be carried convectively with the electron. This fact can be expressed mathematically by:

$$\frac{\partial}{\partial t} = -\vec{v} \cdot \vec{\nabla} \tag{22.3}$$

This expression indicates that any field parameter at a given point changes by the same amount in a time dt that it would change if at a fixed time one would compare it to the same field parameter evaluated at a distance -vdt displaced in the direction of the motion of the electron. Let us write one component of the inhomogeneous wave equation (19.22) as:

$$(22.4) \quad \Box \Psi (p) = - g(p)$$

and let us take the velocity parallel to the x direction. Using (22.3) the inhomogeneous wave equation therefore becomes:

$$(22.5) \quad (1 - \beta^2) \frac{\partial^2 \Psi}{\partial x^2} + \frac{\partial^2 \Psi}{\partial y^2} + \frac{\partial^2 \Psi}{\partial z^2} = -g(x,y,z)$$

which by the substitution:

$$(22.6) \quad x_1 = \frac{x}{\sqrt{1 - \beta^2}} ; \; y_1 = y \; ; \; z_1 = z$$

can be reduced to a simple electrostatic Poisson equation:

$$(22.7) \quad \nabla_1^2 \Psi = - g \left(\sqrt{1 - \beta^2}\, x_1, y_1, z_1\right)$$

the solution of which is the ordinary Coulomb potential:

$$(22.8) \quad \Psi (p_1) = \frac{1}{4\pi} \int_{v'} \frac{g(p'_1)dv'_1}{r(p_1, p_1')}$$

Transforming back, this becomes:

$$(22.9) \quad \Psi(p) = \frac{1}{4\pi} \int_{v} \frac{g(p')dv'}{s}$$

where:

$$(22.10) \quad s = \sqrt{(x - x')^2 + (1 - \beta^2) \left[(y - y')^2 + (z - z')^2 \right]}$$

Note that this is the same potential which we obtained from the Wiechert-Liénard potentials. However, in this case the question regarding the propagation velocity of the corresponding wave and the relation between present and retarded potentials does not enter, since by a suitable transformation we have succeeded in transforming the equation to be solved (22.4) to a static equation (22.7). As we shall see later the purely mathematical process (22.6) is in reality a Lorentz transformation in which we transform the observer's position to a frame at rest relative to the electron whose field is to be computed. Note also that these fields are non-radiating fields. This can be shown directly by evaluating the Poynting vector over a spherical surface enclosing the charge. It can also be

shown by noting that the electron obviously does not radiate in the primed frame in which it is at rest and therefore presumably will not radiate from the stationary observer's point of view in the unprimed frame.

The scalar and vector potentials of the charge can therefore be written in the symetric forms:

$$(22.11) \qquad \phi = \frac{e}{4\pi k_o s}$$

$$(22.12) \qquad \vec{A} = \frac{e\vec{\beta}}{4\pi k_o c s}$$

from which the fields can be computed by the method described earlier. It is instructive to calculate the force which would be exerted by these fields on another electron which is moving with a velocity $\vec{v}$ parallel to the original electron producing the field. This force is given by the Lorentz force expression (16.15):

$$(22.13) \qquad \vec{F} = e(\vec{E} + \vec{v} \times \vec{B})$$

which, using the potentials (22.11) and (22.12) and Equations (11.16) and (18.13) becomes:

$$(22.14) \qquad \vec{F} = e(\vec{E} + \vec{v} \times \vec{B}) = \frac{e^2}{4\pi k_o}\left[-\vec{\nabla}\left(\frac{1}{s}\right) + (\vec{\beta} \cdot \vec{\nabla})\frac{\vec{\beta}}{s} + \vec{\beta} \times (\vec{\nabla} \times \frac{\vec{\beta}}{s})\right]$$

which by expansion of the vector product becomes:

$$(22.15) \qquad \vec{F} = -\frac{e^2}{4\pi k_o}\vec{\nabla}\left[\frac{1 - \beta^2}{s}\right]$$

which can be written in the form:

$$(22.16) \qquad \vec{F} = -\vec{\nabla}\Psi$$

where:

$$(22.17) \qquad \Psi = \frac{e^2(1 - \beta^2)}{4\pi k_o s}$$

is called the convection potential. The force of one electron on the other is therefore derivable from a scalar potential, Ψ, but this scalar potential does

not have spherical symmetry about the position of the field-originating electron. In particular, since the direction of the force must be perpendicular to the surface of equal convective potential, we would conclude that the force $\vec{F}_2$ exerted by the electron e_1 at $(x_1 y_1 z_1)$ on the electron e_2 at $(x_2 y_2 z_2)$ is perpendicular to the ellipsoid:

$$(22.18) \qquad s = \sqrt{(x_1 - x_2)^2 + (1 - \beta^2)\left[(y_1 - y_2)^2 + (z_1 - z_2)^2\right]} = \text{constant}$$

shown in Figure 22.1.

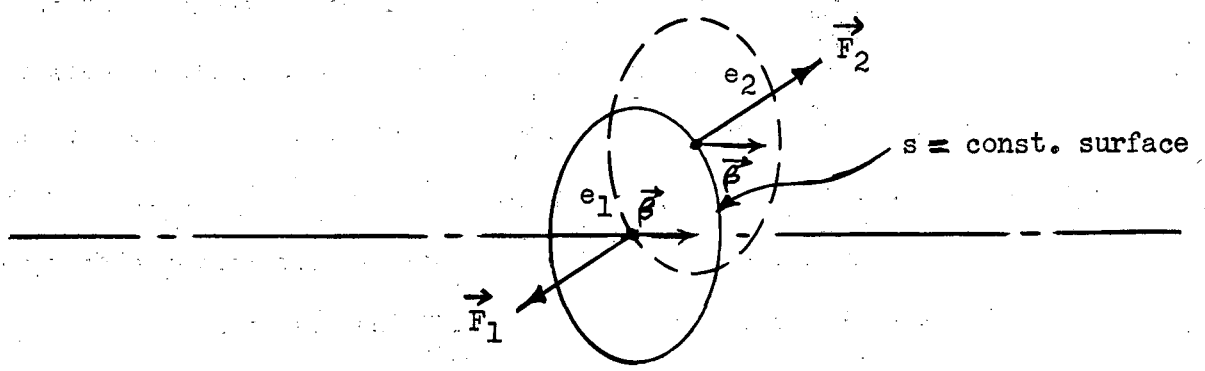


Figure (22.1)

On the other hand, the reaction force $\vec{F}_1$ on the electron e_1 is perpendicular to the corresponding ellipsoid (shown dashed in, Figure 22.1) referred to the co-moving electron e_2. Hence, except when the line between the electrons is parallel or perpendicular to the direction of motion, the forces of action and reaction do not appear to be co-linear and therefore if the two electrons were connected by a rigid bar, there would be a couple acting about the axis perpendicular to the line between the electrons and the direction of motion. Such a couple should be experimentally detectable. Trouton and Noble attempted to measure the torque on a suspended charged condenser at various parts of the year, when presumably the velocity of the condenser relative to the "ether" would differ by an amount of the general order of 30 kilometers per second, which is the magnitude of the velocity of the earth in its orbit about the sun. Actually, no such effect was observed. This null effect can only be explained correctly

by the theory of relativity. We can see immediately that if we interpret the velocity as being simply the relative velocity of the observer and the co-moving charges, then since the observer is at rest relative to the suspended condenser, no effect would be expected. On the other hand, if we would observe the condenser from a frame which is also in motion relative to the suspended condenser, then we should observe such a couple. However, the question arises as to how such a couple would be observed, and the answer is by either observing an angular acceleration or by balancing the couple with another known couple such as a torque in an elastic suspension. We shall find that not only the magnitude of the electric torque is changed, but also the magnitude of all mechanical torques is changed by exactly the same law (Eq. (22.6)) as a function of the velocity of the observer. Therefore the state of equilibrium remains invariant independent of the observer's state. Hence relativistically we would obtain a zero torque either in a frame moving with the electrons or not moving with the electrons.

CHAPTER 23

RADIATION FROM AN ACCELERATED CHARGE

Let us calculate the complete electric and magnetic fields from an electric charge e which is assumed to be moving in accordance with the condition:

$$x_i'(t') = \text{given} \tag{23.1}$$

where x_i' are the coordinates of the charge and t' is the time of emission of the signal emitted at x_i' which will arrive at x_i at time t. Note that this statement assumes the existence of a universal time scale which can be established by signals travelling with infinite velocity -- strictly a non-relativistic concept. Nevertheless, this procedure, owing to the fact that we are treating the process only in one frame in which the wave equation is valid, leads to relativistically correct results as will be shown later.

The velocity and acceleration of the charge:

$$\vec{v} = \frac{\partial \vec{x}'}{\partial t'} \qquad \dot{\vec{v}} = \frac{\partial^2 \vec{x}'}{\partial t'^2} \tag{23.2}$$

are thus considered as given, by virtue of the influence of external fields or possibly by the influence of the field of part of the charge on the other part. We shall discuss this point later.

The Wiechert-Liénard potentials thus represent the following functional relationships:

$$\phi(x_i,t) = \frac{e}{4\pi k_o} \frac{1}{s\left\{\left[x_i - x_i'(t')\right], \beta(t')\right\}} = \frac{e}{4\pi k_o} \frac{1}{s} \tag{23.3}$$

$$\vec{A}(x_i,t) = \frac{e}{4\pi k_o c} \frac{\vec{\beta}(t')}{s\left\{\left[x_i - x_i'(t')\right], \beta(t')\right\}} = \frac{e}{4\pi k_o} \frac{\vec{\beta}}{sc} \tag{23.4}$$

where (x_i,t) refers to the field point and (x_i',t') to the source point. The field point $p(x_i)$ and source point $p'(x_i')$ are connected by the condition:

$$r(p,p') = \left[\sum_{i=1}^{3} (x_i - x_i')^2\right]^{1/2} = c(t - t') \tag{23.5}$$

and s is given by:

$$s = r - \vec{\beta} \cdot \vec{r} \tag{23.6}$$

When we derive the fields $\vec{E}$ and $\vec{B}$ from the potentials (23.3) and (23.4) by the regulations:

(11.16) $\quad \vec{B} = \vec{\nabla} \times \vec{A}$

(18.13) $\quad \vec{E} = -\vec{\nabla}\phi - \dfrac{\partial \vec{A}}{\partial t}$

we must notice that the operator $\vec{\nabla}$ is a vector operator whose components $\dfrac{\partial}{\partial x_i}$ are partial derivatives at constant time t, and therefore <u>not</u> at constant time t'. Partial differentiation with respect to x compares the potentials at neighboring points at the same time. These potential signals originated from the charge at different times. Similarly $\dfrac{\partial}{\partial t}$ means $\left.\dfrac{\partial}{\partial t}\right|_{x_i=\text{constant}}$ and hence refers to the comparison of potentials at a given point over an interval of time during which the coordinate of the signal originating charge will have changed. Since only the time variations with respect to t' are given (in the original description of the problem), we must transform $\left.\dfrac{\partial}{\partial t}\right|_{x_i}$ and $\left.\vec{\nabla}\right|_t$ to expressions in terms of $\left.\dfrac{\partial}{\partial t'}\right|_{x^i}$.

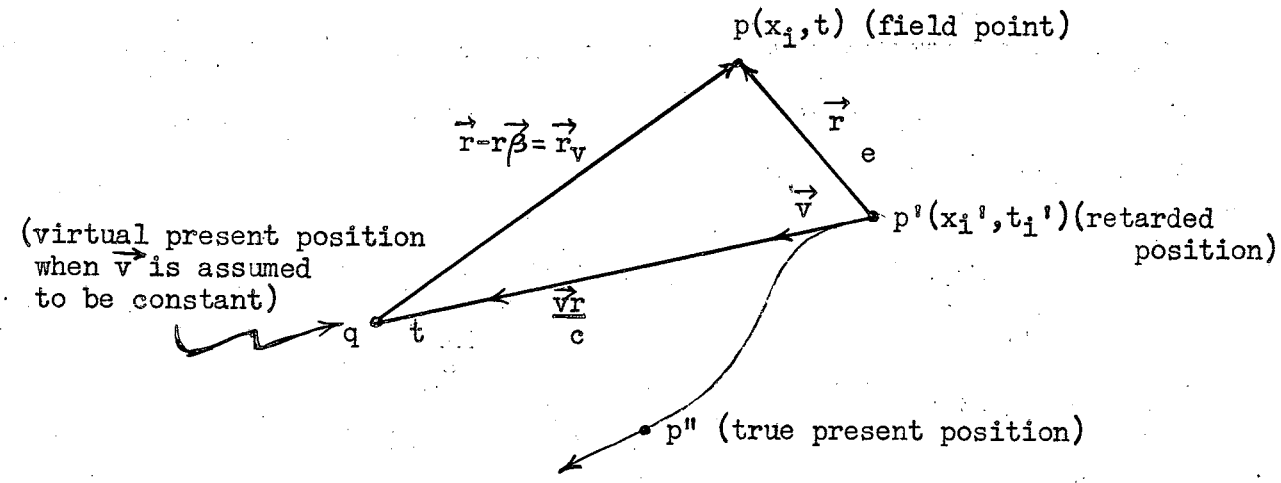


Figure (23.1)

Note that it is in general impossible to express the fields in terms of the "present position" of the accelerating electron as it was in Chapter 20 in the case of uniform motion, but that the distance $\vec{r}$ refers to the retarded position p' of the charge, as seen in Figure (23.1). Note that:

$$(23.7)\qquad \left.\frac{dr}{dt'}\right|_{x_i=\text{constant}} = -\frac{\vec{r}\cdot\vec{v}}{r}$$

and, vectorially:

$$(23.8)\qquad \frac{d\vec{r}}{dt'} = -\vec{v};\quad \frac{d^2\vec{r}}{dt'^2} = -\dot{\vec{v}}$$

To relate $\frac{\partial}{\partial t}$ to $\frac{\partial}{\partial t'}$, we note* that explicitly $\vec{r}$ is given by the functional relation:

$$(23.9)\qquad r = f\left[x_i,\ x_i'(t')\right] = g\left[x_i,\ t'\right]$$

and, functionally (23.5) can be solved for t' in terms of x_i and t, since x_i' is a given function of t'; hence:

$$(23.10)\qquad r = g\left[x_i,\ t'\right] = g\left[x_i,\ h(x_i,t)\right]$$

where $t' = h(x_i,t)$. And thus:

$$(23.11)^x\qquad \frac{\partial r}{\partial t} = c\left[1 - \frac{\partial t'}{\partial t}\right] = \frac{\partial r}{\partial t'}\frac{\partial t'}{\partial t} = -\frac{\vec{r}\cdot\vec{v}}{r}\frac{\partial t'}{\partial t}$$

or, using (23.5), (23.7) and (23.8):

$$(23.12)\qquad \frac{\partial t'}{\partial t} = \frac{1}{1 - \frac{\vec{r}\cdot\vec{v}}{rc}} = \frac{r}{s}$$

Similarly:

$$(23.13)\qquad \vec{\nabla} r = -c\vec{\nabla} t' = \vec{\nabla}'(r) + \frac{\partial r}{\partial t'}\vec{\nabla} t' = \frac{\vec{r}}{r} - \frac{\vec{r}\cdot\vec{v}}{r}\vec{\nabla} t'$$

where by $\vec{\nabla}'$ we mean differentiation with respect to the first argument of the function g in Equation (23.10) only; that is, differentiation at constant retarded time t'. Therefore:

$$(23.14)\qquad \vec{\nabla} t' = -\frac{\vec{r}}{sc}$$

and we have:

$$(23.15)\qquad \begin{cases} \dfrac{\partial}{\partial t} = \dfrac{r}{s}\dfrac{\partial}{\partial t'} \\ \vec{\nabla} = \vec{\nabla}' - \dfrac{\vec{r}}{sc}\dfrac{\partial}{\partial t'} \end{cases}$$

for the required transformation of the differential operator from the field point

---

x Note that $\frac{\partial}{\partial t}$ implies differentiation at constant x_i but <u>not</u> at constant x_i'.

coordinates to the radiator's coordinates.

From (18.13), (23.3), (23.4) and (23.15), we have:

$$(23.16)\quad \frac{4\pi k_o}{e}\vec{E} = \frac{1}{s^2}\vec{\nabla}s - \frac{\partial}{\partial t}\frac{\vec{\beta}}{sc} = \frac{1}{s^2}\vec{\nabla}'s - \frac{\vec{r}}{cs^3}\frac{\partial s}{\partial t'} - \frac{r}{s^2c^2}\dot{\vec{v}} + \frac{r\vec{\beta}}{cs^3}\frac{\partial s}{\partial t'}$$

Using $\vec{\nabla}'s = \frac{\vec{r}}{r} - \vec{\beta}$, and using (23.7) and (23.8) and a series of vector identities, we obtain:

$$(23.17)\quad \frac{4\pi k_o}{e}\vec{E} = \frac{1}{s^3}(\vec{r} - r\vec{\beta})(1 - \beta^2) + \frac{1}{s^3c^2}\left\{\vec{r} \times \left[(\vec{r} - \vec{\beta}r) \times \dot{\vec{v}}\right]\right\}$$

Similarly:

$$(23.18)\quad \frac{4\pi k_o\vec{B}}{e} = \frac{1}{c}\vec{\nabla} \times \left(\frac{\vec{\beta}}{s}\right) = -\frac{1}{s^2c^3}(\vec{r} \times \dot{\vec{v}}) + \frac{\vec{\beta} \times \vec{r}}{s^2c}\left[\frac{1}{r} + \frac{1}{s}\left(\frac{\vec{r}}{r} \cdot \vec{\beta} + \frac{1}{c^2}\dot{\vec{v}} \cdot \vec{r} - \beta^2\right)\right]$$

$$\frac{4\pi k_o}{e}\vec{B} = \frac{-\vec{r} \times \dot{\vec{v}}}{s^2c^3} + \frac{\vec{v} \times \vec{r}}{s^3c^2}\left(1 - \beta^2 + \frac{\vec{r} \cdot \dot{\vec{v}}}{c^2}\right)$$

$$\frac{4\pi k_o\vec{B}}{e} = \frac{\vec{v} \times \vec{r}}{c^2s^3}(1 - \beta^2) + \frac{1}{c^3s^3}\frac{\vec{r}}{r} \times \left\{\vec{r} \times \left[(\vec{r} - r\vec{\beta}) \times \dot{\vec{v}}\right]\right\}$$

and after simplification and comparison with (23.17) we see that:

$$(23.19)\quad \vec{B} = \frac{\vec{r} \times \vec{E}}{rc}$$

Hence the magnetic field is always perpendicular to $\vec{E}$ and the <u>retarded</u> radius vector $\vec{r}$.

The electric field is composed of two separate terms. The first term in Equation (23.17) varies as $1/r^2$ for large distances and is formally identical with the "convective" field of a uniformly moving charge. We might call $\vec{r}_v = \vec{r} - r\vec{\beta} = (\vec{qp})$ (see Figure 23.1) the "virtual present radius vector", i.e. the position the charge would occupy "at present" if it had continued with uniform velocity from the point p'. The $1/r^2$ term is thus simply:

$$(23.20)\quad \vec{E}_{\text{induction}} = \frac{e\vec{r}_v}{4\pi k_o s^3}(1 - \beta^2)$$

This is identical with the field (20.23) of a charge in uniform motion. This field is analogous to the quasi-static or induction field which we discussed in

connection with the radiation from variable current and charge systems in Chapter 20. Equation (23.20) does represent a non-radiating term in the sense that it does not contribute to the energy flow over an infinitely distant surface; it may however contribute to the net energy loss of the electron. The field is neither static nor convective and in general a net change in field energy will take place which must be supplied. This energy loss will cause a reaction on whatever outside force is responsible for the electron's motion. Detailed calculations in the next chapter show that this reaction force will be proportional to $-\dot{\vec{v}}$ and therefore has the character of an inertia or "electromagnetic mass."

The second term in Equation (23.17):

$$(23.21) \qquad \vec{E}_{rad} = \frac{e}{4\pi k_o s^3 c^2} \vec{r} \times (\vec{r}_v \times \dot{\vec{v}})$$

is of order $1/r$ and therefore does represent a radiation field in the sense of contributing to the energy flux over a large sphere. Similar conclusions hold for the two terms of (23.18). Let us now consider several important cases of this radiation field:

Case 1: Radiation at Low Velocity

In this case $\vec{r} \cong \vec{r}_v$, $s \cong r$

$$(23.22)(a) \qquad \vec{E}_{rad} = \frac{e}{4\pi k_o c^2 r^3}\left[\vec{r} \times (\vec{r} \times \dot{\vec{v}})\right]$$

$$(b) \qquad \vec{B}_{rad} = \frac{e}{4\pi k_o c^3 r^2} (\dot{\vec{v}} \times \vec{r})$$

Equations (23.22)(a) and (b) represent a field that is formally identical to the field of a radiating dipole (Chapter 21) of moment $\vec{p_1} = \dot{\vec{v}}\, e/\omega^2$. The angular energy distribution shown in Figure (23.2) of the radiation is therefore simply the $\sin^2\theta$ distribution discussed there.

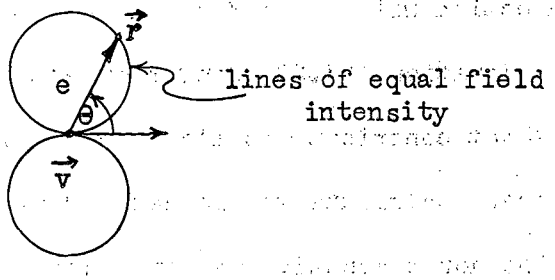


Figure (23.2)

Integrating the Poynting vector of (23.22) over a sphere, one obtains for the rate of energy loss from the accelerated charge:

(23.23) $$* \quad \frac{-dW}{dt'} = \frac{e^2 \dot{v}^2}{6\pi k_o c^3}$$

---

* $\frac{-dW}{dt'} = \frac{2e^2 \dot{v}^2}{3c^3}$ in Guassian units; all remaining equations in Chapters (22 - 25) are written so that they can be reduced to Guassian units by putting $\left[\frac{e^2}{4\pi k_o}\right]_{MKS} = \left[e^2\right]_{(Gaussian)}$

Case 2: Bremsstrahlung, if $\vec{v} \parallel \dot{\vec{v}}$:

In this case, $\vec{v}$ is not necessarily small, but $\vec{v}$ and $\dot{\vec{v}}$ are parallel. Under these conditions, the fields are simply:

(23.24)(a) $$\vec{E} = \frac{e}{4\pi k_o c^2 s^3}\left[\vec{r} \times (\vec{r} \times \dot{\vec{v}})\right]$$

(b) $$\vec{B} = \frac{er}{4\pi k_o c^3 s^3}(\dot{\vec{v}} \times \vec{r})$$

which differs only by the factor $\frac{r^3}{s^3} = (1 - \beta \cos\theta)^{-3}$ from the slow electron (dipole) case of Equation (23.22). The qualitative effect of this factor is to increase the radiated energy in the forward direction, compared to the backward direction, as shown in Figure (23.3).

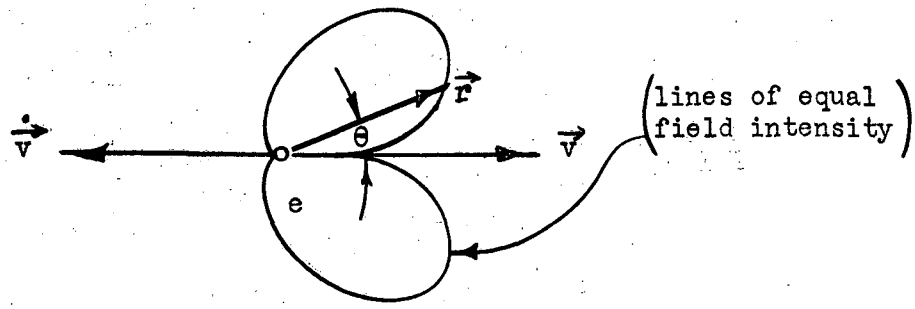


Figure (23.3)

To calculate the angular distribution of the radiated energy quantitatively considerable care has to be taken in the definition of what is meant by "the rate of radiation" of the charge. We shall compute the amount of energy lost by the electron per unit time corresponding to a time interval dt' during emission of the signa that is, the rate of energy loss $\frac{-dW}{dt'}$ of the electron itself. At a given field point, the Poynting vector $\vec{N}$ represents the energy flow per unit time. Therefore the rate of energy loss of the electron into a given solid angle $d\Omega$ is given by:

$$(23.25)\qquad \frac{-dW}{dt'}(\theta)d\Omega = \left|\vec{N}\right|\frac{dt}{dt'}\,r^2 d\Omega = \left|\vec{E}\times\vec{H}\right|\frac{dt}{dt'}\,r^2 d\Omega$$

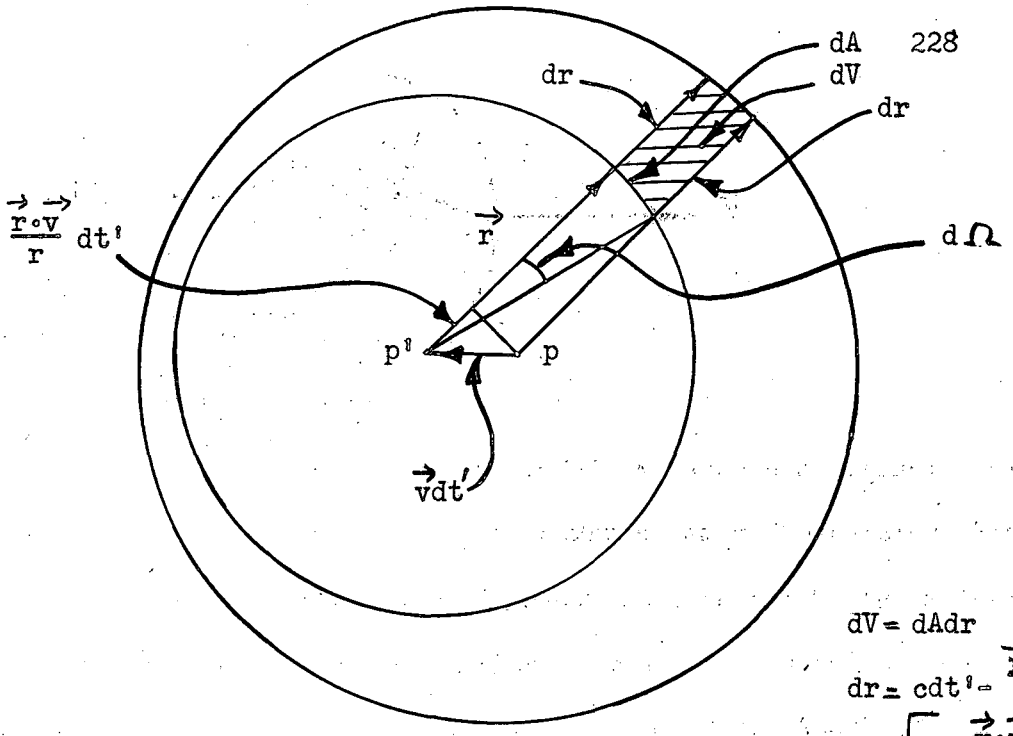
$$= k_o c E^2 r^2 \frac{dt}{dt'}\,d\Omega = k_o c E\,\frac{s}{r}\,r^2 d\Omega$$

from (23.12). Hence the directional rate of energy loss for the accelerated electron is:

$$(23.26)\qquad \frac{-dW(\theta)}{dt'} = \frac{\dot{v}^2}{c^3}\left[\frac{e^2}{16\pi^2 k_o}\right]\frac{\sin^2\theta}{(1-\beta\cos\theta)^5}$$

which is the equation giving the energy loss for the situation shown in Figure (23.3).

The correction s/r from time of signal arrival t to time of emission t' can be physically interpreted as follows: The energy emitted by the electron in a time dt' is located in the volume between two spheres, one of radius r about a fixed point p' and the other of radius r+cdt' about a point p which is a distance $-\vec{v}dt'$ from the first point as seen in Figure (23.4).



$dV = dAdr$

$dr = cdt' - \frac{\vec{r}\cdot\vec{v}}{r}dt'$

$dr = \left[c - \frac{\vec{v}\cdot\vec{r}}{r}\right]dt'$

Figure (23.4)

Consider the element of this volume dV subtending a solid angle $d\Omega = \frac{dA}{r^2}$ at p'. By geometry:

$$(23.27) \qquad dV = dA\left(c - \frac{\vec{v}\cdot\vec{r}}{r}\right)dt' = \frac{cs}{r}\,dAdt'$$

and the energy $dWd\Omega$ contained within this volume in the solid angle $d\Omega$ is:

$$(23.28) \qquad dWd\Omega = \frac{(k_o E^2 + \mu_o H^2)}{2}\,\frac{cs}{r}\,dAdt' = k_o E^2\,\frac{cs}{r}\,dAdt'$$

in agreement with our former energy loss considerations contained in Equation (23.25).

The principal application of these considerations is to the calculation of the radiation from an electron which is retarded along its own direction of motion (Bremsstrahlung). For an exact classical calculation, it would be necessary to put the exact variation of acceleration with time into the equations, using the stopping power of the material upon which the electron impinges.[x] For a simplified discussion, let us assume that $\dot{\vec{v}}$ is constant while the velocity decreases from v_o to 0. This gives a resultant "pulse" of radiation of energy:

[x] See Compton and Allison, "X-Rays in Theory and Experiment."

$$(23.29)\qquad dW = -\frac{e^2\sin^2\theta\,|\dot{\vec{v}}|}{(4\pi)^2 k_o c^3}\int_{v_o}^{0}\frac{\dot{v}\,dt'}{(1-\beta\cos\theta)^5}\,d\Omega$$

$$-dW = \frac{e^2\sin^2\theta\,|\dot{\vec{v}}|}{64\pi^2 k_o c^2\cos\theta}\left[\frac{1}{(1-\beta_o\cos\theta)^4}-1\right]d\Omega$$

The angular distribution is again tipped forward, in the direction of motion. The radiation is polarized with the electric vector lying in the plane of the radius vector and the direction of deceleration $\dot{\vec{v}}$. Equation (23.29) can be used to estimate the total efficiency of an X-ray tube. In practice, however, both the angular distribution and the polarization of the outgoing radiation are greatly modified by scattering of the electrons in the target material.

The frequency spectrum of the outgoing radiation can easily be obtained by Fourier analysis of the outgoing field. Let us assume for simplicity that a change $\Delta\vec{v}$ in the velocity takes place in a very short time $\Delta t'$ and that $\beta << 1$. During this time interval, the radiation field is:

$$(23.30)\qquad E(t) = \frac{e\sin\theta}{4\pi k_o c^2 r}\frac{\Delta v}{\Delta t'} = \frac{e\sin\theta}{4\pi k_o c^2 r}\Delta v\,\delta(t-t_o)$$

where $\dot{v} = \frac{\Delta v}{\Delta t'}$, and we have expressed $\dot{v}$ as a δ function, $\dot{v} = \delta(t-t_o)\Delta v$

$\int_{-\infty}^{\infty}\dot{v}\,dt = \Delta v$, where t_o is the time at which the radiation takes place.

If we put:

$$E(t) = \int_{-\infty}^{\infty} E_\omega\, e^{i\omega t}\,d\omega$$

then:

$$(23.31)\qquad E_\omega = \frac{1}{2\pi}\int_{-\infty}^{\infty} E(t)e^{-i\omega t}\,dt = \frac{e\sin\theta\,\Delta v}{8\pi^2 k_o c^2 r}\,e^{-i\omega t_o}$$

which, except for phase, is independent of ω. The frequency spectrum of the energy loss, corresponding to a field whose Fourier components are E_ω, can be obtained as follows:

$$-W = \int -\frac{dW}{dt'}\, dt' = k_o c \int_S \int_{-\infty}^{\infty} E^2\, dt\, dS \qquad (\beta << 1)$$

but:

$$(23.32)\quad \int_{-\infty}^{\infty} E^2 dt = \int_{-\infty}^{\infty} dt \left\{ \left[\int_{-\infty}^{\infty} E_\omega e^{i\omega t}\, d\omega \right] \left[\int_{-\infty}^{\infty} E_{\omega'} e^{i\omega' t}\, d\omega' \right] \right\}$$

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} E_\omega E_{\omega'} \left[\int_{-\infty}^{\infty} e^{i(\omega+\omega')t}\, dt \right] d\omega d\omega'$$

$$= 2\pi \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} E_\omega E_{\omega'}\, \delta(\omega+\omega')\, d\omega\, d\omega'$$

$$= 2\pi \int_{-\infty}^{\infty} E_\omega E_{-\omega} d\omega = 4\pi \int_0^{\infty} |E_\omega|^2\, d\omega$$

since $E_{-\omega} = E_\omega^*$. Hence the energy loss corresponding to a given frequency band $d\omega$ is:

$$(23.33)\quad -W_\omega d\omega = 4\pi k_o c \left[\int_S |E_\omega|^2\, dS \right] d\omega$$

Hence for our specific spectrum (23.31):

$$(23.34)\quad -W_\omega d\omega = \frac{e^2}{4\pi k_o c} \cdot \left(\frac{\Delta v}{c}\right)^2 \frac{d\omega}{2\pi} \int_0^{\pi} \sin^2\theta \sin\theta\, d\theta$$

$$= \frac{e^2}{3\pi k_o c} \left(\frac{\Delta v}{c}\right)^2 \frac{d\omega}{2}$$

The spectrum is thus constant on a frequency scale. Actually the spectrum will be cut off at the point where the kinetic energy of the electron:**

$$(23.35)\quad K.E. = 1/2\ mv^2 = h\nu_{max}$$

The ideal spectra on a frequency and wavelength scale are shown in Figure (23.5).

Equation (23.34) can be expressed in terms of the number of quanta $dN = -W_\omega\, d\omega/\hbar\omega$ which are "shaken off" during the velocity change. This gives:

$$dN_\omega = \frac{e^2}{4\pi k_o \hbar c} \cdot \frac{2}{3\pi}\left(\frac{\Delta v}{c}\right)^2 \frac{d\omega}{\omega} = \frac{1}{137} \cdot \frac{2}{3\pi}\left(\frac{\Delta v}{c}\right)^2 \frac{d\omega}{\omega}$$

The number of low frequency quanta diverges although their energy is finite. This feature is also true in the more exact treatment.*

---

* See e.g. Heitler, "Quantum Theory of Radiation."

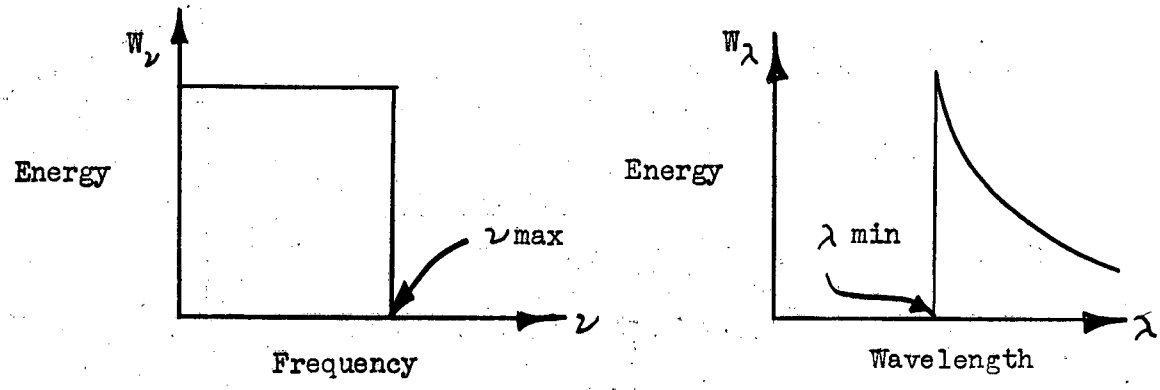


Figure (23.5)

Case 3: No Restrictions on the Acceleration or Velocity

In this case, $\vec{v}$ is not necessarily small or parallel to $\dot{\vec{v}}$. The general radiation field (23.17), in combination with the rate of radiation considerations that lead to (23.25) give the following general relation for the directional rate of radiation:

$$(23.36) \qquad \frac{-dW}{dt'}(\theta)\, d\Omega = \frac{e^2 r}{16\pi^2 k_o s^5 c^3}\left[\vec{r} \times (\vec{r}_v \times \dot{\vec{v}})\right]^2 d\Omega$$

It turns out that, when integrated over the total solid angle, this leads to:

$$(23.37) \qquad \frac{-dW}{dt'} = \frac{e^2}{6\pi k_o c^3}\left[\frac{\dot{v}^2 - (\vec{\beta} \times \dot{\vec{v}})^2}{(1-\beta^2)^3}\right]$$

for the total rate of radiation. We shall see later that this expression can be derived by a simple relativistic transformation of Equation (23.23) to a moving frame of reference.

Let us now consider the practically important case of an electron moving in

**Even classically the extent of the spectrum to high ω is only due to our simplifying assumption of zero collision time. A Fourier analysis of a finite collision time process will automatically remove the high frequency components.

a circle of radius a with a constant angular velocity ω, as shown in Figure (23.6).

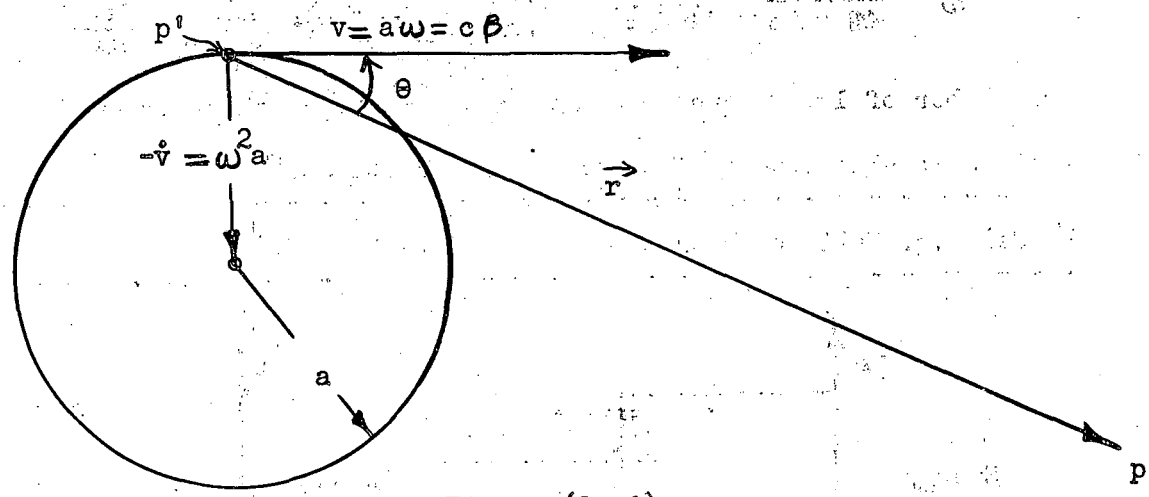


Figure (23.6)

Since $\vec{r}_v = \vec{r} - \vec{\beta} r$; $\vec{\beta} \cdot \vec{r} = \beta r \cos\theta$; $\dot{\vec{v}} \cdot \vec{r} = \dot{v} r \sin\theta$; we have:

$$(23.38) \qquad \left[\vec{r} \times (\vec{r}_v \times \dot{\vec{v}})\right]^2 = \left[(\vec{r} - \vec{\beta} r) r\dot{v} \sin\theta - \dot{\vec{v}} r^2 (1 - \beta \cos\theta)\right]^2$$
$$= \dot{v}^2 r^4 \left[\beta - \cos\theta\right]^2$$

So we have:

$$(23.39) \qquad \frac{-dW}{dt'}(\theta)\, d\Omega = \frac{e^2 \dot{v}^2}{16\pi^2 k_o c^3} \frac{(\beta - \cos\theta)^2}{(1 - \beta\cos\theta)^5}\, d\Omega \,.$$

The resultant pattern has zeros at $\theta = \cos^{-1}\beta$. For large β, the pattern is very much more intense in the forward direction than in the backward direction as is shown in Figure (23.7). As $\beta \to 1$, the radiation becomes a sharp forward ray.

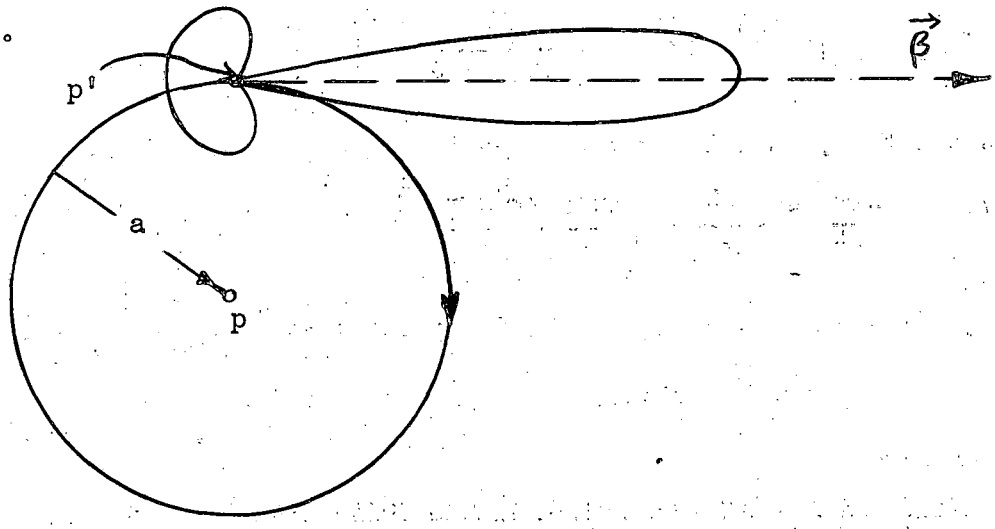


Figure (23.7)

The total rate of radiation becomes, from (23.37) or integration of (23.39):

$$(23.40)^{*}\quad -\frac{dW}{dt'} = \frac{e^2\dot{v}^2}{6\pi k_o c^3}\frac{1}{(1-\beta^2)^2} = \frac{e^2a^2\omega^4}{6\pi k_o c^3}\frac{1}{\left[1-(a\omega/c)^2\right]^2} = \frac{e^2a^2\omega^4}{6\pi k_o c^3}\left(\frac{W}{W_o}\right)^4$$

---

* Since, relativistically the energy of the electron is given by $W = W_o/\sqrt{1-\beta^2}$, the rate of radiation is proportional to $\left(\frac{W}{W_o}\right)^4$ for high velocities at constant radius and to $\left(\frac{W}{W_o}\right)^3$ in a constant magnetic field but with a variable radius. $W_o = m_o c^2$ where m_o is the rest mass of the electron. W = total energy of the radiating electron.

In the general case of arbitrary direction of $\vec{v}$ and $\dot{\vec{v}}$, the position of the nodes can be constructed graphically, as shown in Figure (23.8). Since the radiation fields vanish when $\vec{r}_v = \vec{r} - \vec{\beta}r$ is parallel to $\dot{\vec{v}}$, we can find two points A and B lying on a circle of radius r about O and which lie on a straight line parallel to $\dot{\vec{v}}$ which passes through a point Q such that $\vec{OQ} = r\vec{\beta}$. OA and OB are then the two nodal lines; the nodal lines are always in the plane defined by $\vec{v}$ and $\dot{\vec{v}}$.

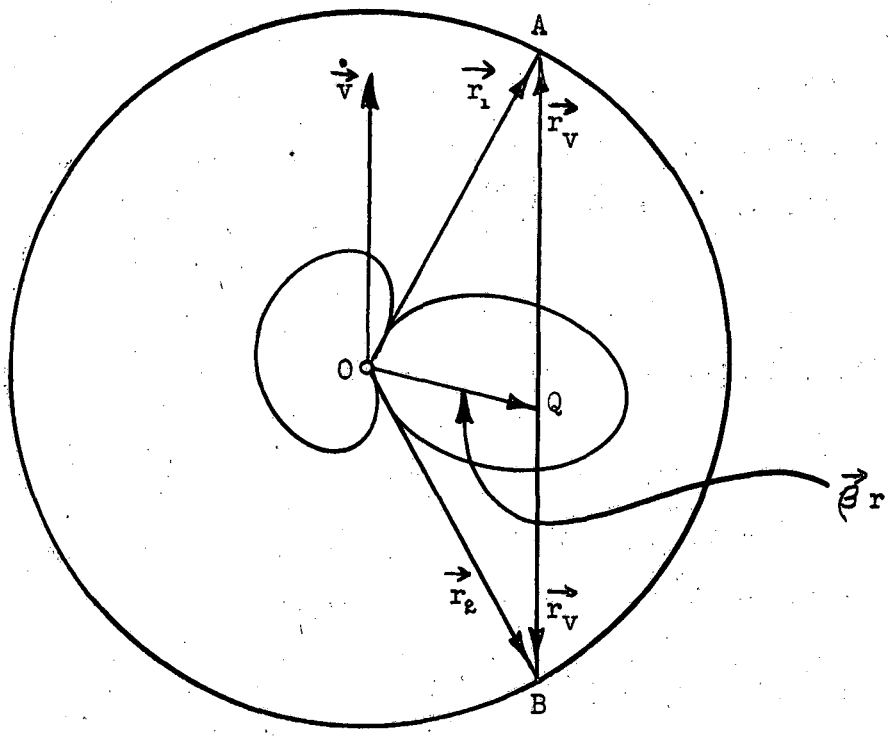


Figure (23-8)

Classical Cross Section for Bremsstrahlung in a Coulomb Field

At low energies Eq. (23.29) (due to Sommerfeld) gives the approximate yield of low energy x-rays if empirical values of $\dot{v}$ are used. At higher energies and particularly for heavier targets, the x-ray yield has to be considered to be due to radiation in deflection of the electron in the Coulomb field of the nucleus. We can make a simple classical calculation which contains essentially all of the physical features of the more detailed treatment.

Consider an electron of velocity $\vec{v}$ passing a nucleus of charge Ze at a distance b. The electron will suffer a transverse acceleration:

$$(23.41) \quad \dot{v} = \frac{Z e^2}{4\pi k_o m} \quad \frac{b}{(b^2 + v^2 t^2)^{3/2}}$$

In order to calculate the radiation loss and the spectral distribution using Eq. (23.33) we have to Fourier analyze $\dot{v}$. This gives:

$$(23.42) \quad \dot{v}_\omega = \frac{1}{2\pi}\int_{-\infty}^{\infty} \dot{v}(t)\, e^{-i\omega t} dt = \frac{1}{2\pi} \cdot \left(\frac{2Ze^2 b}{4\pi k_o m}\right) \int_0^\infty \frac{\cos \omega t}{(b^2 + v^2 t^2)^{3/2}}\, dt$$

$$= \left(\frac{1}{2\pi}\right) \frac{Z e^2}{4\pi k_o m} \quad \frac{2\omega}{v^2} K_1 \left(\frac{\omega b}{v}\right)$$

The function $K_1\left(\frac{\omega b}{v}\right)$ has the property that $K_1\left(\frac{\omega b}{v}\right) \approx \frac{v}{\omega b}$ for $\frac{\omega b}{v} < 1$ and $K_1\left(\frac{\omega b}{v}\right)$ decreases exponentially for large values of the argument. We shall therefore take:

$$(23.43) \quad \dot{v}_\omega = \frac{1}{\pi} \quad \frac{Z e^2}{4\pi k_o m} \quad \frac{1}{vb} \qquad \omega < \frac{v}{b}$$

$$= 0 \qquad \omega > \frac{v}{b}$$

The radiated energy/unit solid angle is thus, from Eq. (23.33):

$$(23.44) \quad -W_\omega\, d\omega\, d\Omega = \frac{Z^2 r_o^2}{\pi^2} \cdot \left(\frac{c}{vb}\right)^2 \quad \frac{e^2 \sin^2\theta}{4\pi k_o c}\, d\omega\, d\Omega$$

Here $r_o = \frac{e^2}{4\pi k_o m c^2}$ is the classical electron radius. This can be expressed as a cross section for emitting quanta of energy $\hbar\omega$:

$$(23.45) \quad \frac{d\sigma}{d\omega} = \frac{4\pi Z^2 r_o^2}{\beta^2} \left(\int_0^\pi \sin^3\theta d\theta\right) \int_{b_{min}}^{b_{max}} \frac{db}{b} \quad \left(\frac{d\omega}{\omega}\right) = \frac{16}{3} \frac{Z^2 r_o^2}{\beta^2} \left[\log\left(\frac{b_{max}}{b_{min}}\right)\right] \left(\frac{d\omega}{\omega}\right)$$

$$\omega < \frac{v}{b_{min}}$$

The limit b_{max} is defined by the effective limits to the Coulomb field due to the "screening" of extra nuclear electrons. The limit b_{min} is defined by the fact that the electron is localized only to a distance of the order of $\hbar/mv$. This same limit also puts the cut-off of the spectrum at $\hbar\omega \approx mv^2 \approx$ kinetic energy.

This classical calculation is of course approximate. The approximations made here are: 1. Only the transverse accelerations are considered. 2. Relativistic considerations are ignored. 3. The quantum cut-off is introduced only in order of magnitude.

CHAPTER 24

ENERGY AND MOMENTUM BALANCE - RADIATION REACTION

It was shown in Chapter 16 that the law of conservation of mementum could be reconciled with an electrical interaction of the form:

$$\vec{F} = e(\vec{E} + \vec{v} \times \vec{B}) \tag{24.1}$$

only if a momentum:

$$\vec{g} = \frac{\vec{N}}{c^2} = \frac{\vec{E} \times \vec{H}}{c^2} \tag{24.2}$$

is assumed to be contained per unit volume of the electromagnetic field. Similarly the electromagnetic field carries an energy density of the amount:

$$U_v = \frac{\vec{E} \circ \vec{D} + \vec{H} \circ \vec{B}}{2} \tag{24.3}$$

in order that the law of conservation of energy of a closed system containing both matter and radiation be obeyed.

If we attempt to apply these considerations to the radiation field of an electron, we are led to the following conclusions:

1. If an electron is in uniform motion, its field will add to the momentum of the electron, since for a small virtual change in velocity, the momentum of both the particle and the field would change simultaneously.
2. If an electron radiates by virtue of its acceleration, produced by an external force, the external force must supply both the radiated energy and the momentum required by the change in fields. The only way the external force can supply the energy is by a reaction force of the electron produced by action of the radiation field of the electron on the electron itself.

Let us examine these two conclusions in somewhat greater detail. Let us first consider the electron in uniform motion moving with a velocity $v << c$. If the velocity is changed by a small amount $\delta\vec{v}$, the magnetic field changes

by an amount:

$$(24.4) \qquad \delta\vec{B} = \frac{\mu_0}{4\pi}\frac{e\,\delta\vec{v} \times \vec{r}}{r^3}$$

Consider the magnetic flux passing through an area in a narrow strip normal to $\delta\vec{v}$ and passing from the electron to ∞ as shown in Figure (24.1).

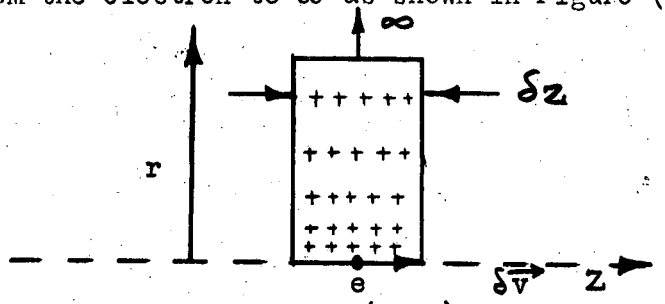


Figure (24.1)

The flux change thus produces a net voltage around the path indicated. At a point near the electron the field is given by $\frac{\partial E_z}{\partial r} = \frac{\partial B}{\partial t}$, and hence the impulse received by the charge due to a change $\delta\vec{v}$ in velocity is of the order:

$$(24.5) \qquad \delta g = e\,E_z\,\delta t = e\int_{\infty}^{r_o} \delta B\,dr = \frac{e^2}{4\pi k_o c^2}\,\frac{\delta v}{r_o}$$

We see immediately that this impulse would be infinite if the electron were a point electron; if the electron has a finite extent of order r_o, then the magnetic field would react its maximum at r_o and then go to zero at the center of the electron. For a finite electron we would have approximately:

$$(24.6) \qquad \delta\vec{g} \cong \frac{e^2}{4\pi k_o r_o c^2}\,\delta\vec{v}$$

which as far as the accelerating force producing $\delta\vec{v}$ is concerned is equivalent to an effective "mass" of the order:

$$(24.7)^* \qquad m \cong \frac{e^2}{4\pi k_o r_o c^2}$$

---

* $m = \frac{e^2}{r_o c^2}$ in Gaussian units

This mass expression was derived by directly computing the inertial reaction of the field on the driving force. By the general considerations of Chapter 16, we should also obtain the same expression if we integrate $\vec{g} = \frac{\vec{E} \times \vec{H}}{c^2}$ over the field. Again, if we take the field to be that of a point charge, the integration will diverge so that a lower limit corresponding to a finite radius r_o of the electron must be introduced into the region of integration. For small values of β, this integration can be carried out easily, using the fields of Chapter 20, Equation (20.20) ff. The result is*, taking $s \cong r$:

$$(24.8) \qquad \vec{G} = \int_{r_o}^{\infty} \vec{g}\, dv = \frac{4}{3} \frac{U_o}{c^2} \vec{v}$$

* See Becker, "Theorie der Elektrizitat," Vol. II, Pg. 42 ff.

where:

$$(24.9) \qquad U_o = \frac{1}{2} \int_{r_o}^{\infty} \vec{E} \cdot \vec{D}\, dv \cong \frac{1}{8\pi k_o} \frac{e^2}{r_o}$$

is the electrostatic field energy of a point charge e of finite radius r_o. With the exception of the numerical factor, this expression is again in agreement with an electromagnetic mass of the order:

$$(24.10) \qquad m \cong \frac{e^2}{4\pi k_o r_o c^2}$$

as before. The agreement between (24.10) and (24.7) is of course required by the general considerations of Chapter 16.

At high velocities ($\beta \to 1$) the fields of Chapter 20 must be used with $s \neq r_o$. The resultant relation between momentum and velocity depends on the assumption that is made about the behavior of the shape of the electron on its motion. It may be assumed that the shape of the electron remains rigid, independent of the motion, as in the Abraham electron, or contracts, as was assumed by Lorentz, in the ratio $\sqrt{1-\beta^2}$ in the direction of motion. If this

latter assumption, which is plausible today by virtue of relativistic arguments, is made, we obtain:

$$(24.11)\quad \vec{G} = \frac{4}{3}\,\frac{U_o}{c^2}\,\frac{\vec{v}}{\sqrt{1-\beta^2}}$$

where U_o is given by (24.9). With the exception of the factor 4/3, (24.11) is of the general form:

$$(24.12)\quad \begin{cases} \vec{G} = m\vec{v} \\ m = \dfrac{m_o}{\sqrt{1-\beta^2}} \\ U_o = m_o c^2 \end{cases}$$

m_o = electronic rest mass.

which we shall later prove are the correct relativistic expressions relating mass, energy, and momentum. The extra factor 4/3 is however, significant and indicates that the purely "electromagnetic mass" calculated here is too big by this factor.* In order to make the total mass a relativistically correct quantity, an additional mass $\frac{-U_o}{3c^2}$ must be present, whose origin is not electromagnetic, which will lead to the overall observed quantities which obey the relativistically correct relations (24.12). This extra mass (or energy) $-U_o/3c^2$ presumably represents the non-electromagnetic binding which must be present to make the charge system of the electron stable.

*Note that the ratio of electromagnetic momentum to total field energy is independent of r_o and hence independent of any of the approximations made in the fields.

We have considered how the momentum of the field affects the effective mass of an electron and also how physically this mass term is due to a reaction on the force which tends to change the velocity of an electron which is in steady motion. We now would like to obtain the reaction force which must be present if the electron is accelerated with an acceleration $\dot{\vec{v}}$, giving rise to a rate of energy loss:

$$(23.23)\quad \frac{-dW}{dt'} = \frac{e^2\dot{v}^2}{6\pi k_o c^3}$$

where we are taking β to be small.**

**This restriction is not basic since, as will be explained later, it can be removed by a relativistic transformation.

This reaction force should be given by:

$$(24.15)\quad \vec{F}\cdot\vec{v} + \frac{e^2}{6\pi k_o}\,\frac{\dot{\vec{v}}^2}{c^3} = 0$$

This equation obviously has no solution for $\vec{F}$ which can be instantaneously correct for all time, since $\vec{v}$ and $\dot{\vec{v}}$ are basically uncorrelated. It is

therefore not possible to solve (24.15) instantaneously but only when averaged over a large enough period of time. This means of course that we are not balancing energy between the force and the radiation field instantaneously but that an extra fluctuation will be available which will be stored in the induction field. On this basis, Equation (24.15) becomes:

$$(24.16) \qquad \int_{t_1}^{t_2} \left(\vec{F} \cdot \vec{v}\right) dt' + \int_{t_1}^{t_2} \frac{e^2}{6\pi k_o c^3} \dot{v}^2 dt' = 0$$

integrating by parts:

$$(24.17) \qquad \int_{t_1}^{t_2} \left[\vec{F} - \frac{e^2 \ddot{\vec{v}}}{6\pi k_o c^3}\right] \cdot \vec{v} dt' + \left[\frac{e^2}{6\pi k_o c^3} \left(\vec{v} \cdot \dot{\vec{v}}\right)\right]_{t_1}^{t_2} = 0$$

The integrated term is the "fluctuation" term referred to above; for a periodic motion or an acceleration occurring over a limited time it will not affect the integrated energy balance. On the average, energy will be preserved if we put:

$$(24.18) \qquad \vec{F}_{rad} = \frac{e^2}{6\pi k_o c^3} \ddot{\vec{v}}$$

as the radiation reaction force. The total reaction caused by the interaction of the fields of the electron on itself demanded by conservation of energy and momentum is therefore:

$$(24.19) \qquad \vec{F} = \frac{e^2}{6\pi k_o c^3} \ddot{\vec{v}} - m_{el.mag.} \dot{\vec{v}}$$

where:

$$(24.20) \qquad m_{el.mag.} = \frac{4}{3} \frac{U_o}{c^2} \cong \frac{e^2}{6\pi k_o c^2 r_o}$$

Note that we have arrived at this reaction force from the point of view of the conservation laws rather than by consideration of the direct interaction.

It will be instructive to also study the direct computation of (24.19), which was originally due to Lorentz, in which $\vec{F}$ is derived by direct integration of the interaction of the radiation field, of part of the electron, on the other parts of the electron. This calculation very clearly shows the limitations of the theory. In making the calculation we shall make the following four assumptions.

(1). We shall choose a frame such that the element de of the electron on which another element de' acts, is at rest. $\vec{v}(t)_{de} = 0$

(2). None of the quantities $\vec{v}$, $\dot{\vec{v}}$, $\ddot{\vec{v}}$ etc. change greatly during the time it takes for an electromagnetic signal to cross the electron. This is equivalent to:

$$v << c \tag{24.21}$$

$$\dot{v} << \frac{c^2}{r_o} \tag{24.22}$$

$$\ddot{\vec{v}} << \frac{\dot{v}c}{r_o} \text{ etc.} \tag{24.23}$$

The solution is effectively a power series expansion in $\tau_o = \frac{r_o}{c} = \frac{e^2}{4\pi k_o mc^3}$ *,

---

* $\tau_o = \frac{e^2}{mc^3}$ in Guassian units.

---

(3). The fields based on the Wiechert-Liénard potentials will be used even if their correctness at dimensions of the order r_o is doubtful. Therefore, only terms not containing r_o explicitly will be considered as having physical significance.

(4). The electron distribution has spherical symmetry.

Consider an element de of the charge at p affected by the element de' at p', as shown in Figure (24.2).

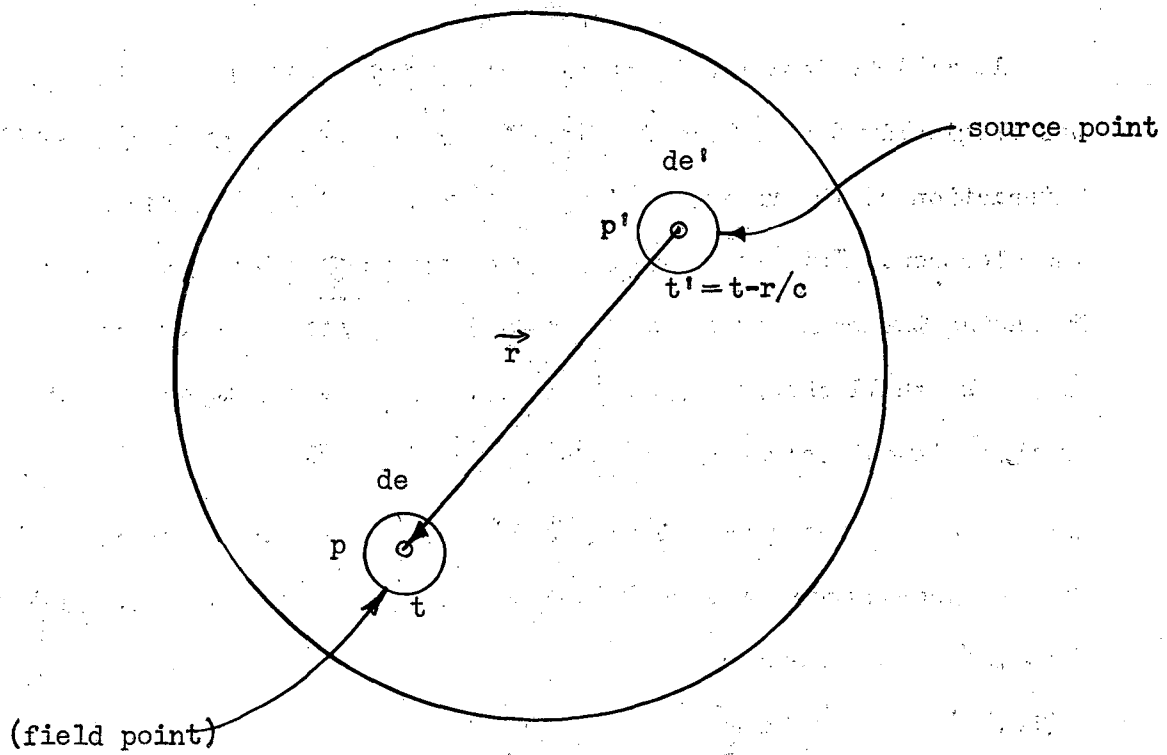


Figure (24.2)

Using Equation (23.17), the electric field at p is given by:

$$(24.24)\qquad 4\pi k_o d\vec{E}(t) = \frac{de'}{s^3}\left\{\frac{1}{c^2}\vec{r}\times\left[\left(\vec{r}-\frac{\vec{v}(t')r}{c}\right)\times\dot{\vec{v}}(t')\right]+\left(1-\frac{v^2(t')}{c^2}\right)\left(\vec{r}-\frac{\vec{v}(t')r}{c}\right)\right\}$$

Note that the field is expressed in terms of the electron's condition at t' and hence for totally arbitrary motions, the problem would be insoluble without knowledge of the electron's entire past. It is only the restriction (2) above which permits discussion of the problem at all. In order to allow integration of (24.24) over the entire electron, all velocities and accelerations must be referred to the same time of arrival t of the signal at p. Using assumption (1), $\vec{v}(t) = 0$, and assumption (2), we have:

$$(24.25)\qquad \dot{\vec{v}}(t') = \dot{\vec{v}}(t) - \frac{r}{c}\ddot{\vec{v}}(t)$$

$$(24.26)\qquad \vec{v}(t') = \frac{-\dot{\vec{v}}(t)r}{c} + \frac{\ddot{\vec{v}}(t)}{2}\left(\frac{r}{c}\right)^2$$

$$(24.27)\qquad s^{-3} = r^{-3}\left[1 - \frac{3\dot{\vec{v}}(t)\cdot\vec{r}}{c^2} + \frac{3}{2}\left(\frac{\ddot{\vec{v}}(t)\cdot\vec{r}}{c^3}\right)r\right]$$

Let us carry terms to $1/c^3$ only; to this order $v^2/c^2 = O(1/c^4)$ will be neglected.

Substituting:

$$(24.28)\quad 4\pi k_o d\vec{E} = de'\left[\frac{\vec{r}(\vec{r}\cdot\dot{\vec{v}})}{r^3c^2} - \frac{\dot{\vec{v}}}{rc^2} - \frac{\vec{r}(\vec{r}\cdot\ddot{\vec{v}})}{c^3r^2} + \frac{\ddot{\vec{v}}}{c^3} + \frac{\vec{r}}{r^3} - \frac{3\vec{r}(\dot{\vec{v}}\cdot\vec{r})}{r^3c^2} + \frac{3}{2}\frac{\vec{r}}{c^3r^2}(\ddot{\vec{v}}\cdot r) + \frac{\dot{\vec{v}}}{c^2r} - \frac{\ddot{\vec{v}}}{2c^3}\right]$$

$$= de'\left[\frac{-2\vec{r}(\dot{\vec{v}}\cdot\vec{r})}{r^3c^2} + \frac{1}{2}\frac{\vec{r}(\ddot{\vec{v}}\cdot\vec{r})}{r^2c^3} + \frac{\vec{r}}{r^3} + \frac{\ddot{\vec{v}}}{2c^3}\right]$$

In averaging this expression over a sphere, let us write (24.28) in tensor notation:

$$(24.29)\qquad dE_i = \frac{de'}{4\pi k_o}\left[\frac{-2r_i(\dot{v}_j r_j)}{r^3c^2} + \frac{1}{2}\frac{r_i(r_j\ddot{v}_j)}{r^2c^3} + \frac{r_i}{r^3} + \frac{\ddot{v}_i}{2c^3}\right]$$

When averaging over a region of spherical symmetry, we have $\overline{r_i} = 0$, $\overline{r_i r_j} = \frac{1}{3}r^2\delta_{ij}$, and hence:

$$(24.30)\qquad dE_i = \frac{de'}{4\pi k_o}\left[-\frac{2}{3}\frac{\dot{v}_i}{c^2 r} + \frac{2}{3}\frac{\ddot{v}_i}{c^3}\right]$$

Integrating over de' and de we obtain for the self-force:

$$(24.31)\qquad \vec{F} = \frac{e^2\dot{\vec{v}}}{6\pi k_o c^3} - m_{el.mag.}\dot{\vec{v}}$$

where:

$$(24.32)\qquad m_{el.mag.} = \frac{4}{3c^2}\iint\frac{de\,de'}{8\pi k_o r} = \frac{4}{3c^2}U_o$$

where U_o is the electrostatic energy of the electron in its own field. This is in complete agreement with the results (24.19) required by the conservation laws.

Note that the $\ddot{\vec{v}}$ term in (24.30) is independent of the extent of the electron and therefore presumably independent of the detailed structure of the electron. The $\dot{\vec{v}}$ term, i.e., the mass term, is structure-dependent, but its relation to the electrostatic energy is not. It is clear, however, that the electromagnetic mass is at least of the order of the experimental mass m. It

would be useless to carry this calculation to higher order, since the following terms would form an ascending series in r_o and would therefore be structure-dependent and presumably not of physical significance.

The preceding discussion shows that there are the following unsolved difficulties in classical electron theory:

(1). The theory gives a relation between electrostatic field energy and mass, but does not indicate why either of them should be finite for a point field source.

(2). The electromagnetic mass of the electron is insufficient to account for the entire mass; the electron must have non-electromagnetic mass of unknown origin to account for its stability.

(3). The condition (24.23) which was necessary to permit the radiation calculation is equivalent to putting:

$$(24.33) \qquad \left| m\dot{\vec{v}} \right| >> \left| \frac{e^2 \ddot{\vec{v}}}{6\pi k_o c^2 r_o} \; \frac{r_o}{c} \right| = \left| \vec{F}_{rad} \right|$$

This means that there must be an external force large compared to the radiation reaction force in order for the theory to be valid. Hence, the force equations derivedhere do not apply to a free electron but only to a bound electron. If (24.31) were applied to a free electron, we would have a differential equation of the type:

$$(24.34) \qquad k\ddot{\vec{v}} - \dot{\vec{v}} = 0$$

$$k \cong \tau_o \cong \frac{e^2}{4\pi k_o m c^3} \cong 10^{-23} \text{ sec.}$$

which leads to an integral involving terms in e^{t/τ_o} which leads to an extreme instability of position for a free electron which is contrary to fact.

Quantum mechanics has not solved these difficulties, but has in fact added others to them, which cause similar difficulties to exist for the charge as

they do classically for the mass of the electron. The experimental values of e and m can of course be used as before to describe the external behavior of the electron; the real problems concern themselves with the detailed description of the electron itself.

CHAPTER 25

RADIATION FROM A BOUND ELECTRON

The theory of radiation reaction discussed in Chapter 24 can with some confidence be applied to radiation from a bound electron, provided that none of the conditions given there are violated.

Transient Vibrations:

Consider an electron bound to a center of force with a harmonic force $\vec{F}=-k\vec{x}$ corresponding to a natural frequency $\omega_o = \sqrt{k/m}$. The equation of motion, including the self force of the electron, is then:*

$$(25.1) \qquad \ddot{x} + \omega_o^2 x = \frac{e^2}{6\pi k_o c^3 m} \dddot{x} = \frac{2}{3}\tau_o \dddot{x}$$

In accordance with Eq. (24.33), this relation is valid only if the right-hand side is small. If this is true, we can write:

$$(25.2) \qquad \dddot{x} \simeq -\omega_o^2 \dot{x}$$

and hence if we let $\gamma = \frac{2}{3}\tau_o \omega_o^2 = \frac{e^2}{6\pi k_o c^3 m}\omega_o^2$, we have:

$$(25.3) \qquad \ddot{x} + \gamma\dot{x} + \omega_o^2 x = 0$$

whose solution for small γ is:

$$(25.4) \qquad x = Ae^{i\omega_o t} e^{-(\gamma/2)t}$$

Eq. (25.4) corresponds to a damped wave train after a given amplitude has been excited by an external impulse. The energy of the wave falls off as

$$(25.5) \qquad e^{-\gamma t}$$

and hence γ^{-1} is the mean duration of the radiated pulse when averaged over energy. This is the classical quantity corresponding to the quantum mechanical "life-time" of the excited state produced by the external impulse.

The condition for validity of Eq. (25.3) is that

---

* The mass reaction term is of course included in the $m\ddot{x}$ term.

$$\frac{\gamma}{\omega_o} \ll 1; \quad \text{or } \omega_o \ll \tau_o^{-1}$$

or in terms of the quantum energy of the outgoing radiation:

$$(25.6) \qquad \hbar\omega_o \ll \left(\frac{4\pi k_o \hbar c}{e^2}\right) mc^2 = 137\, mc^2 \cong 70 \text{ Mev.}$$

This limitation is therefore entirely unimportant in the optical or X-ray frequency regions.

The line-width of the emitted radiation can be obtained directly by Fourier analysis of Eq. (25.4). If:

$$E = \int_{-\infty}^{\infty} E_\omega e^{i\omega t}\, d\omega = E_o e^{i\omega_o t}\, e^{-(\gamma/2)t}$$

then:

$$E_\omega = \frac{E_o}{2\pi} \int_0^\infty e^{i\omega_o t}\, e^{-(\gamma/2)t}\, e^{-i\omega t}\, dt$$

or:

$$(25.7) \qquad E_\omega = \frac{E_o}{2\pi} \frac{1}{i(\omega - \omega_o) + \frac{\gamma}{2}}$$

Corresponding to an intensity:

$$(25.8) \qquad I_\omega = \frac{I_o \gamma}{2\pi} \frac{1}{(\omega - \omega_o)^2 + \frac{\gamma^2}{4}}$$

Normalized such that:

$$(25.9) \qquad \int_{-\infty}^{\infty} I_\omega\, d\omega = I_o$$

The full width at half energy is therefore:

$$(25.10) \qquad \Delta\omega \cong \gamma = \frac{2}{3} \tau_o \omega_o^2$$

corresponding to a width in wavelength:

$$(25.11) \qquad \Delta\lambda = \frac{2\pi \Delta\omega_o}{\omega_o^2} = 2\pi \cdot \frac{2}{3} c\tau_o = \frac{4\pi}{3} r_o \sim 10^{-12} \text{ cm.}$$

Note that this width is constant, independent of the frequency used. In practice this is of course not the only source of the width of a spectral line, since interruption of the wave train due to interatomic collision and Doppler frequency shifts also contribute to the line width.

The relation between line width and mean life γ^{-1}:

$$\Delta\omega\gamma^{-1} = 1 \tag{25.12}$$

which follows from the above relations, is of course equivalent to the relation:

$$\Delta E \Delta t \sim \hbar \tag{25.13}$$

where $\Delta E = \hbar\Delta\omega$, corresponds to the quantum mechanical relation between the lifetime and the energy width of a state.

Forced Vibrations

The previous equations give the motion and the consequent radiation of a bound electron following a transient disturbing impulse; this is the classical theory of spectral emission. Let us now consider the steady state motion of an electron in a sinusoidal external field. The resultant radiation will be coherent with the external field and will, in general, interfere with it. The relations that result from this consideration will give the classical theories of light scattering, absorption and dispersion.

The equation of motion of the bound electron in an external field $\vec{E} = \vec{E}_o e^{i\omega t}$ of angular frequency ω is:

$$\ddot{\vec{x}} + \gamma\dot{\vec{x}} + \omega_o^2\vec{x} = \frac{e}{m}\vec{E}_o e^{i\omega t} \tag{25.14}$$

where γ is the same as it was in Equation (25.3). The resultant steady state motion is given by:

$$\vec{x} = \vec{E}_o \frac{e}{m}\left[\frac{1}{\omega_o^2 - \omega^2 + i\omega\gamma}\right] e^{i\omega t} \tag{25.15}$$

Let us analyze this solution as applied to several cases.

1. Free Electron: (a) Scattering by Single Electron (Thomson Scattering)

The displacement and acceleration become, for the unbound and weakly accelerated electron:

$$(25.16) \qquad \vec{x} = -\frac{e\vec{E}}{\omega^2 m} \; ; \; \ddot{\vec{x}} = \frac{e\vec{E}}{m} \; ; \; \omega_o \cong 0 \; ; \; \gamma \cong 0$$

If $\dot{x} \ll c$, Equation (25.16) using Equation (23.21), gives rise to a radiation field:

$$(25.17) \qquad E' = \frac{e(\sin\alpha)}{4\pi k_o r c^2} \ddot{x}$$

where α is the angle between $\vec{E}$ and $\vec{r}$ as seen in Figure (25.1).

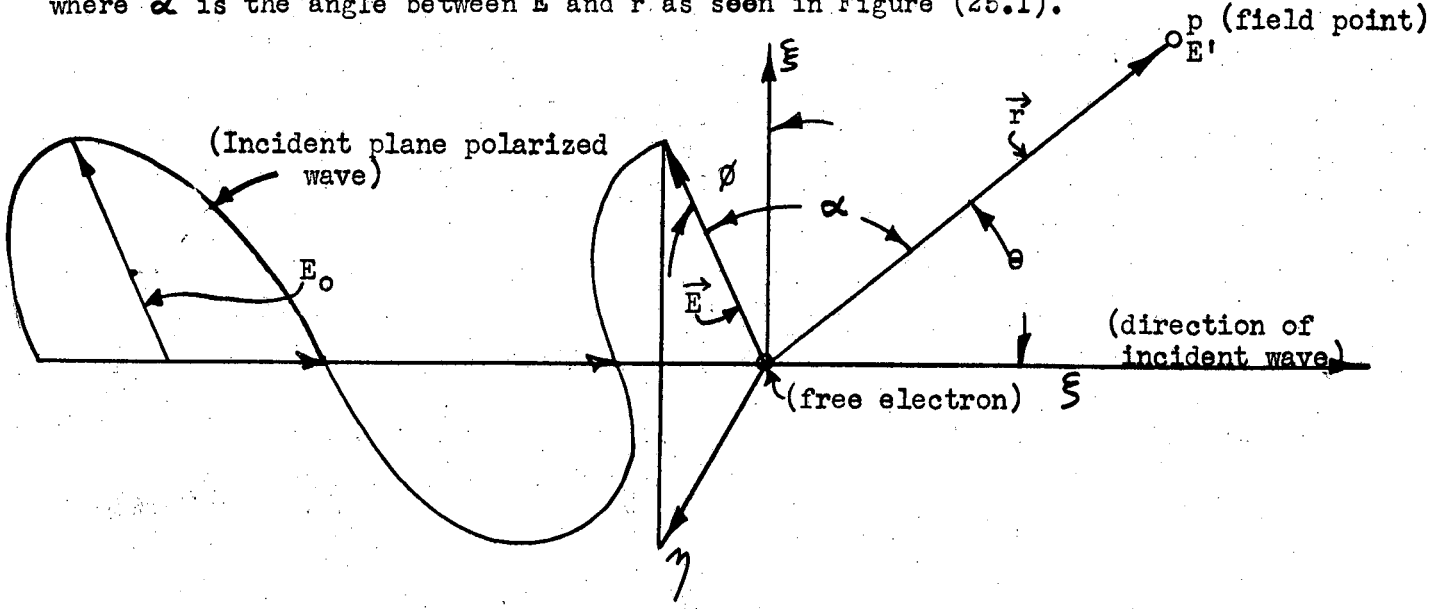


Figure (25.1)

The rate at which re-radiated energy crosses a unit area is:

$$(25.18) \qquad |N| = \sqrt{\frac{k_o}{\mu_o}} \left[\frac{e(\sin\alpha)\ddot{x}}{4\pi k_o r c^2}\right]^2$$

$$(25.19) \qquad |N| = r_o^2 \sin^2\alpha \left(\frac{I_o}{r^2}\right)$$

where $r_o = \dfrac{e^2}{4\pi k_o m c^2}$ is the classical electron radius and:

(25.20) (a) $$I_o = \frac{1}{2}\sqrt{\frac{k_o}{\mu_o}}\, E_o^2$$

is the primary intensity.

(25.19) is better expressed in terms of the scattering angle θ and the polarization angle ϕ, which are related by: (See Figure 25.1)

$$\cos\alpha = \cos\phi \sin\theta$$

i.e. $$\sin^2\alpha = 1 - \cos^2\phi\,(1 - \cos^2\theta)$$

If the primary wave is unpolarized, (i.e. randomly polarized) we have, averaging over ϕ:

(25.20) (b) $$\sin^2\alpha = \frac{1}{2}\,(1+\cos^2\theta)$$

since:

$$\overline{\cos^2\phi} = \frac{1}{2}$$

The total rate at which energy is re-radiated is therefore:

(25.21) $$\frac{dW}{dt'} = -\iint r^2 N d\Omega \qquad \ldots\ldots\ldots\ldots d\Omega = \sin\theta\, d\theta d\phi$$

$$\frac{dW}{dt'} = -I_o r_o^2 \int_0^{\pi} \frac{(1+\cos^2\theta)}{2} \cdot 2\pi \sin\theta\, d\theta$$

(25.22) $$\frac{dW}{dt'} = \frac{8\pi I_o r_o^2}{3}$$

This corresponds to an effective scattering cross-section per electron:

(25.23) $$\sigma_o = \frac{8\pi}{3} r_o^2$$

Thompson X-sect

and a differential scattering cross-section per unit solid angle:

(25.24) $$\frac{d\sigma_o}{d\Omega} = \frac{r_o^2}{2}\,(1+\cos^2\theta)$$

This cross-section is shown graphically in Figure (25.2): Note that classically

the scattering is symmetrical in the forward and backward directions.

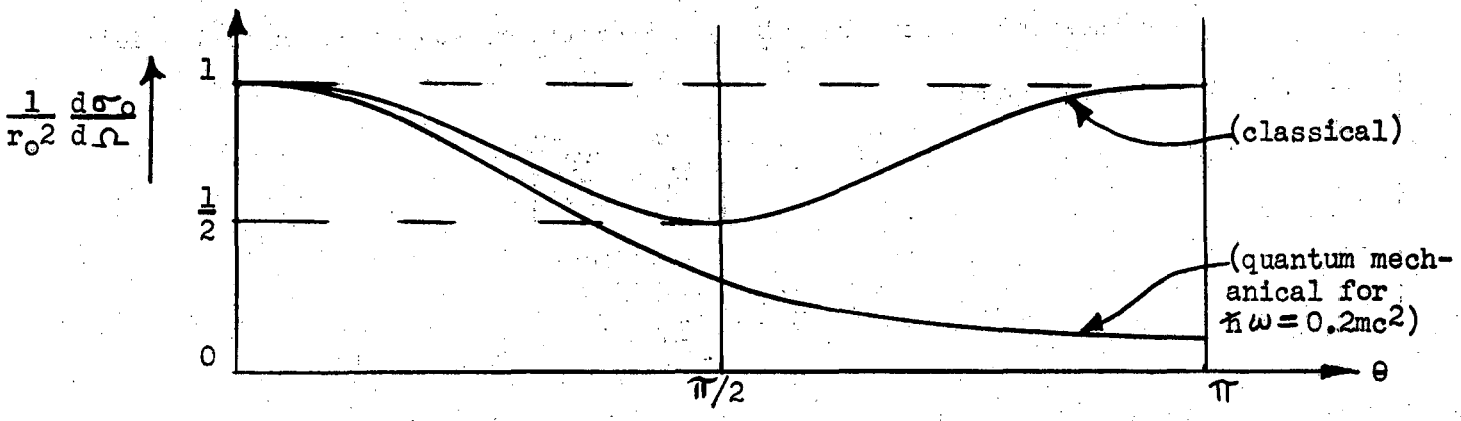


Figure (25.2)

Quantum-mechanically* the scattered radiation is more concentrated in the forward direction, even for relatively moderate energies. (See e.g. the case of scattering at 100 KEV in Figure (25.2).

* See e.g. Heitler, "The Quantum Theory of Radiation," second ed., Pp. 149-161.

(b). Volume Distribution of Free Electrons

This case applies to the following cases of practical interest:

1. Refraction of X-rays.
2. Refraction of electromagnetic waves in the ionosphere.
3. Re-radiation by a "plasma" in an electric discharge.

If we have N electrons per unit volume, each electron, when affected by an infinite incident plane wave will scatter the radiation in accordance with the Thomson scattering formula. This scattered radiation will combine coherently into a plane wave which will interfere with the incident wave, thus modifying the effective velocity of the wave. The case could be analyzed therefore by analyzing the superposition of Thomson scattered wavelets. It is, however, simpler mathematically to treat the re-radiation, due to certain polarization $\vec{P}$, from an entire volume element rather than that from a single electron of displacement $\vec{x}$. These two quantities are related by:

$$\vec{P} = Ne\vec{x} \tag{25.24}$$

The effect of the polarization is to add the polarization current $\frac{\partial \vec{P}}{\partial t}$ to the vacuum displacement current $k_o \frac{\partial \vec{E}}{\partial t}$ as a circulation source of magnetic field. Equation (14.6) becomes:

(25.25) $$\vec{\nabla} \times \vec{B} = u_o \left[k_o \frac{\partial \vec{E}}{\partial t} + \frac{\partial \vec{P}}{\partial t} \right]$$

This equation, when combined with the equation:

(14.6)(3) $$\vec{\nabla} \times \vec{E} = - \frac{\partial \vec{B}}{\partial t}$$

in the usual way, yields the homogeneous wave equations:

(24.26) $$\nabla^2 \vec{E} - \frac{1}{v^2} \frac{\partial^2 \vec{E}}{\partial t^2} = 0; \quad \nabla^2 \vec{B} - \frac{1}{v^2} \frac{\partial^2 \vec{B}}{\partial t^2} = 0$$

where:

(25.27) $$v = \frac{c}{n}$$

is the propagation velocity and:

(25.28) $$n = \sqrt{1 + \frac{|\vec{P}|}{k_o |\vec{E}|}} = \sqrt{k}$$

is the refractive index. The local field (see Eq. 3.41) is given by:

(25.29) $$\vec{E}_\ell = \vec{E} + \frac{\vec{P}}{3k_o}$$

in isotropic dielectrics and simply by $\vec{E}_\ell = \vec{E}$ in dilute systems.

We have therefore for effectively dilute systems (always true in the free electron case):

(25.30) $$k = n^2 = 1 + \frac{Ne|\vec{x}|}{k_o |\vec{E}|}$$

and from (25.16):

(25.31) $$k = n^2 = 1 - \frac{Ne^2}{k_o \omega^2 m}$$

If we put $\lambda\!\!\!^{-} = \frac{c}{\omega}$ for the reduced wavelength and $r_o = \frac{e^2}{4\pi k_o m c^2}$

for the classical electron radius, this can be written as:

$$k = n^2 = 1 - 4\pi N r_o \lambda^2 \tag{25.32}$$

The index is therefore less than unity and thus radiation penetrating free electron materials can exhibit the pheonmenon of total reflection.

Note that the refractive index, being a measure of the **net** phase velocity caused by interference of the primary wave and the induced polarization radiation, is sensitive to the sign of the scattered wave amplitudes, while the measurement of scattering cross sections is not. Investigations of the refractive index therefore yield more information about the binding of the electrons than do scattering measurements. This has recently been demonstrated in analogous experiments on scattering vs. refraction of neutrons in crystals.

2. Bound Electrons

(a) Scattering by a Bound Electron

The radiation scattering cross section of a bound electron becomes, for unpolarized primary radiation:

$$\sigma = \sigma_o \frac{\omega^4}{(\omega_o^2 - \omega^2)^2 + (\gamma\omega)^2} \tag{25.33}$$

where $\sigma_o = \frac{8\pi}{3} r_o^2$ is the Thomson cross section. The angular distribution is the same as in the free electron case. For strong binding; i.e. $\begin{cases} \omega_o >> \omega \\ \gamma << \omega_o \end{cases}$ this becomes:

$$\sigma = \sigma_o \left(\frac{\omega}{\omega_o}\right)^4 \tag{25.34}$$

giving a cross section depending on the inverse 4th power of the incident wavelength. Equation (25.34) not only holds for scattering by individual bound electrons but also for scattering by small polarizable regions if they are small compared to the incident wavelength. This is the case in scattering by small regions caused by density fluctuations in the atmosphere. Equation (25.34) then

gives the Rayleigh scattering of a gas of fluctuating density, which is responsible for the blue of the sky.

The scattering cross section can become very large if $\omega \approx \omega_o$ (resonance scattering). In this case:

$$\sigma = \sigma_o \left(\frac{\omega_o}{\gamma}\right)^2 \tag{25.35}$$

which can give a cross section greatly in excess of the classical area of the electron.

(b). Index of Refraction of Material Containing Bound Electrons

For bound electrons, Equation (25.30) becomes:

$$n^2 = k = 1 + \frac{Ne^2}{k_o m} \frac{1}{(\omega_o^2 - \omega^2 + i\omega\gamma)} \tag{25.36}$$

In the high density case, if the distinction between local and external field becomes important, this becomes (since $\vec{E}_o$ in Eq. (25.15) is now to be identified with the local field $\vec{E}_\ell$ given by (25.29)):

$$3\left(\frac{k-1}{k+2}\right) = 3\left(\frac{n^2-1}{n^2+2}\right) = \frac{4\pi N r_o}{K_o^2 - K^2 + iK\Gamma} \tag{25.37}$$

where $K_o = \frac{\omega_o}{c}$; $K = \frac{\omega}{c}$; $\Gamma = \frac{\gamma}{c}$.

If not all the electrons have the same binding or damping but only a fraction f_i has binding wave numbers K_{io} and damping width (in wave numbers) Γ_i, we obtain for the "Molar refraction:"

$$\frac{M}{g}\left(\frac{n^2-1}{n^2+2}\right) = \frac{4\pi}{3} N_o r_o \sum_i \frac{f_i}{K_{oi}^2 - K^2 + iK\Gamma_i} \tag{25.38}$$

where:
- N_o = Avogadro's number.
- M = Molecular weight.
- g = density.

This is the general dispersion formula which gives both the real index of refraction and the absorption coefficient.

CHAPTER 26

THE EXPERIMENTAL BASIS OF THE SPECIAL THEORY OF RELATIVITY

The velocity $\vec{v}$ appearing in the formulae for the radiation from a moving electron, as derived in the previous chapter, denotes the velocity relative to a frame of reference in which the wave equation:

$$\nabla^2\phi - \frac{1}{c^2}\frac{\partial^2\phi}{\partial t^2} = -g(x,y,z;t) \tag{26.1}$$

is valid. Substitution of the Galilean transformation for uniform translation parallel to the x axis:

$$\begin{aligned} x' &= x - vt \\ y' &= y \\ z' &= z \\ t' &= t \end{aligned} \tag{26.2}$$

will not preserve the form of (26.1); electromagnetic effects will therefore not be the same if observed from different frames moving with a constant velocity relative to one another. Specifically the velocity of propagation of a plane wave in vacuo would not retain its value $c = 1/\sqrt{\mu_o k_o}$. If we accept the basic correctness of Maxwell's equations and classical kinematic laws, then there exists a privileged frame of reference, the classical "ether frame", which is the only one in which Maxwell's equations are valid and in which light is propagated with the velocity c.

The concept of a privileged frame is foreign to classical mechanics. The basic equation of motion of n point particles interacting with potential functions V depending on their separation:

$$m_i\ddot{\vec{x}}_i = -\sum_j \vec{\nabla}\left[V(x_i - x_j)\right] \tag{26.3}$$

is not changed by a substitution of the type (26.2). This invariance of the laws of mechanics is only true for transformations of the type (26.2), representing linear translations with uniform velocity and non-accelerated reference frames. A frame in which a body on which no forces are acting is unaccelerated

is called an "inertial" frame; we can therefore formulate the invariance of the laws of mechanics under the transformations (26.2) by saying that in classical mechanics all inertial frames are equivalent.*

*The definition of an "inertial frame" given above is at best very unsatisfactory, since the absence of forces can only be detected by the absence of acceleration, unless the sources of the force are known. In a field theory an inertial frame is therefore undefinable and a force field and a state of acceleration of a frame are basically indistinguishable. This fact is only recognized in the general theory of relativity; we will assume here that an "inertial frame" can be realized.

This statement is sometimes known as the principle of Galilean relativity.

Therefore, we have seen that the principle of Galilean relativity does apply to the laws of mechanics but does not apply to electrodynamics. We are therefore forced to choose between the following alternatives:

(a). A principle of relativity exists for mechanics, but not for electrodynamics. A preferred inertial frame (ether frame) exists in electrodynamics.

(b). A principle of relativity exists for both mechanics and electrodynamics, but electrodynamics is not correct in the Maxwell formulation.

(c). A principle of relativity exists for both mechanics and electrodynamics, but the laws of mechanics in the Newtonian form need modification.

The choice between these alternatives can only be made on the basis of experimental results. We shall see by analysis of the relevant experiments that alternative (c), in the form of the special theory of relativity, is essentially correct.

The experiments in question essentially fall into three classes:

(a). Attempts to locate a preferred inertial frame for the laws of electrodynamics.

(b). Attempts to obtain deviations from the laws of classical electrodynamics.

(c). Attempts to observe deviations from classical mechanics.

A reference list of relevant experiments is given below:

REFERENCES ON THE EXPERIMENTAL BASIS OF THE SPECIAL THEORY OF RELATIVITY

1. Miller; Rev. Mod. Phy. 5, 203, 1933.
Discusses in detail Miller's small positive effect on the ether drag experiment.

2. Kennedy; Proc. Nat. Acad. Sci., 12, 621, 1926.
Michelson-Morley experiment using step mirror technique. Null result to ± 2 km/sec.

3. Illingworth; Phys. Rev., 30, 692, 1927.
Repetition of Ref. 2. Null results to ± 1 km/sec.

4. Kennedy and Thorndike; Phys. Rev., 42. 400, 1932.
Interferometer with unequal arms. Null result to ± 10 km/sec.

5. Comstock; Phys. Rev., 30, 267, 1910.
DeSitter, Proc. Amst. Acad., 15, 1297, 1913.
16, 395, 1913.
Spectroscopic work on binary stars.

6. Tolman; Phys. Rev., 31, 26, 1910.
Thomson, J. J.; Phil. Mag., 19, 301, 1910.
Stewart; Phys. Rev., 32, 418, 1911.
Emission theories; new source theory and ballistic theory.

7. Ritz; Ann. de Chim. et. Phys., 13, 145, 1908
Original source emission theory.
See also: Tolman; Phys. Rev., 31, 26, 1910
35, 136, 1912.

8. Majorana; Phil. Mag., 35, 163, 1918.
37, 145, 1919.
Moving source and mirror experiments.

9. Miller; Proc. Nat. Acad., 11, 306, 1925.
M. M. interferometer using light from sun.

10. Kennedy; Phys. Rev., 47, 965, 1935.
Critical discussion of geometrical effects of high order in Michelson-Morley experiment.

11. H. E. Ives; J. Opt. Soc. Am, 215 (1938). - Doppler Shift

This list covers only some of the basic early work. For further references see:

1. Tolman: "Relativity, Thermodynamics and Cosmology."

2. Pauli: "Relativitatstheorie," Encyclopaedia der Mathematischen Naturwissenschaften, Vol. V (special reprint).

3. Bergmann: "Introduction to the Theory of Relativity."

Let us first consider some of the experiments which have a bearing on the question of an absolute ether frame. In Chapter 22 we have already discussed the experiment of Trouton and Noble on measuring the torque on a suspended charged condenser. Such a torque was predicted on the basis that the wave equation, and therefore the retarded potential solutions, are valid only in a given ether frame, and assuming that the earth's orbital motion would assure that the earth would not always move with this ether frame. However, when the experiment was performed, a null result was obtained.

Another attempt to localize the ether frame was the Michelson-Morley experiment shown in Figure (26.1). Light from a source L is split into two paths by a half-silvered mirror at P. The split beams are reflected at mirrors S_1 and S_2 respectively and return through the half-silvered mirror to an observing telescope at F. Interference fringes are

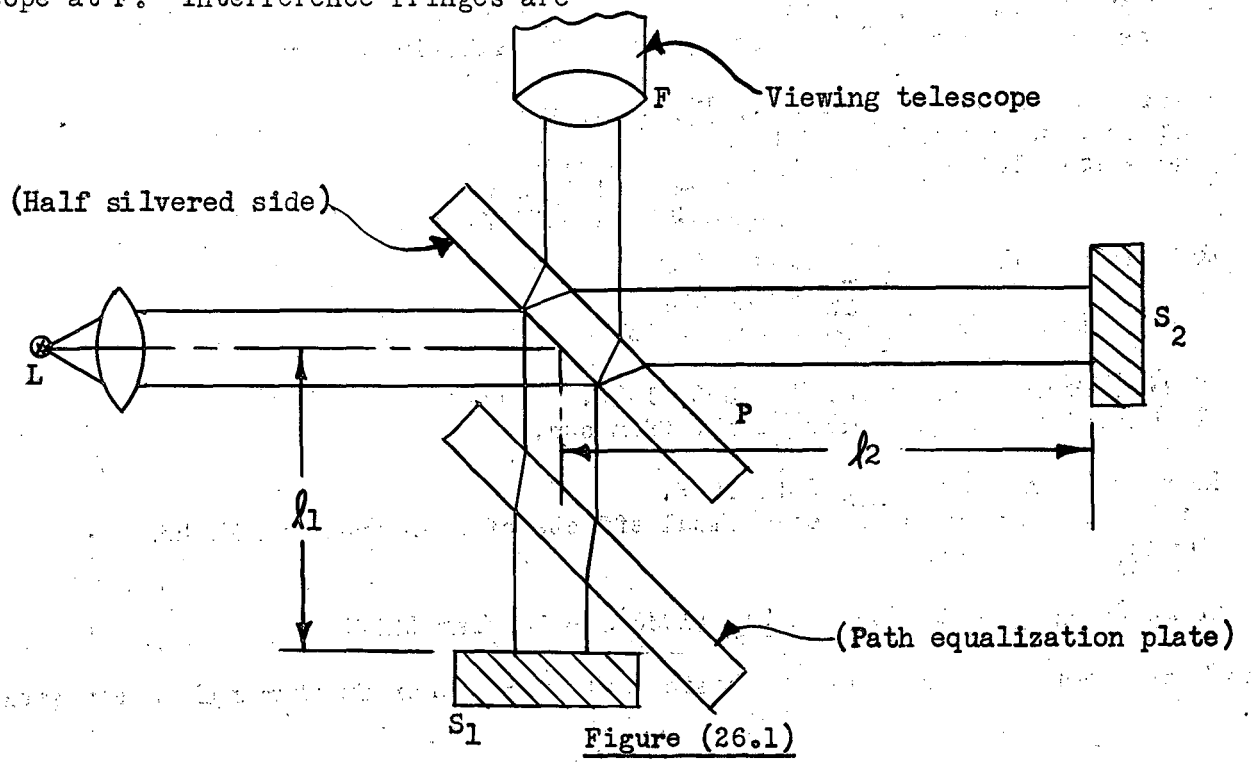


Figure (26.1)

observed at F. Let us assume that the instrument is moving with a velocity $\vec{v}$ parallel to S_1P relative to a stationary ether. Classically the time for the beam PS_1P to complete its passage is:

$$t_{S_1} = \ell_1\left[\frac{1}{c-v} + \frac{1}{c+v}\right] = \frac{2\ell_1/c}{1-\beta^2} \tag{26.4}$$

The time for the beam PS_2P must be corrected for the fact that P will have moved by a distance δ (see Fig. 26.2) during the time light travels from P to S_2; δ is given by:

$$\frac{\delta}{\sqrt{\delta^2 + \ell_2^2}} = \beta; \qquad \delta = \frac{\beta\,\ell_2}{\sqrt{1-\beta^2}} \tag{26.5}$$

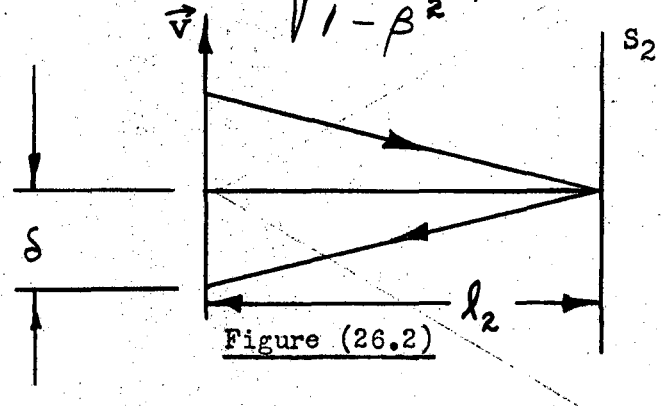


Figure (26.2)

and hence:

$$t_{S_2} = \frac{2}{c}\sqrt{\ell_2^2 + \delta^2} = \frac{2\ell_2/c}{\sqrt{1-\beta^2}} \tag{26.6}$$

The difference Δ in optical path is therefore:

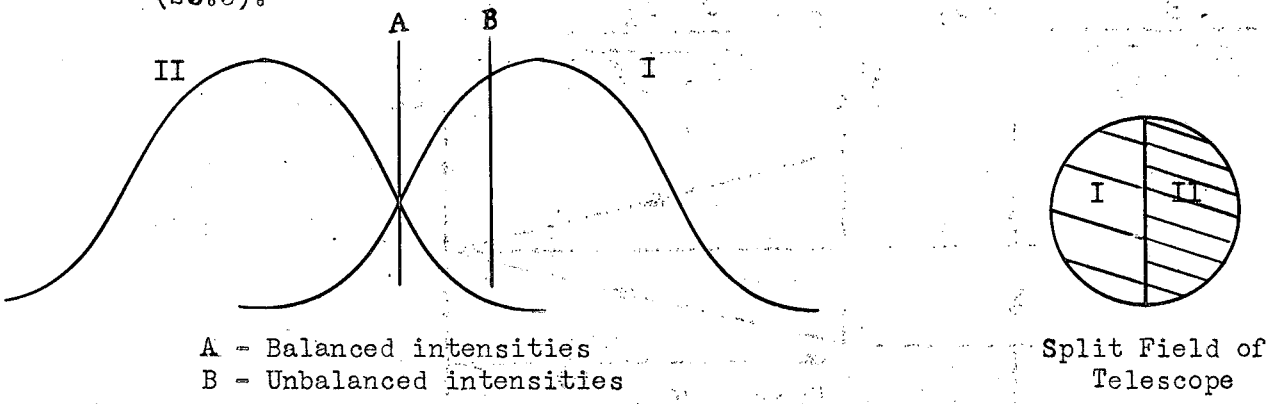
$$\Delta = \frac{2}{\sqrt{1-\beta^2}}\left[\frac{\ell_1}{\sqrt{1-\beta^2}} - \ell_2\right] \tag{26.7}$$

If the instrument is rotated through 90° (i.e. ℓ_1 and ℓ_2 interchanged), the number of fringes N that are expected to shift is approximately:

$$N = \frac{(\ell_1 + \ell_2)}{\lambda}\beta^2 \tag{26.8}$$

The experiment was originally performed by Michelson and Morley with a null result accurate to $\pm$ 10 km/sec. That is the velocity of the earth relative to

any "ether frame" must be less than $\pm$ 10 km/sec. Later experiments by Miller[1] gave a positive indication of a velocity of 10 km/sec. (which is less than the earth's velocity of 30 km/sec. in its orbit!) apparently directed toward a certain point in space. The accuracy of the experiment was greatly improved by Kennedy[2] and Illingworth[3], using a step mirror technique: the mirror S_1 is divided into two halves of slightly different thickness; the telescope is focussed on the mirror. The field then has two halves whose intensity corresponds to two slightly displaced interference patterns, as seen in Figure (26.3).



A - Balanced intensities
B - Unbalanced intensities

Split Field of Telescope

Figure (26.3)

The field is balanced for fringe position A, but not for B; since the match occurs on the steep sides of the intensity curves, much greater sensitivity ($\sim$1/1000 fringe) was attained. Kennedy obtained a null result to $\pm$ 2 km/sec. and Illingworth to $\pm$ 1 km/sec. We can therefore accept the null result with confidence.

The search for a preferred frame for electrodynamics without further modification of electrodynamics or mechanics therefore appears to have been unsuccessful. An attempt to preserve the concept of the preferred ether frame is contained in the Lorentz-Fitzgerald contraction hypothesis which proposes that motion relative to the stationary ether frame contracts all bodies in the ratio $\sqrt{1-\beta^2}$ in the direction of motion. ℓ_1 in Fig. (26.1) is therefore given by

$\ell_1 = \ell_1^\circ\sqrt{1-\beta^2}$ and $\ell_2 = \ell_2^\circ$, where ℓ_1° is the length of ℓ_1 when at rest relative to the ether. Hence (26.7) becomes:

$$\Delta = \frac{2}{\sqrt{1-\beta^2}}\left[\ell_1^\circ - \ell_2^\circ\right] \tag{26.9}$$

and no fringe shift is obtained on interchanging arms. If $\ell_1^\circ \cong \ell_2^\circ$, as was the case in the experiments mentioned, no fringe shift occurs as a function of velocity. If $\ell_1^\circ \neq \ell_2^\circ$, then even with the Lorentz contractions a fringe shift of:

$$N = \left[\frac{\ell_1^\circ - \ell_2^\circ}{\lambda}\right](\beta_2^2 - \beta_1^2) \tag{26.10}$$

is expected from the velocity change due to the term $(\beta_2^2 - \beta_1^2)$. Kennedy [4] constructed an interferometer using a path difference essentially as long as coherence of the source permitted. The square of the velocity of the instrument is presumably given by:

$$c^2\beta^2 = \left[\begin{pmatrix}\text{velocity of earth in}\\ \text{orbit around sun}\end{pmatrix} + \begin{pmatrix}\text{surface velocity}\\ \text{of earth}\end{pmatrix} + \begin{pmatrix}\text{velocity}\\ \text{of sun}\end{pmatrix}\right]^2 \tag{26.11}$$

$$c^2\beta^2 = (\quad v_E \quad + \quad v_R \quad + \quad v_S \quad)^2$$

which should change by:

$$\Delta v^2_{(12\ \text{hr})} = 4\,(v_S + v_E)v_R \tag{26.12}$$

every 12 hours and by:

$$\Delta v^2_{(6\ \text{mo})} = 4\,(v_S + v_R)v_E \tag{26.13}$$

every six months. Neither effect was observed, in contradiction to the Lorentz contraction hypothesis.

A further alternative in which the concept of the ether could be preserved would be to consider the ether frame as attached to ponderable bodies. This would automatically give a null result for terrestrial interferometer experiments. However, the assumption of a local ether is in direct contradiction to the following well-established experimental facts:

(a). The aberration of "fixed" stars.

(b). The Fizeau convection coefficient (Eq.19.12) agrees with a convection of the $\partial \vec{P}/\partial t$ term of the magnetic field sources only, but the $k_0 \frac{\partial \vec{E}}{\partial t}$ term remains unchanged. If the ether moved with the medium, the velocity would simply become $c \pm u$. Note that the idea of an ether attached to ponderable bodies leads to discrepancies in the first order in β.

These considerations appear to make the idea of a preferred frame unacceptable even if it is attempted to have such a frame not generally stationary but only locally stationary. We are therefore led to the conclusion that a principle of relativity is also valid in electrodynamics; this is possible only by either modifying the basic postulates of electrodynamics or the bases of mechanics. The types of modification of electrodynamics considered are the so-called emission theories in which the velocity of a light wave remains associated with the source rather than with a local or universal frame. Such an idea does not admit a field description since the effects at a field point cannot be made independent of the source. As another alternative, we can consider the special theory of relativity which modifies the kinematic description of mechanics such that a principle of relativity applies to both mechanics and electrodynamics. Let us tabulate these comparative ideas.

Theory / Properties	Emission Theory	Classical Ether Theory	Special Theory of Relativity
Reference System	No reference system.	Stationary ether as reference system.	No reference system.
Velocity Dependence of Light	Velocity of light depends on motion of source	Velocity of light is independent of motion of source.	Velocity of light is independent of motion of source.
Space-time Connection	Space and time are independent.	Space and time are independent.	Space and time are interdependent.

Table (26.1)

Let us briefly discuss the emission theories. All emission theories pro-

pose that the velocity of light remains c relative to the original source independent of the state of the transmitting medium. The only difference between the emission theories is the change in velocity which would occur on reflection from moving mirrors. Three alternatives exist here:

1. The velocity remains c relative to the original source.*

---

* This theory, proposed by Ritz[7], was developed to a fair degree of completeness. Ritz retained the two Maxwell equations:

$$\vec{\nabla} \cdot \vec{B} = 0$$

$$\vec{\nabla} \times \vec{E} = - \frac{\partial \vec{B}}{\partial t}$$

with $\vec{E}$ and $\vec{B}$ derived from $\vec{B} = \vec{\nabla} \times \vec{A}$

$$\vec{E} = - \vec{\nabla} \phi - \frac{\partial \vec{A}}{\partial t}$$

while the other two equations (the source equations) are replaced by the requirement that

$$\phi = \frac{1}{4\pi k_o} \int \frac{[\rho]_{[t - r/(c+v)]}}{r} \, dv$$

where v is the velocity of the source relative to the ether. This equation replaces the ordinary retarded potential solution. In this way fields due to a moving source are definable.

---

2. The velocity becomes c relative to the last mirror from which reflection occurred.

3. The velocity relative to its mirror image remains c.

The first theory was proposed by Ritz and is the only theory which does not lead to coherence difficulties concerning the reflected light. These three theories will give differences in the first order in β in experiments on the interference between light beams reflected from moving mirrors such as those of Thompson, Majorana[8], and Stewart. All these experiments gave results in disagreement with

the theories 2 and 3, but in agreement, within experimental error, with the Ritz theory. The reason for this agreement is that in any closed system of interfering beams, the Ritz theory will differ only by terms of the second order in β from the results obtained at constant velocity. If, e.g., light travels from a stationary source to a mirror moving with velocity v and returns, then the effective velocity remains c, as it does relativistically.

If the source is moving with velocity v, then using the original source theory, the time required for light to cover a given distance ℓ on the forward trip is $\frac{\ell}{c-v}$ while on a return trip after reflection from a mirror, it is $\frac{\ell}{c+v}$; combined, these give a time:

$$\Delta t = \frac{2\ell}{c}\left(\frac{1}{1-\beta^2}\right) \tag{26.14}$$

which differs only by terms in the second order from the constant light velocity expression $\Delta t = \frac{2\ell}{c}$. Hence any terrestrial moving source and mirror experiments fail to give a first order contradiction to the Ritz emission theory.

There are, however, two extraterrestrial experiments which contradict any form of emission theory. One of these is the observed dynamics of eclipsing binary stars by DeSitter. If light, emitted by a star from the approaching and receding leg of the orbit, travelled with different velocity, then the time interval observed between successive eclipses of binary stars would become highly asymmetrical. Actually no such effect is observed; in fact DeSitter concluded that if $v_{light} = c + k\, v_{star}$ then $k \ll 0.002$. The second extraterrestrial evidence is the experience of Miller that the Michelson-Morley experiment does not exhibit any change in result when light from the sun is used instead of terrestrial sources.

This outline of the experimental basis shows that experiment contradicts any reasonable alternative to the theory of relativity, rather than any single experiment proving the theory. The experimental tests are summarized in Table

(26.2). For the sake of completeness, experiments are included which have not been discussed here and whose bearing on the subject will not become evident until the theory of relativity has been discussed in more detail.

EXPERIMENTAL EVIDENCE OF THE SPECIAL THEORY OF RELATIVITY

Table (26.2)

Theory \ Experiments	Light Propagation Experiments								Experiments from Other Fields					
	Aberration	Fizeau Convection Coefficient	Michelson-Morley	Ives	Kennedy-Thorndike	Moving Sources and Mirrors	DeSitter Spectroscopic Binaries	Michelson-Morley with light from sun	Variation of Mass with Velocity	General Mass-energy Equivalence	Uni-polar induct. with perm. magnet	Radiation from moving charges	Meson Decay at High Velocity	Trouton-Noble
Stationary ether, no contraction	√	√	N	N	N	√	√	N	N	O	N	√	O	N
Lorentz contr. and stationary ether	√	√	√	N	N	√	√	√	√	O	N	√	O	√
Ether attached to ponderable bodies	N	N	√	N	√	√	√	√	N	O	O	O	O	√
Original Source	√	√	√	N	√	√	N	N	O	O	O	N	O	O
Ballistic	√	O	√	N	√	N	N	N	O	O	O	N	O	O
New Source	√	O	√	N	√	N	N	√	O	O	O	N	O	O
Special Theory of Relativity	√	√	√	√	√	√	√	√	√	√	√	√	√	√

√ = agrees
N = contradicts
O = does not apply

The experiments outlined above present evidence that:

(1). The presence of an ether, either stationary or convectively carried, cannot be established.

(2). Modification of electrodynamics of the emission theory type is untenable.

The conclusions then make it plausible to look upon the basic laws of mechanics as in need of modification.

In 1905 Einstein proposed as a solution, compatible with the experiment facts known at that time, the following postulates:

(1). All laws of electrodynamics (including, of course, propagation of light with the velocity c in free space) shall be the same in all inertial frames, as are the laws of mechanics.

(2). It shall be impossible to devise any experiment defining a state of absolute motion or to determine a preferred inertial frame having special properties for any physical phenomena.

It is clear that if the laws of physics obeyed these postulates, all the experimental facts outlined above would be in agreement with these postulates. We must now examine the consequences of these postulates, known as the postulates of special relativity.

CHAPTER 27

RELATIVISTIC KINEMATICS

We have seen in Chapter 26 that much experimental evidence exists that a principle of relativity exists for all fields of physics, including electrodynamics. This implies, that among other things, the velocity of propagation, c, of plane electromagnetic waves in free space must be independent of the particular observer's inertial frame.* This statement, plausible as it may seem,

* The question is often raised why none of the following discussions apply e.g. to the velocity of propagation of sound or any other velocity. The reason is that the velocity of sound does require a material medium for its propagation and therefore does have a preferred reference system.

runs grossly contrary to our intuition. Consider for example a light pulse starting from a point P, and consider this event as recorded by observers stationed in two frames, one frame containing P at the origin, while the other frame moves relative to P with a velocity $\vec{v}$. Let the origins of the two frames coincide at the start of the pulse. According to the statement above, both observers must see the light wave propagating as a spherical wave centered at their respective origins! If we consider the position of the wavefront to be an event permitting description independently in space and in time, then this statement cannot be true. The independence of the velocity of light of the particular frame therefore requires a revision of the accustomed ideas of the possibility of specifying the position coordinates of an event referred to a particular frame, but specifying the time of the event by a "universal" time scale. The above paradox would not exist if there were no such universal time scale, but if the simultaneity of the wavefront passing through two points was an observation which was not independent of the frame of the observer. If such a disagreement as to the simultaneity of time of passage through a set of points was permitted to exist, then presumably a kinematics could be constructed in which a spherical light

wave would be seen in both frames of reference.

Therefore, we are led to re-examine the concept of simultaneity. If we must abandon the existence of a universal time as not corresponding to reality, then we must establish a mechanism whereby simultaneity can be established in a given frame. This mechanism must be such that a measurement of the velocity of light in the particular frame using its time and distance scale must give c. This means that the only way in which simultaneity can be defined is by means of the velocity of light itself. This conclusion gives c a much more fundamental significance than just the velocity of propagation of electromagnetic waves; it introduces c into all the relations of physics. Among other things, the utilization of c as the defining element of simultaneity precludes the existence of the "ideal rigid body" of mechanics; if there were such a body, its ends would move simultaneously as observed by any frame and it would therefore be used as a means of establishing a "universal time," in violation of our former conclusions.

We therefore consider two instants of time t_1 and t_2 observed at two points x_1 and x_2 in a particular frame as simultaneous if:

(a) A light wave omitted at the geometrically measured midpoint between x_1 and x_2 arrives at the time t_1 and x_1 and at the time t_2 at x_2.

Or if:

(b) A light wave emitted at x_1 at the time t_1 arrives at x_2 at a time $t_1 + \frac{x_2 - x_1}{c} = t_2$

The first definition of simultaneity will automatically assure that a light pulse emitted at the origin will reach all equidistant points simultaneously and that the wave surface is therefore a sphere in a particular reference frame. Simultaneity of two events at two spatially separated points therefore does not have a significance independent of the frame. The relation of the time intervals

observed by two different frames is therefore dependent on the spatial interval between the events; the Galilean transformation Eq. (26.2) which transformed temporal intervals as observed by two frames independently of spatial coordinates therefore cannot be in agreement with the simultaneity definition in terms of c. We must therefore attempt to derive the corresponding transformations from an $(x,y,z;t)$ frame to a $(x',y',z';t')$ frame which will supersede the Galilean transformation. Such a transformation must remain linear, to assure mathematical equivalence of all points in space and time, but the spatial and temporal coordinates need not transform independently.

The desired transformation gives the relation between the space-time coordinates of an arbitrary event $(x,y,z;t)$ as observed in the $\sum$ frame and the space-time coordinates $(x',y',z';t')$ as observed in the $\sum'$ frame. This transformation must obey the postulates of special relativity (see Chap. 26) for an event of any type. We shall therefore construct a set of "Gedanken-experimente"* each of which will incorporate only one additional feature of the transformation in order to indicate how the basic postulates of relativity necessitate the nature of the transformation.

* Thought experiments.

Experiment I - Comparison of parallel measuring sticks oriented perpendicular to their direction of relative motion.

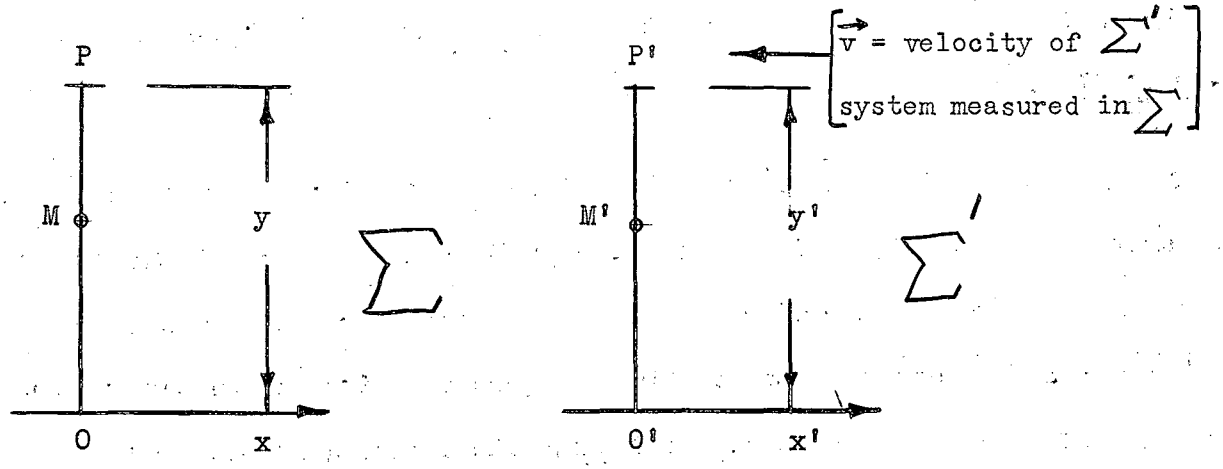


Figure (27.1)

Let us assume that it is possible to bring the frames Σ and Σ' of Figure (27.1) to relative rest and to adjust the two measuring sticks to equality. It is assumed explicity that the properties of a given body of specified structure are independent of its past history when observed in a frame where that body is at rest (called the proper frame)*, it is therefore not essential whether the adjustment of lengths referred to is possible or not; it could for instance be specified that the length of each rod should be a given number of wavelengths of a specified spectral line measured in each frame.

* The length of a rod when measured in a frame in which the rod is at rest is called its "proper length".

Let the two systems approach each other such that the midpoints M and M' coincide. Let light signals be sent from O and P at the time when O and P coincide with the y' axis. Since OM' remains equal to PM' during the motion O and P will appear to cross the y' axis simultaneously for both systems, and similarly O' and P' will cross the y axis simultaneously in both systems. We therefore conclude that along a direction perpendicular to the direction of relative motion simultaneity will be the same in both systems. Both observers can therefore compare the positions of the end markers at time of cross-over and arrive at the same result since the time of observation for both ends is defined identically in both systems. Hence both observers would conclude either $OP \geqslant O'P'$, or $O'P' \geqslant OP$; since both systems are fully equivalent as to their state of motion, an asymmetric solution would provide a means of determining absolute velocity, which is ruled out by the postulates. We therefore put:

$$y' = y \tag{27.1}$$

and similarly:

$$z' = z \tag{27.2}$$

Experiment II - Comparison of Clock Rates

In comparing clock rates among moving systems, we are faced with a fundamental difficulty: it is impossible to compare one clock in $\sum$ with one clock in $\sum'$ since they will not stay in coincidence; we must compare two clocks in $\sum$ with one clock in $\sum'$ and synchronize the two clocks in $\sum$ by light signals. Consider that in $\sum'$ at the clock position a light signal is emitted normal to $\vec{v}$ and reflected at a mirror normal to the z' axis at a distance z' from the clock and returned to the clock. (Figure 27.2)

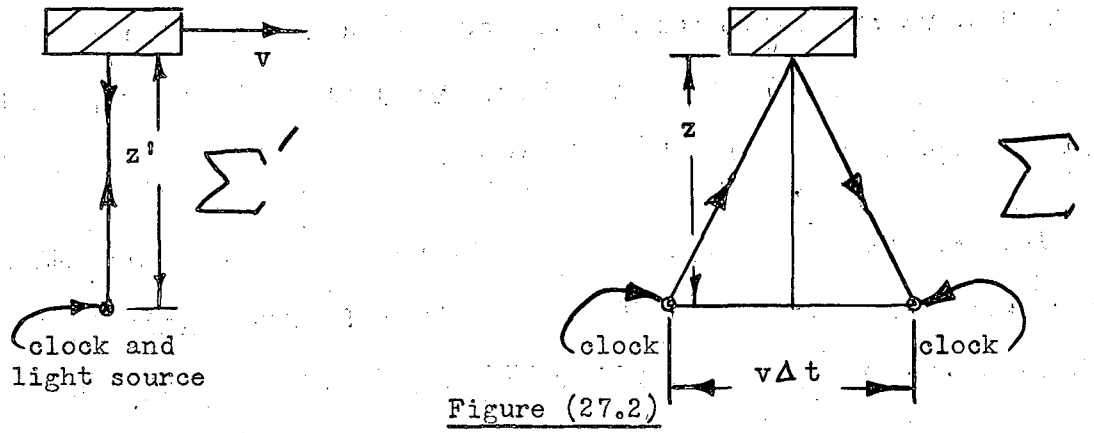


Figure (27.2)

$\sum'$ will define the time interval between sending and receiving of the pulse as:

$$\Delta t' = \frac{2z'}{c} \qquad (27.3)$$

$\sum$ can record the time interval Δt between the same events with two clocks spaced $v\Delta t$ apart. $\sum$ concludes therefore, since c is independent of frame:

$$c\Delta t = 2\sqrt{(z)^2 + \left(\frac{v\Delta t}{2}\right)^2} \qquad (27.4)$$

$$\Delta t = \frac{2z}{c}\,\frac{1}{\sqrt{1-\beta^2}} \qquad (27.5)$$

or, from (27.1) and (27.3):

$$\Delta t = \frac{\Delta t'}{\sqrt{1-\beta^2}} \qquad (27.6)$$

($\Delta t'$ is the proper time interval)

Note that the apparent asymmetry causes no paradox, since this is not a symmetrical situation. $\Delta t'$ is the time interval between two events occurring at the same

place in the Σ' frame. $\Delta t'$ is called the "proper time interval" between the two events. Δt on the other hand, is not a proper interval since it is measured by two clocks at different places. A proper time interval, similar to the case of a proper length discussed above, is a definite function of the physical nature of the clock; e.g. a particular radioactive decay constant or the natural frequency of a crystal of specified proper dimensions is a constant in the frame where such time intervals are observable at a single point; i.e. in a frame where such a "clock" is at rest.

Note therefore that Σ will find that his (not proper) time interval is longer than the proper time interval, measured in Σ'. This phenomenon is known as time dilation.* Σ' cannot judge the rate of the Σ clocks, but as we shall see later would consider them out of synchronism, which is reasonable since simultaneity between two events displaced along the x axis is not an invariant* property.

---

* The lifetime of a high velocity meson disintegrating in flight appears lengthened to a ground observer. The lifetime in the proper (i.e. the meson's) frame is invariant.

---

*We shall use the word "invariant" to mean "independent of choice of inertial frame".

Experiment III - Comparison of Lengths Parallel to the Direction of Motion

Let us consider a rod which has a length x' in the frame Σ' in which it is at rest; i.e., let its proper length be x'. In the Σ frame its length x would be the distance between the ends of the rod measured "simultaneously" in Σ in the sense of the simultaneity definition in terms of c. To separate the length comparison from the simultaneity calculation, let us consider the following event: Let a light source S' at one end of the rod send a light pulse to a

mirror M' at the other end where it is reflected back to the source,

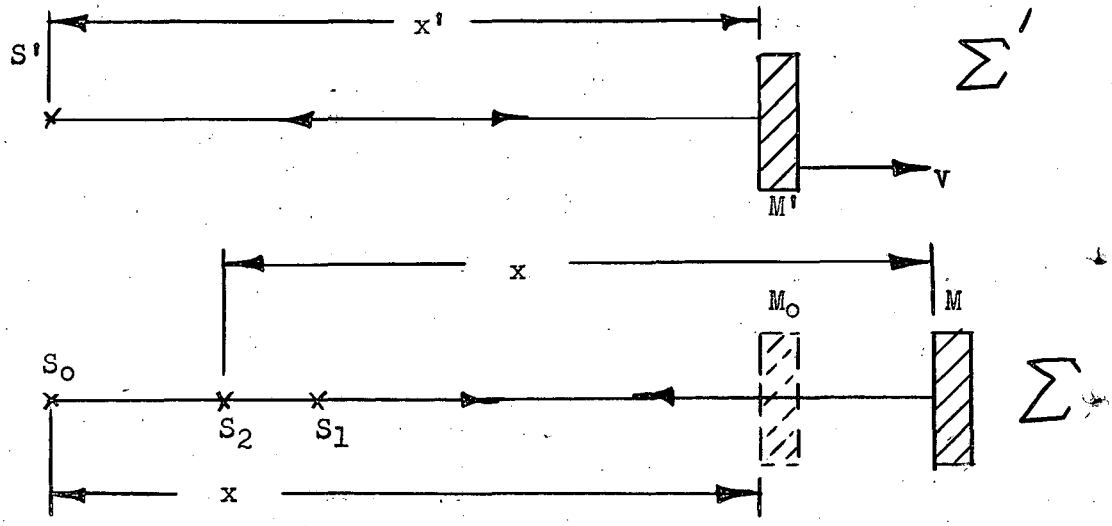


Figure (27.3)

Let $\Delta t'$ be the time interval between the time of emission and the time of arrival of the signal. Note that $\Delta t'$ is a proper time interval, being observable with a single clock at one point. Evidently:

$$(27.7) \qquad \Delta t' = \frac{2x'}{c}$$

In Σ these same events appear to be more complicated. At the time of emission the source S' was at S_0 and the mirror M' at M_0. (This statement has only unique meaning when referred to Σ'.) At the time of reflection, the mirror M' has moved to M and the pulse returns to the source S' when it is at S_1. The time interval Δt is thus measured between the points S_0 and S_1 with two clocks (Δt is not proper) as in Experiment II. Equation (27.6) therefore applies here also. By definition, we mean by x the distance S_0M_0. Since M_0 has moved to M with velocity v while the light moved from S_0 to M with velocity c, we have:

$$(27.8) \qquad S_0M = x + \frac{v}{c}\, S_0M, \qquad S_0M = \frac{x}{1-\beta}$$

and similarly since the source has moved from S_2 to S_1 with velocity v (where $S_2M = x$) while the light has traveled from M to S_1 with velocity c:

(27.9) $\qquad MS_1 = x - \frac{v}{c} MS_1 \qquad\qquad MS_1 = \frac{x}{1+\beta}$

Hence:

(27.10) $\qquad \Delta t = \frac{S_oM + MS_1}{c} = \frac{2x}{c(1-\beta^2)}$

From (27.6) and (27.7), this becomes:

(27.11) $\qquad x = x'\sqrt{1-\beta^2}$

This relation, called the Lorentz contraction, again is asymmetrical in x and x', since it gives the relation between measurement of a proper length x' (at rest) in $\sum'$ and an improper length x not at rest in $\sum$. The length x in $\sum$ was definable only by assuming the constancy of the velocity of light.

Experiment IV - The Synchronization of Clocks

By proper design of Experiments II and III, we have been able to derive the transformation of temporal and spatial intervals from proper to non-proper frames without calculating explicitly the error in synchronization δ (which $\sum$ would conclude to exist) of two clocks separated by a distance x' in $\sum'$ which appear synchronized in $\sum'$.

Consider two clocks synchronized in $\sum'$ and located a distance x' apart, as seen in Figure (27.4). Let there be a single clock in $\sum$ which will record the times, t_o and t_1, when it passes the ends of x'; the corresponding times in $\sum'$ are recorded on the two clocks in $\sum'$ and are denoted by t_o' and t_1'.

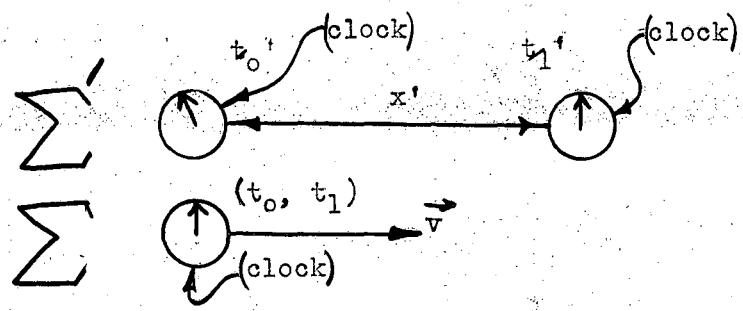


Figure (27.4)

Since $t_1 - t_o$ is a proper time interval in $\sum$, we can apply (27.6) in reverse, giving:

$$(27.12) \qquad t_1' - t_o' = \frac{t_1 - t_o}{\sqrt{1 - \beta^2}}$$

The observer on $\sum$ could apply the dilation equation (27.6) to the rates of the individual clocks at the ends of x' but would conclude that they are out of step by an amount δ; i.e.;

$$(27.13) \qquad t_1 - t_o = \frac{t_1' - t_o' + \delta}{\sqrt{1 - \beta^2}}$$

Both observers have to obtain the same value for the relative velocity v between the frames since if one obtained either a larger or smaller value, the frames would not be equivalent. Hence:

$$(27.14) \qquad t_1' - t_o' = \frac{x'}{v}$$

$$(27.15) \qquad t_1 - t_o = \frac{x}{v}$$

but since x' is a proper length in $\sum'$, we obtain:

$$(27.11) \qquad x = x' \sqrt{1 - \beta^2}$$

combining (27.13), (27.14), (27.15) and (27.11), we have:

$$(27.18) \qquad \delta = - \frac{x' \beta^2}{v}$$

The negative sign indicates that in the opinion of $\sum$ the leading clock (t_o') in $\sum'$ should have indicated a larger time*, in order to make the elapsed time between t_o' and t_1' as observed by $\sum$ smaller.

---

* The clock that is ahead (i.e. met first) is behind (in time).

These four experiments have thus demonstrated four kinematic relations:

I. Distances transverse to the direction of motion are invariant.

II. A time interval Δt measured in a frame moving with velocity $\pm v$ relative to a frame in which the time interval $\Delta\tau$ between two events is proper (i.e. the two events occur at one place) is given by:

$$\Delta t = \frac{\Delta\tau}{\sqrt{1-\beta^2}} \qquad (27.19)$$

III. The length Δx of a rod measured in a frame moving with velocity $\pm v$ relative to a frame in which the rod is at rest and has the proper length $\Delta\lambda$ is given by:

$$\Delta x = \Delta\lambda\sqrt{1-\beta^2} \qquad (27.20)$$

IV. Two clocks, synchronous in a given frame, and separated by a distance $\Delta\lambda$ in that frame appear to be out of synchronism as observed by a frame moving with a relative velocity -v to the clock frame by an amount given by:

$$\delta = -\frac{\Delta\lambda v}{c^2} \qquad (27.21)^*$$

---

* See footnote on previous page.

CHAPTER 28

THE LORENTZ TRANSFORMATION

In Chapter 27 we derived a series of special kinematic effects from the fundamental postulates of special relativity, each effect applied to a given set of conditions arranged such that only one effect applied. We shall now combine these "effects" to give the general relation between the time and space coordinates of a particular event as observed from inertial frames in relative motion.

Consider a point event at a point P moving with the Σ' frame, as seen in Figure (28.1). Let this event occur at time t' in Σ' and let the coordinates of P in Σ' be (x', y', z'). Now consider this same event observed from a frame Σ in motion relative to Σ' with a velocity $|\vec{v}|$. Let us choose the x and x' axis

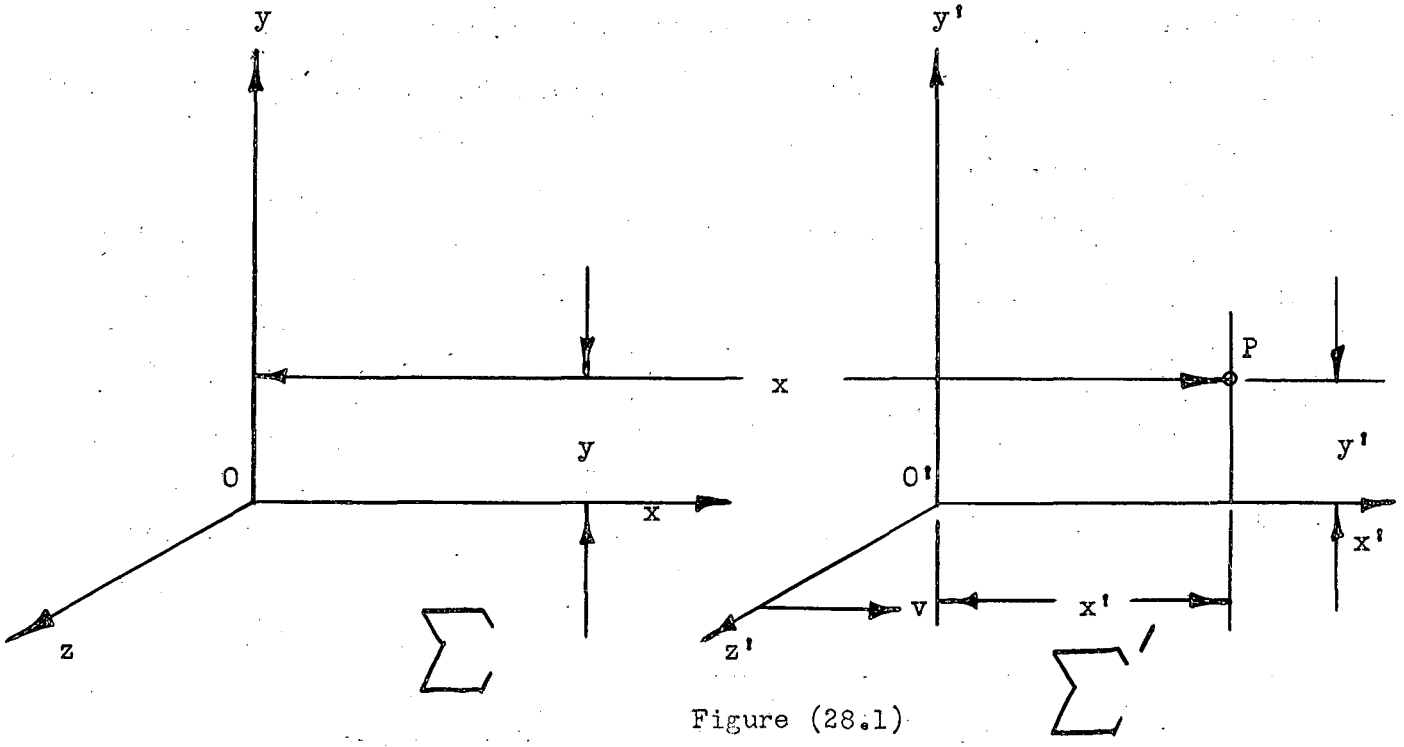


Figure (28.1)

as the direction of relative motion and let the origins and the zero point of time be so chosen that at $t = t' = 0$, the two origins coincide. By the time t or t' respectively, we mean therefore the time elapsed as measured by an observer in Σ or Σ' respectively since the coincidence of origins. Let v be positive if the origin of Σ' moves along the positive x direction in Σ.

From the Σ point of view, $OO' = vt$, but x' being a proper length in Σ' is shortened by the Lorentz contraction. Hence:

$$x = vt + x'\sqrt{1-\beta^2} \tag{28.1}$$

and hence:

$$x' = \frac{x - vt}{\sqrt{1-\beta^2}} \tag{28.2}$$

Again from the Σ point of view, clocks located at P and O' (and thought by an observer in Σ' to be synchronized) are out of synchronism by an amount:

$$\Delta t' = \frac{x'v}{c^2} \tag{28.3}$$

The clocks at O and O' were synchronous at $t = t' = 0$. Since that time, according to Σ, the Σ' clocks have been running at a rate which must be dilated by $1/\sqrt{1-\beta^2}$ to make it equal to the rate of the clock on Σ. Combining these two effects, we have:

$$t = \frac{t' + x'v/c^2}{\sqrt{1-\beta^2}} \tag{28.4}$$

This, by the use of Eq. (28.2), can be reduced to:

$$t' = \frac{t - xv/c^2}{\sqrt{1-\beta^2}} \tag{28.5}$$

Excepting for the sign of $\vec{v}$, Σ and Σ' are equivalent, in agreement with the second postulate. We can also show from (28.4) and (28.2) that:

$$x = \frac{x' + vt'}{\sqrt{1-\beta^2}} \tag{28.6}$$

again in agreement with (28.2) except for the sign of $\vec{v}$. It also follows from the first "experiment" of Chapter 27 that:

$$y = y' \tag{28.7}$$

$$z = z' \tag{28.8}$$

Equations (28.2) to (28.8) are the general transformations desired, subject

to the restrictions as to choice of origin and orientation of axes as given above. The classical Galilean transformations corresponding to these relations would be:

(28.9) $x' = x - vt$

(28.10) $y' = y$

(28.11) $z' = z$

(28.12) $t' = t$

Let us examine some of the comparative properties of the Lorentz and Galilean transformations. We can easily show algebraically that if the Lorentz transformation is valid:

(28.13) $x^2 + y^2 + z^2 - c^2t^2 = x'^2 + y'^2 + z'^2 - c^2t'^2$

This means that if a light signal is propagated in all directions with velocity c from O at t = 0 as observed by Σ, then a light signal is propagated from O' in all directions with velocity c at t' = 0 as observed by Σ'. The transformations are therefore in agreement with the first postulate and resolve the apparent paradox mentioned earlier in Chapter 27. For the Galilean transformations, (28.9) to (28.12), Eq. (28.13) is not true; in its place, however:

(28.14) $(x_1' - x_2')^2 + (y_1' - y_2')^2 + (z_1' - z_2')^2 = (x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2$

$$(t_1' - t_2') = (t_1 - t_2)$$

which shows that <u>independently</u> the spatial interval and the temporal interval between two events $(x_1, y_1, z_1; t_1)$ and $(x_2, y_2, z_2; t_2)$ are invariant, while in special relativity the <u>combined</u> space-time interval:

(28.15) $(x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2 - c^2(t_1 - t_2)^2$

is invariant. For a differential interval between two events, the quantity:

(28.16) $ds^2 = c^2dt^2 - dx^2 - dy^2 - dz^2$

is therefore invariant. If in a given frame:

(28.17) $dx^2 + dy^2 + dz^2 < c^2dt^2$

then a Lorentz frame can be found in which the spatial part of ds is zero; i.e. where the two events occur at the same place; ds = cdt is then c times the "proper" time interval as defined before. We therefore conclude if the "space-time" interval ds fulfills the inequality (28.17) (time-like interval), then ds/c represents the proper time interval between the events. Conversely, if in a given frame:-

(28.18) $dx^2 + dy^2 + dz^2 > c^2dt^2$

then a frame can be found in which dt = 0; in that frame the two events are simultaneous and ids is then their mutal distance. Hence if ds fulfills (28.18) (space-like interval) then ids represents the proper length of the increment. No Lorentz transformation with real $\vec{\beta}$ can reverse the sign of the inequalities (28.17) and (28.18), so that the physical significance of ds; either being a proper time interval (times c) or a proper length interval respectively, depends uniquely on the nature of the interval described.

It can be easily shown algebraically that two successive Lorentz transformations with velocity parameters β_1 and β_2 are equivalent to a single Lorentz transformation of parameter:

(28.19) $$\beta = \frac{\beta_1 + \beta_2}{1 + \beta_1\beta_2}$$

Lorentz transformations therefore form a mathematical "group".

It is possible to obtain the Lorentz transformation equations in several ways using simply the demand that the interval ds in (28.16) be invariant and that the transformations be linear; the second demand arises from the fact that all points in space and time should have identical transformation character as long as only inertial frames are considered. We must also have that if

$x = f(x', t', v)$, then $x' = f(x, t, -v)$. As an example of such a derivation, we assume:

(28.20) $$kx = x' - vt'$$

(28.21) $$kx' = x + vt$$

where k is to be an even function of the velocity. By simple algebra:

(28.22) $$x'^2 - x^2 = v^2(1-k^2)^{-1}\,(t^2-t'^2)$$

To make this agree with the invariant interval (28.13) we must take:

(28.23) $$c^2(1-k^2) = v^2$$

(28.24) $$k = \sqrt{1-\beta^2}$$

giving the Lorentz transformation.

The Lorentz transformation treats x and t as entirely equivalent variables. It was suggested by Minkowski to introduce ct simply as a fourth coordinate. Let us put:

(28.25) $$x^1 = x$$
$$x^2 = y$$
$$x^3 = z$$
$$x^4 = ct$$

as a set of variables in four dimensional space.* The space-time interval

---

* Superscripts rather than subscripts are used here for a reason to be explained later.

(28.16) is therefore:

(28.26) $$ds^2 = -dx^{1^2} - dx^{2^2} - dx^{3^2} + dx^{4^2}$$

The Lorentz transformation is therefore in a general sense the set of linear transformations in 4-dimensional space which leaves ds^2 invariant.

The interval ds^2 can be written in a more familiar but physically less obvious form if we write:

(28.27) $x^4 = ix^4$; $dS = ids$

and hence:

(28.28) $dS^2 = dx^{1^2} + dx^{2^2} + dx^{3^2} + dX^{4^2}$

Since dx^2 and dx^3 are not affected by our special choice of coordinates in the Lorentz transformation, let us consider only the invariance of the two-dimensional sub-interval:

(28.29) $dX^2 = dx^{1^2} + dX^{4^2}$

Other than translations of origin, this interval is also invariant to rotations to an angle θ in the $x^1 - X^4$ plane, (Figure 28.2),

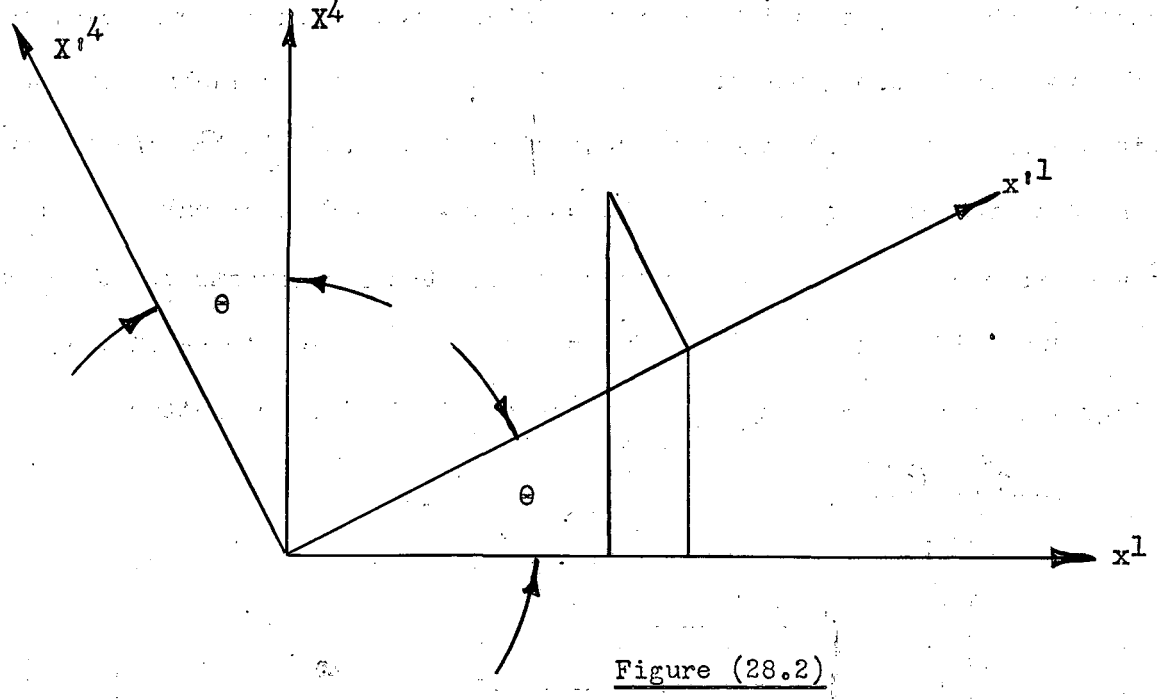


Figure (28.2)

to a new set of axes $x'^1 - X'^4$. By geometry:

(28.30) $x'^1 = x^1 \cos\theta + X^4 \sin\theta$

(28.31) $X'^4 = -x^1 \sin\theta + X^4 \cos\theta$

or from (28.25) and (28.27) and putting $\theta = i\phi$:

(28.32) $x' = x \cosh\phi - ct \sinh\phi$

(28.33) $ct' = -x \sinh\phi + ct \cosh\phi$

This is identical with the Lorentz transformation if we put:

(28.34) $\sinh\phi = \dfrac{\beta}{\sqrt{1-\beta^2}}$; $\cosh\phi = \dfrac{1}{\sqrt{1-\beta^2}}$; $\tanh\phi = \beta$

Hence the Lorentz transformation is simply a **rotation** in the 4-dimensional space x^1, x^2, x^3, X^4. An "event" is therefore conveniently described by the four coordinates in such a space-time system. Temporal and spatial coordinates are equivalent in this notation. Eq. (28.19) which gives the relation between the equivalent β for two successive Lorentz transformations corresponds simply to the addition formula for tanh ϕ.

$$\tanh(\phi_1 + \phi_2) = \frac{\tanh\phi_1 + \tanh\phi_2}{1 + \tanh\phi_1 \tanh\phi_2} \tag{28.35}$$

This representation of the Lorentz transformation as a rotation in the 4-dimensional space x, y, z, ict is a very useful concept but an artificial one since the coordinate along one of the axes is an imaginary variable. Let us investigate what the geometrical representation of the Lorentz transformation is in the real four-dimensional x, y, z, ct space. Let us plot only $x^1 = x$ and $x^4 = ct$ to permit representation in a plane. Under a Lorentz transformation these axes will transform into x'^1 and x'^4, as seen in Figure (28.3). This diagram is called the Minkowski Diagram and in it the trajectory of an

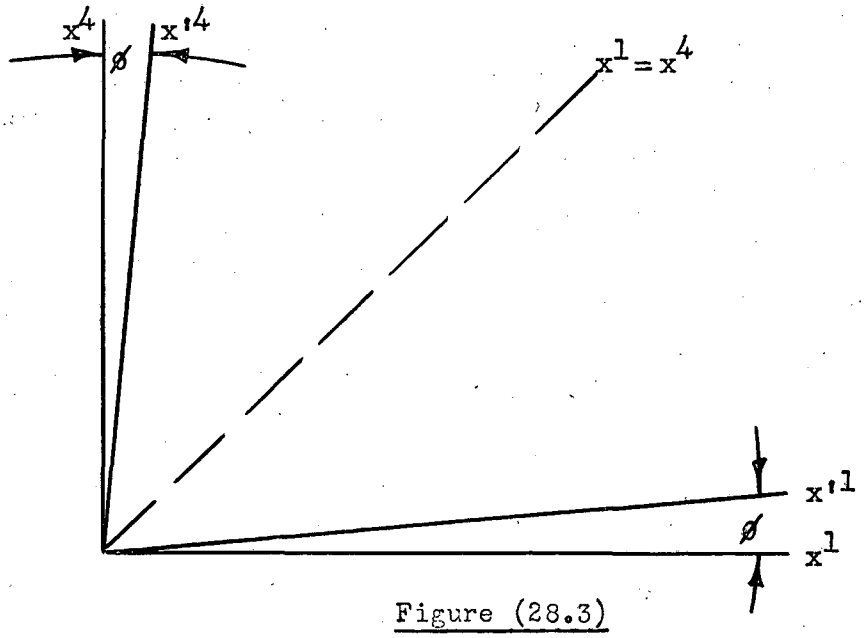


Figure (28.3)

event as a function of space and time is called a world-line.

The diagram must be interpreted with great care. The reason is that

distances on this diagram cannot be measured by the quantities $x^{1^2} + x^{4^2}$ as in Figure (28.2) where we artificially produced a "Euclidean" geometry by the imaginary transformation (28.27). A substitute for a distance measurement in this real space can be obtained however by noting in Figure (28.4) that the family of hyperbolas:

$$x^{1^2} - x^{4^2} = x'^{1^2} - x'^{4^2} = \text{constant} \tag{28.36}$$

lays out a convenient measuring net which permits comparison of the various quantities involved.

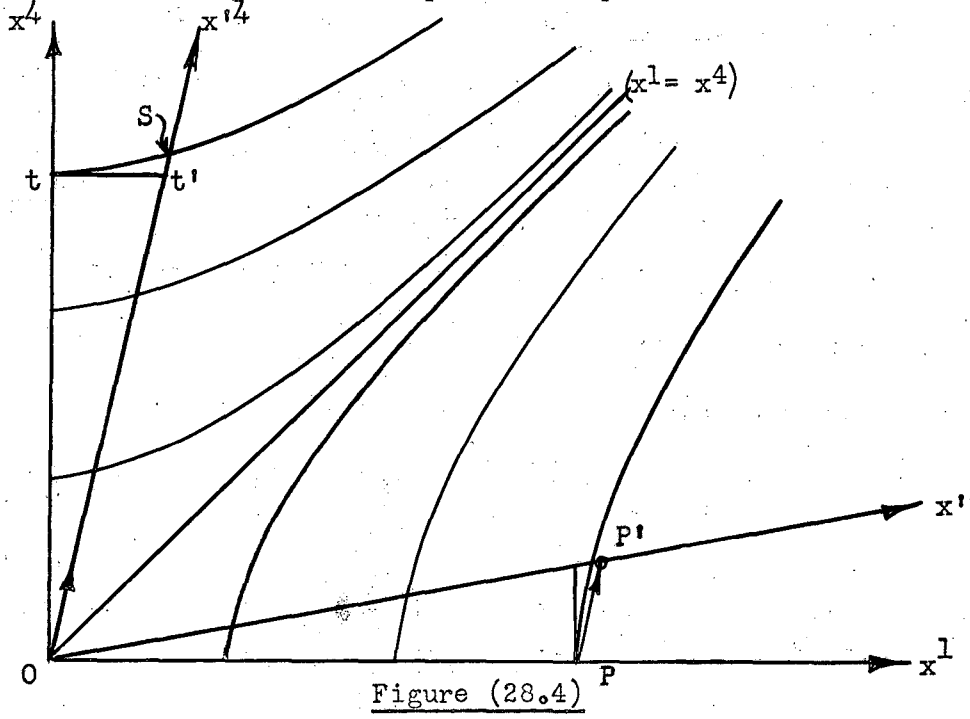


Figure (28.4)

Let us see how we can interpret the various phenomena of relativistic kinematics in this diagram. Consider first the Lorentz-Fitzgerald contraction, Equation (27.20). The Lorentz contraction considers the transformation of a proper length x'^1 in $\sum'$ to the $\sum$ frame. Consider a rod OP' at rest in $\sum'$. The world line of the end point P' on the Minkowski plane will move parallel to the $x'^4(ct')$ axis, from P to P', being at rest (proper) in $\sum'$. Similarly the point 0 will move along the x'^4 axis (see arrows in Figure (28.4). When the length of this rod is measured in $\sum$, the distance between its endpoints is observed when they are simultaneous in $\sum$, i.e., along the x^1 axis. The length of the rod in $\sum$

is thus the length OP. In comparing the lengths OP and OP' we must be careful to refer the measurements to the hyperbolic grid discussed above. It is easily seen that the hyperbola:

$$x^{1^2} - x^{4^2} = OP^2 = x'^{1^2} - x'^{4^2} \tag{28.37}$$

crosses Ox'^1 between O and P' and hence $OP < OP'$ in accordance with the Lorentz contraction of proper lengths observed from a moving frame.

Similarly let us consider the time dilation. Consider a single clock at rest in $\sum'$ at $x' = t' = x = t = 0$. As time progresses the time interval relative to $t' = t = 0$ will be represented by a world line moving along the x'^4 axis; the time interval measured in $\sum'$, since the coincidence of the time origins, is thus Ot'. In $\sum$, O and t' are not at the same spatial point, however, in $\sum$ the point considered simultaneous with t' will be at t where tt' is parallel to the x^1 axis.

Considering that in the hyperbolic "metric" Ot = OS, we find that Ot > Ot'. Hence $\sum$ will observe a longer elapsed time than the proper time interval measured in $\sum'$, in agreement with our former result.

The Minkowski diagram shows the symmetry between the $\sum$ and $\sum'$ frames despite the apparent asymmetry of the time-dilation and Lorentz contraction. In our examples, $\sum'$ was taken to be the proper frame for both spatial and temporal intervals. If $\sum$ had been chosen the proper frame, then we would have $PP' \| Ot$ and $tt' \| OP'$ which would have reversed the contraction and dilation relations.

A Lorentz transformation with $\beta > 1$ becomes complex and thus impossible. We shall later give a more concise formulation to this statement. On the Minkowski diagram, this means that the x'^1 and x'^4 axes cannot pass the cone (light cone) $x^1 = \pm x^4$. This means that a time-like interval $ds^2 > 0$ cannot

become space-like in any frame but can become purely temporal if referred to the proper frame; conversely a space-like interval $ds^2 < 0$ cannot become time-like in any frame but can become purely spatial if referred to a proper frame.

The light cone $x^1 = \pm x^4$ thus divides the Minkowski space into four regions (as seen in Figure 28.5) which have invariant significance:

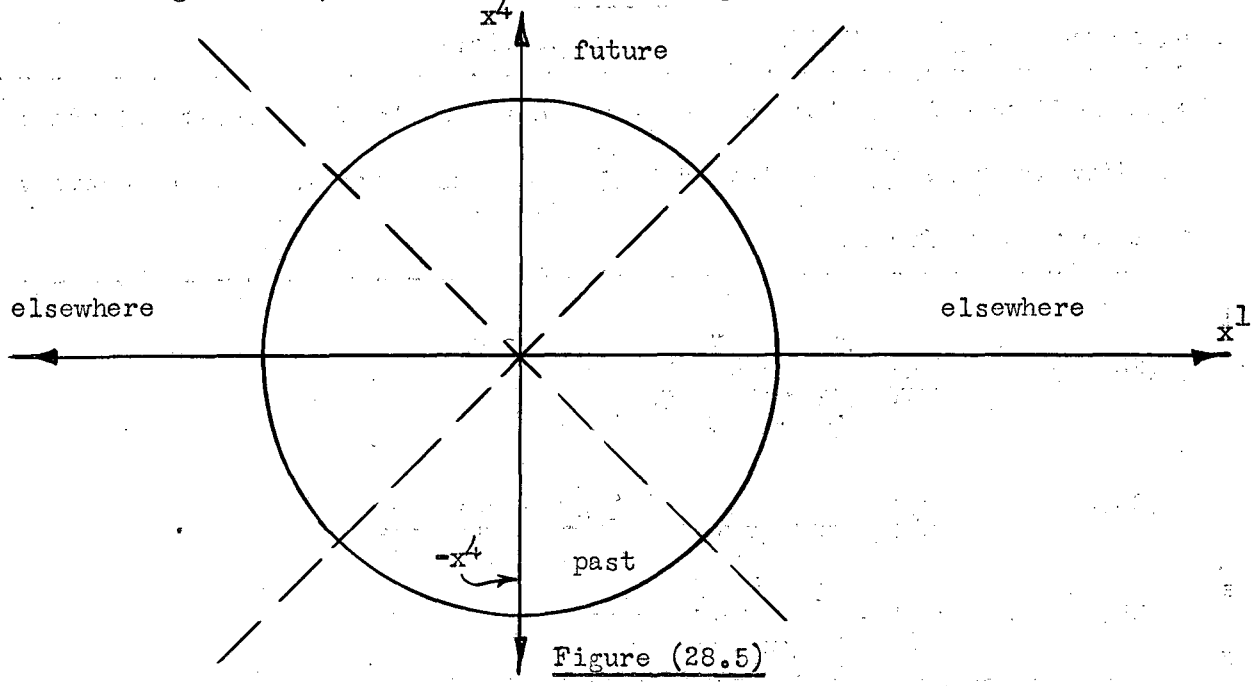


Figure (28.5)

1). The two spatial interval regions labeled "elsewhere" represent regions where events are located which from no inertial frame can be considered to occur at the origin.

2). The temporal region labeled "future" represents events whose temporal interval relative to the origin is positive from any inertial frame.

3). The temporal region labeled "past" represents events whose temporal interval relative to the origin is negative from any inertial frame.

Let us now extend the Lorentz transformation to other kinematic quantities. In the next chapter we shall introduce more general methods for obtaining

transformation relations between various physical quantities in different inertial frames. Let us here simply derive the transformation equations of simple kinematic quantities which can be obtained directly from the Lorentz transformation.

Let us write the Lorentz transformation in terms of differential intervals. Differentiating and dividing by dt' we obtain:*

* We shall use the symbols v and $\beta = v/c$ to denote the velocity of frame Σ' relative to frame Σ, evaluated in Σ, and $u_x = dx/dt$, etc. to denote velocities in a given frame.

$$\frac{dx'}{dt'} = \left[\frac{dx/dt - v}{\sqrt{1-\beta^2}}\right]\frac{dt}{dt'} \tag{28.38}$$

$$\frac{dy'}{dt'} = \frac{dy'}{dt}\,\frac{dt}{dt'} \qquad \frac{dz'}{dt'} = \frac{dz'}{dt}\,\frac{dt}{dt'} \tag{28.39}$$

$$\frac{dt'}{dt} = \frac{1 - \frac{\beta}{c}\frac{dx}{dt}}{\sqrt{1-\beta^2}} \tag{28.40}$$

or putting $u_x = dx/dt$; $u_x' = dx'/dt'$, etc., and substituting (28.40) into (28.38) and (28.39), we obtain:

$$u_x' = \frac{u_x - v}{1 - u_x v/c^2} \tag{28.41}$$

$$u_y' = \left[\frac{\sqrt{1-\beta^2}}{1-u_x v/c^2}\right] u_y \qquad u_z' = \left[\frac{\sqrt{1-\beta^2}}{1-u_x v/c^2}\right] u_z \tag{28.42}$$

Eq.(28.41) the "longitudinal velocity addition formula" is in agreement with Eq. (28.19) for the successive Lorentz transformations, since (28.19) must remain true as u_x may represent the motion of the origin of another Lorentz frame relative to the Σ' frame.

Let us consider Eq. (28.40) in the form:

$$(28.43) \qquad \frac{dt'}{dt} = \frac{1 - vu_x/c^2}{\sqrt{1 - \beta^2}}$$

If it were possible to make:

$$(28.44) \qquad \frac{dt'}{dt} < 0$$

by a suitable choice of v, then the temporal sequence of two events would be reversed between the two frames under consideration. This would be considered a logical contradiction if:

1) The two events represent a cause and effect.

2) The sense of time has an invariant significance.

The latter point has been a question of some discussion. Let us assume here that the sense of time can be ascertained independently in any frame, for instance, by use of the Second Law of Thermodynamics. This could be done by taking the positive direction of time to be the direction of statistically increasing disorder or increasing entropy and of degradation of heat. Assuming this point, we therefore conclude that, in order that the sequence between cause and effect be preserved as observed from any frame, that in any particular frame:

$$(28.45) \qquad u \leq c$$

where u represents the velocity of propagation of any event which can connect cause and effect. Obviously velocities like phase velocities, or velocities of geometrical significance only, are not affected by this restriction. The restriction also applies to the relative velocities v of possible inertial frames so that it is not necessary to discuss the significance of the Lorentz transformation when $\sqrt{1 - \beta^2}$ becomes complex.

Let us consider some simple applications of the velocity transformation relations:

1. Fizeau Convection Coefficient - (cf. Ch. 19, Eqs. (19.12) to (19.20)).

Let u' be the velocity of light which an observer finds for light propagated

in a plane wave in a fluid moving with a velocity v relative to him. In a frame Σ' moving with the fluid the light would be propagated with a phase velocity:

$$u' = c/n \tag{28.46}$$

where n is the refractive index. Transforming to a frame in relative motion to the fluid with velocity -v and applying (28.41):

$$u = \frac{c/n + v}{1 + v/nc} \cong \frac{c}{n} + v\left[1 - 1/n^2\right] \tag{28.47}$$

to the first order in v/c. This is in agreement with the experimental facts and the classical electrodynamic result derived in Ch. 19. Note that in the classical discussion a relatively complicated mechanism was involved, namely re-radiation from the moving secondary radiators in the fluid, which led to (28.47). The relativistic discussion on the other hand, led to this relation without any detailed information about the mechanism. We shall frequently meet situations in which an end result is demanded by relativistic considerations but where the mechanism of attaining the result is far less obvious.

2. Aberration - Consider a beam of light emitted by a star in the direction of the earth which strikes the earth in a direction transverse to the velocity of the earth (considered along the x axis) as seen in Figure (28.6). In the

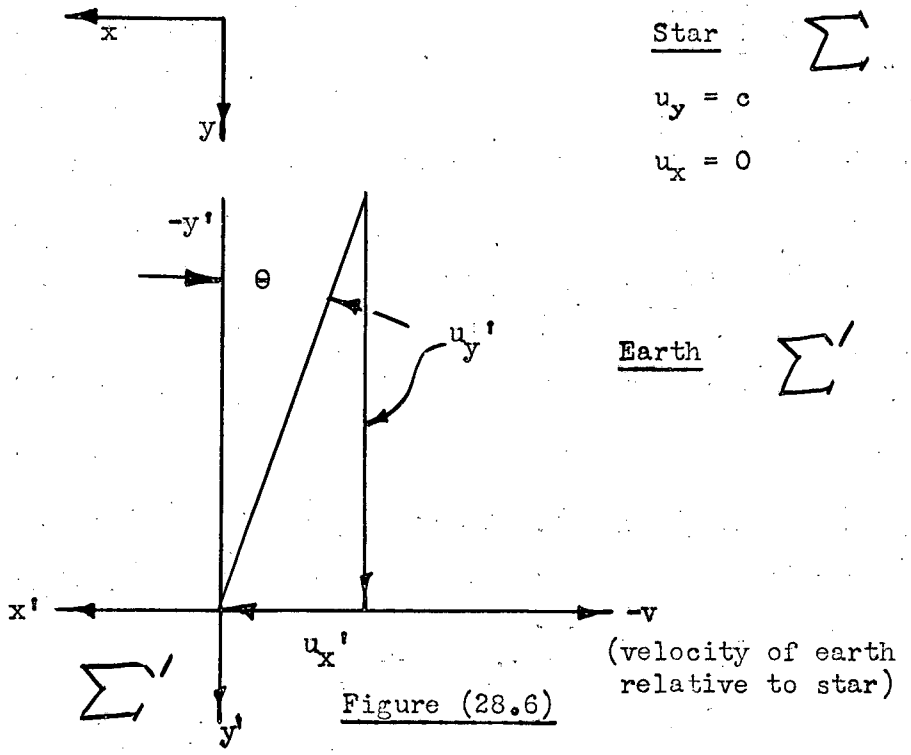


Figure (28.6)

star's frame the process is simply the emission of light with velocity components $u_y = c$; $u_x = 0$. In the earth's frame, the velocity components become from (28.41) and (28.42):

$$u_y' = \sqrt{1-\beta^2}\ c \qquad u_x' = v \tag{28.48}$$

Therefore the angle of incidence to the normal becomes:

$$\tan\theta' = \frac{u_x'}{u_y'} = \frac{\beta}{\sqrt{1-\beta^2}} \quad ; \text{ or}$$

$$\sin\theta' = \beta \tag{28.49}$$

A mechanical emission picture or stationary ether assumption would give $\tan\theta' = \beta$ which is in practice indistinguishable from (28.49). On the other hand, any convectively carried ether theory would contradict (28.49) as it would the Fizeau experiment.

We shall obtain (28.49) again from more detailed consideration of the propagated wave; it is, however, again characteristic that any consideration of mechanism can be by-passed.

CHAPTER 29

COVARIANCE

In the previous chapters, we have investigated the bearing of the principles of special relativity on the laws of kinematics. We obtained new relations by applying the principles of special relativity to two regions of physics; kinematics and the propagation of plane electromagnetic waves. The principles of relativity apply to all fields of physics, however, for by no experiment in any field should it be possible to detect a preferred inertial frame. The bearing of the principles of relativity on other fields could of course be discussed by designing "Gedanken Experimente" in those fields and thus obtaining new laws valid in those fields. Or we could attempt to obtain transformation relations for physical quantities in these fields by applying the Lorentz transformation to the time and space coordinates of the pertinent pre-relativistic equations and then trying to deduce the transformation relations for the remaining quantities. Both of these approaches are useful; in particular we shall use the thought experiment approach in our discussion of collisions. The direct transformation approach is frequently tedious. However, it was used by Einstein in his original work to deduce the transformation equations for the electromagnetic fields and to show that Maxwell's equations are in agreement with relativistic principles.

There is a third approach which is by far the most powerful one in extending relativity to other fields. This approach is to rewrite the equations of these fields in a form which explicitly makes evident in which way the quantities would behave under a change to a different inertial frame. If an equation has a form which is invariant to a change in inertial frame, then an experiment based on this equation obviously could not give a result depending on the particular frame of reference. The equation then describes a phenomenon which would be in agreement with the principles of special relativity. An equation written in such a way that its form is independent of the choice of inertial frame is said to be "Lorentz covariant."

The Lorentz transformation can be written as a linear transformation of four components of a vector x^j, $(x^1, x^2, x^3, x^4 = ct)$, in a four-dimensional space; i.e. it can be written as:

$$x'^i = Q_j{}^i\, x^j \qquad \text{(summation convention!)} \tag{29.1}$$

where $Q_j{}^i$ is given by the matrix:

$$\left(Q_j{}^i\right) = \begin{pmatrix} \gamma & 0 & 0 & -\beta\gamma \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -\beta\gamma & 0 & 0 & \gamma \end{pmatrix} \tag{29.2}$$

where:

$$\beta = v/c \tag{29.3}$$

$$\gamma = \frac{1}{\sqrt{1-\beta^2}} \tag{29.4}$$

We have seen that if the components of the vector x^i transform in accordance with Eq. (29.1), then an experiment involving x^i cannot yield a preferred frame. If therefore any physical relation is written in the form* of a vector equation in four space where the vector components transform in accordance with Eq. (29.1), then such an equation is said to be written in Lorentz covariant form. If we solve (29.1) for x^j, we obtain:

* By "form" we include in special relativity the "numerical content" of all equations; i.e., the magnitude of all general physical constants, etc.

$$x^j = \left(Q_j{}^i\right)^{-1} x'^i \tag{29.5}$$

where:

$$\left(Q_j{}^i\right)^{-1} = \begin{pmatrix} \gamma & 0 & 0 & +\beta\gamma \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ +\beta\gamma & 0 & 0 & \gamma \end{pmatrix} \tag{29.6}$$

is the inverse matrix of (29.2); i.e., the matrix of the transformation corresponding to relative motion of the frames with opposite velocity. If a quantity with four components A_j transforms as the reverse transformation of the x^j, i.e., as:

$$A_i' = (Q_j{}^i)^{-1} A_j \tag{29.7}$$

then a relation equating components of the type A_j is Lorentz covariant also. To sum up, any quantities A_j and B^j are Lorentz covariant if under change of inertial frame they transform as (29.7), or as:

$$B'^i = Q_j{}^i B^j \tag{29.8}$$

respectively. A_j is called a covariant* 4-vector and B^j a contravariant 4-vector.

---

* It is unfortunate that two different uses of the word "covariant" are being made here, but this usage appears to be accepted.

Note that $(Q_j{}^i)^{-1}$ and $Q_j{}^k$ are related by:

$$(Q_j{}^i)^{-1} Q_j{}^k = \delta_i{}^k \tag{29.9}$$

where:

$$\delta_i{}^k = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{29.10}$$

Any quantity of the type:

$$T^{i_1\, i_2\, i_3 - - - i_n}_{j_1\, j_2\, j_3 - - - j_m} \tag{29.11}$$

which transforms like:

$$T'^{i_1\, i_2\, i_3 - - i_n}_{\ell_1\, \ell_2\, \ell_3 - - \ell_m} = Q^{i_1}_{k_1} Q^{i_2}_{k_2} Q^{i_3}_{k_3} - Q^{i_n}_{k_n} \left(Q^{\ell_1}_{j_1}\right)^{-1} \left(Q^{\ell_2}_{j_2}\right)^{-1} \left(Q^{\ell_3}_{j_3}\right)^{-1} - - \left(Q^{\ell_m}_{j_m}\right)^{-1} T^{k_1\, k_2\, k_3 - - k_n}_{j_1\, j_2\, j_3 - - j_m} \tag{29.12}*$$

* Tensors of physical interest can be represented by matrices. Matrix multiplication is usually non-commutative. However, when written out in component form as in (29.12), where the dummy summation convention is being used, the individual elements are numbers and their order does not matter. (29.12) is written in the standard form.

i.e., simply like products of covariant and contravariant 4-vector components, is called a tensor of rank (n + m). An equality between such tensors will also be Lorentz covariant. We shall therefore attempt to extend the principles of relativity to other fields of physics by rewriting their laws in the form of tensor equations; their covariance is then self-evident and the transformation laws of the quantities involved in the equations will follow from Equation (29.2), (29.6) and (29.12). Of course, we have not proved, nor is it necessarily true, that the writing of the laws of physics in tensor form is the only way of describing physics in a covariant way. All we can say is that it is a suitable way of doing so.

Before proceding with this program, let us summarize a number of useful facts concerning tensors with constant transformation coefficients.**

** In the Lorentz transformation of special relativity, the transformation coefficients $Q_i{}^j$ are constants. In general tensor analysis, as is used in the general theory of relativity, these coefficients are themselves functions of the coordinates and represent the partial derivatives of the transformation. I.e. (29.1) becomes, differentially:

$$dx'^i = \left(\frac{\partial x'^i}{\partial x^j}\right) dx^j \qquad (29.13)$$

covariant and contravariant vector components transform respectively as:

$$(29.14) \qquad A_i' = \left(\frac{\partial x^j}{\partial x'^i}\right) A_j \qquad\qquad B'^i = \left(\frac{\partial x'^i}{\partial x^j}\right) B^j$$

---

1) A tensor of rank zero is called an invariant or scalar, i.e.:

$$(29.15) \qquad T_o' = T_o$$

2) The product of a tensor of rank n and a tensor of rank m is a tensor of rank n + m.

3) When a contravariant and a covariant index of a tensor have the same index letter, the resultant summation reduces the rank of the tensor by two. This process is called contraction. The contraction of a tensor $T_j^{\ i}$ is an invariant.*

---

* Summation over two covariant or two contravariant indices does not lead to a new tensor and is therefore not an invariant operation.

---

4) By the basic postulates, the space-time interval in special relativity (line-element):

$$(29.16) \qquad ds^2 = -dx^2 - dy^2 - dz^2 + c^2dt^2$$

is a scalar invariant. It can be written in the form:

$$(29.17) \qquad ds^2 = g_{ij}\, dx^i\, dx^j$$

where:

$$(29.18) \qquad (g_{ij}) = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & +1 \end{pmatrix} = \begin{pmatrix} -1 & & & 0 \\ & -1 & & \\ & & -1 & \\ 0 & & & +1 \end{pmatrix}$$

is called the "metric tensor" corresponding to the line-element.* It can be

---

* In general tensor analysis g_{ij} is a function of the coordinates.

---

proved easily by transformation, using Eq. (29.2), that g_{ij} is actually a tensor.

In Euclidean geometry $g_{ij} = \delta_{ij}$.

5) By use of the relation:

$$B_i = g_{ij} B^j \tag{29.19}$$

each member of a set of covariant components B_i can be associated with a contravariant tensor component. For the special form of g_{ij} given by Eq. (29.18) this process will simply reverse the sign of the first three components. Note that $g_{ij} = g^{ij}$.

6) The covariant components dx_i corresponding to the basic contravariant coordinate interval $dx^i = (dx, dy, dz, cdt)$ is $dx_i = (-dx, -dy, -dz, cdt)$ and:

$$dx^i \, dx_i = ds^2 \tag{29.30}$$

7) The derivative:

$$\frac{\partial}{\partial x^i}$$

when operating on a tensor transforms like an additional covariant tensor component.* The increment of a scalar S can thus be written as a tensor relation:

$$dS = \frac{\partial S}{\partial x^i} \, dx^i \tag{29.21}$$

---

* This is not true if the g_{ij} are functions of the coordinates.

---

8) If the product of a tensor and a symbol of unknown transformation character is a tensor, then that symbol transforms like a tensor also.

A covariant relation in physics can be generated by one of the following processes:

1) The relation is known in a special inertial frame, such as a proper frame where the system under consideration is at rest. If it is possible to write a tensor equation which reduces to the special relation for the special frame, then this tensor equation has general significance.

2) A known tensor relation is converted into a new tensor relation by a covariant tensor operation (the simplest example is multiplication by an invariant).

3) An equation is obtained from a relation valid in a special frame by transformation of the remaining quantities deduced. The resultant quantities are then expressible in tensor form. This process is the direct transformation process referred to above and is usually very tedious.

A tensor theorem frequently of importance is the following: If a four-vector j^i obeys a relation of the type (conservation law; compare Equation (9.1))

$$\frac{\partial j^i}{\partial x^i} = 0 \tag{29.22}$$

and if j^1, j^2, j^3 are different from zero only in a finite spatial region, then the integral in 3-dimensional space

$$\iiint j^4 dv \tag{29.23}$$

is an invariant. To prove this theorem let us apply Gauss' theorem in 4-space

$$\iiiint \frac{\partial j^i}{\partial x^i} \, d^4x = \iiint j^k \, d^4S_k \tag{29.24}$$

★ $d^4x = dx^1 dx^2 dx^3 dx^4$;

d^4S_k = element of 3-dimensional "surface" normal to x^k in 4-space.

to the boundary shown in Figure (29.1):

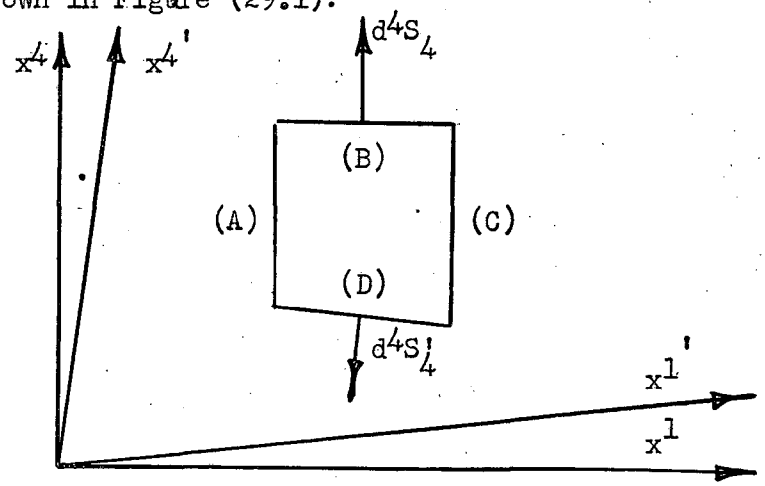


Figure (29.1)

Let the surfaces (A) and (C) be chosen such that the spatial components of j^i vanish on (A) and (C). This can be done since the region of non-vanishing spatial components of j^i was assumed to be finite. Let (B) be chosen normal to the time-axis in a Σ-frame while (D) is chosen normal to the time axis in a Σ'-frame. It then follows from (29.22) and (29.24) that:

$$(29.25) \qquad \iiint_B j^4 dS_4 = \iiint_D j^{4\prime} d^4S'_4$$

and hence, by geometry:

$$(29.26) \qquad \iiint j^4 dv = \iiint j^{4\prime}\, dv' = \text{invariant}$$

which was to be proved. It follows similarly that if a "conservation law":

$$(29.27) \qquad \frac{\partial T^{ij}}{\partial x^i} = 0$$

applies to a tensor of 2^{nd} rank, then:

$$(29.28) \qquad \iiint T^{4j} dv$$

is a 4-vector.

In the next chapters we shall attempt to express the laws of point mechanics and then the laws of electrodynamics in covariant form.

CHAPTER 30

RELATIVISTIC MECHANICS

In the preceding chapter we have studied formal methods which will enable us to generate "covariant" laws of physics in the relativistic sense. Let us use these considerations to formulate the law of conservation of momentum for point particles. This will give us certain formal relations which we shall then show are in agreement with a thought experiment involving an inelastic collision as observed from various frames.

$$(30.1) \qquad \vec{p} = m\vec{u}$$

The vector $\vec{u}$ does not constitute the first three components of a four vector, for $u_x = dx/dt$ etc., and dt is not an invariant. However, if we divide the contravariant vector dx^i by the invariant line element ds we obtain a four vector:

$$(30.2)* \qquad u^i = \frac{dx^i}{ds}$$

---

* Note that the components of this four-vector velocity $\left(\frac{\gamma_o}{c}\vec{u},\ \gamma_o\right)$ where $\gamma_o = \frac{1}{\sqrt{1 - u^2/c^2}}$, do not have the dimensions of velocity but are dimensionless.

---

known as the four-velocity. A covariant expression corresponding to momentum can then be generated by multiplying (30-2) by an invariant quantity $m_o c^2$, which is assumed to be characteristic of the particle, where m_o in the proper (rest) frame of the particle would be its mass. The four-momentum is thus:

$$(30.3)** \qquad p^i = m_o \frac{dx^i}{ds} c^2$$

** Note that the components of the four-vector momentum p^i have the dimensions of of energy.

If we formulate laws involving momenta in terms of p^i we thereby assure their covariant properties. If we want to assure that the law of conservation of momentum shall be preserved in the framework of relativity for two particles interacting at a point, then in order that this law be independent of the inertial frame chosen, it must take the form:

$$(30.4) \qquad p_1{}^i + p_2{}^i = \text{const.}$$

before and after the interaction, rather than the classical law:

$$(30.5) \qquad m_1\vec{u}_1 + m_2\vec{u}_2 = \vec{\text{const.}} \quad .$$

with the assumption that m_1 and m_2 are constant.

We are restricting this discussion to particles in direct interaction, rather than interacting at a distance, since the total momentum of two separated particles at a "given" time has no meaning in relativity and since all interactions are necessarily propagated with finite velocity. Therefore, momentum between separated particles has meaning only if each particle conserves momentum with a field acting on it, or if the interaction is carried by a particle interacting in succession with the interacting mass points. Strictly speaking then, each of the mass points considered here must have zero extension in order that this discussion be rigorous; if its extension were non-zero, it could no longer be considered rigid.

The components of the four-momentum are, using (29.16) for ds:

$$(30.6)^* \qquad (p^i) = \left(m_o \frac{\vec{u}c}{\sqrt{1-u^2/c^2}} \, , \, \frac{m_o c^2}{\sqrt{1-u^2/c^2}} \right) = (\vec{p}c, \, mc^2)$$

---

* If the spatial components of a four-vector conform to a standard three dimensional vector, we shall in enumerating the components of the four-vectors use regular vector notation, e.g.

$$\ldots\ldots (dx^i) = (\vec{dr}, \, cdt)$$

---

where $\vec{u}$ is the ordinary velocity in a given frame. The first three components have the form:

(30.7) $\vec{cp} = (m\vec{u})c$

where:

(30.8) $$m = \frac{m_o}{\sqrt{1-u^2/c^2}}$$

Hence if it is desired to maintain the law of conservation of momentum in the classical form (30.5) and in order to have this law hold independent of the choice of inertial frame, the mass can no longer be considered to be an invariant, but will depend on the velocity $\vec{u}$ as measured in the particular frame, by the relation (30.8). The "variation of mass with velocity" is thus a necessary consequence of formulating the law of conservation of momentum in a two-body collision in a covariant manner.

The law of conservation of mementum (30.4) implies not only the conservation of the three "spatial" components but also the conservation of the fourth component:

(30.9) $$p^4 = \frac{m_o c^2}{\sqrt{1-u^2/c^2}} = mc^2$$

Let us investigate the physical significance of this quantity. Let us consider the time rate of change of this quantity in a given frame. We can show algebraically that:

(30.10) $$\frac{d}{dt} p^4 = \frac{d}{dt}\left[\frac{m_o c^2}{\sqrt{1-u^2/c^2}}\right] = \vec{u} \cdot \frac{d}{dt}\left[\frac{m_c \vec{u}}{\sqrt{1-u^2/c^2}}\right]$$

or

(30.11) $$\frac{dp^4}{dt} = \vec{u} \cdot \frac{d\vec{p}}{dt}$$

Hence if we continue to call:

(30.12) $$\vec{F} = \frac{d\vec{p}}{dt}$$ (Not the space component of a four-vector!)

the force, i.e. if we measure force in terms of rate of change of momentum then:

(30.13) $$P = \vec{F} \cdot \vec{u} = \vec{u} \cdot \frac{d\vec{p}}{dt}$$ P = Power

represents the rate at which work is being done in a particular system. Hence if the law of conservation of energy is to hold in a particular frame, and if E is

the energy:

$$(30.14) \qquad \frac{dp^4}{dt} = \frac{dE}{dt} \; ; \; p^4 = E + \text{const.}$$

The constant of integration has no particular significance since it affects only the zero of energy; since energy only manifests itself when changes of the energy occur, we lose no physical significance if we put:

$$(30.15) \qquad E = mc^2$$

We are thus led to the following conclusions:

1) Energy as measured by "work-content" and mass as measured by the momentum for a given velocity are interchangable concepts; when one exists so does the other. Neither mass nor energy are invariants; their magnitude depends on the frame of the observer by the relation:

$$(30.16) \qquad E = mc^2 = \frac{m_o c^2}{\sqrt{1-u^2/c^2}}$$

where $\vec{u}$ is the observed velocity in a particular frame. We have shown that the change in mc^2 corresponds to work done by mechanical forces; that it corresponds also to change in energy under whatever mechanism might be involved implies an additional assumption whose justification rests with experiment. Experience in other fields of physics, in particular nuclear physics, where the fractional mass changes become very large, certainly prove beyond any reasonable doubt that (30.15) is valid in this more general interpretation.

2) The conservation of energy and the conservation of momentum are not independent principles; one demands the other for a covariant formulation.

3) The four-vector p^i which we can now call the energy-momentum vector has the components:

$$(30.17) \qquad (p^i) = (\vec{cp}, E)$$

and the invariant $p^i p_i$ is:

(30.18) $$p^i p_i = E^2 - c^2 p^2 = (m_o c^2)^2$$

This relation between energy and momentum is thus valid in any frame. In a proper frame ($p = 0$) we have simply:

(30.19) $$E = m_o c^2$$

as expected.

4) p^i is a contravariant vector; thus on changing from a $\sum$ to a $\sum'$ frame by the use of (29.2), we have the transformation relations:

(30.20) $$cp_x' = \frac{cp_x - \beta E}{\sqrt{1-\beta^2}}$$

(30.21) $$p_y' = p_y$$

$$p_z' = p_z$$

(30.22) $$E' = \frac{E - \beta cp_x}{\sqrt{1-\beta^2}}$$

If $\sum'$ is a proper frame for the particle, ($\vec{p}' = 0$, $E' = m_o c^2$) these can be written in the following form, useful for computation:

(30.23) $$\beta = \tanh \phi; \quad m_o c^2 = E_o$$

then, from (30.20):

(30.24) $$cp_x = m_o c^2 \sinh \phi$$

(30.25) $$E = m_o c^2 \cosh \phi$$

This relation corresponds geometrically to rotation through an angle ϕ in a (cp_x, iE) plane (c.f. Chapter 28, Figure 28.2).

5) For small velocities E reduces to the classical kinetic energy plus the "rest energy" $m_o c^2$. By expansion:

(30.26) $$E = m_o c^2 \left[1 + 1/2\left(\frac{u}{c}\right)^2 + 3/8\left(\frac{u}{c}\right)^4 - - -\right] = m_o c^2 + 1/2\, m_o u^2 + \frac{3}{8}\,\frac{m_o u^4}{c^2} - -$$

6) Since mass and energy are equivalent quantities, any transfer of energy

implies transfer of mass and therefore a momentum. Consider a motor M driving a load L at a distance x by means of a belt as seen in Fig. (30.1).

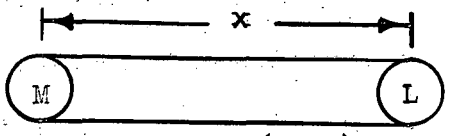


Figure (30.1)

If the motor transfers energy at the rate dE/dt to the load, the mass of L increases correspondingly, the system thus has a momentum:

$$(30.27) \qquad p = \frac{x}{c^2}\frac{dE}{dt}$$

If energy is absorbed by a body at a given rate, then the momentum of the body increases, to conserve the overall momentum we must thus associate a density of momentum $\vec{g}$ with any agent which transmits energy at the rate $\vec{S}$ per unit area in a given direction, given by:

$$(30.28)^* \qquad \vec{g} = \frac{\vec{S}}{c^2}$$

---

* This relation can be deduced formally from the transformation properties of $\vec{p}$ and E. If we consider Σ to contain a system of energy E of zero momentum then from (30.20) in Σ' we will observe a momentum:

$$p_x' = \frac{-\beta E}{c\sqrt{1-\beta^2}} = \frac{-v}{c^2} E'$$

This is equivalent to (30.28) and includes all forms of energy.

This relation is in agreement with our discussion of radiation pressure given in Chapter 17. If we consider electromagnetic radiation which represents an energy flow $\vec{S}$ incident upon an observer, then in the absence of a force sustaining medium or ether, we were forced to attribute a momentum density as given by (30.28) to the electromagnetic field. The fact, as evidenced by the failure of the various experiments which were devised to detect the existence of an ether, that electromagnetic radiation, has the singular property of existing in free space, makes the system of radiation and absorber a <u>closed</u> system. The result

is that Equation (30.28) is required by the conservation of momentum. Equation (30.28) could be made the starting point of our mass-energy equivalence discussions, using only the assumption of conservation of momentum and the absence of an ether. This approach is frequently used in obtaining the relation $E = mc^2$ without the necessity of introducing the entire relativistic kinematics.

7) The force $\vec{F} = \frac{d\vec{p}}{dt}$ in a given frame is not the spatial component of a four-vector. On the other hand, a quantity known as the "Minkowski force":

$$(30.29) \qquad (F^i) = \frac{dp^i}{ds} = \frac{d}{ds}\,(\vec{pc}, mc^2) = \left(\frac{1}{\sqrt{1-u^2/c^2}} \frac{d\vec{p}}{dt}, \frac{1}{\sqrt{1-u^2/c^2}} \frac{d(cm)}{dt} \right)$$

is a contravariant four-vector. The components of F^i can be written as: (from (30.12) and (30.12):

$$(30.30)^* \qquad (F^i) = \left(\frac{1}{\sqrt{1-u^2/c^2}} \vec{F}, \quad \frac{1}{\sqrt{1-u^2/c^2}} \vec{F} \cdot \frac{\vec{u}}{c} \right)$$

---

* Note that the dimensions of F^i <u>are</u> those of a force.

The transformation laws for force can be derived from the four-vector character of F^i. Let us restrict ourselves to the case where $\vec{F}$ is proper in the $\sum$ frame, i.e. $u = 0$, $F^i = (\vec{F}, 0)$. In the $\sum'$ frame (using the fact that $u_x' = v$, since $u_x = 0$) we obtain:

$$(30.31) \qquad F_x' = F_x$$

$$(30.32) \qquad F_y' = F_y \sqrt{1-\beta^2}$$

$$(30.33) \qquad F_z' = F_z \sqrt{1-\beta^2}$$

Let us apply these relations to the equilibrium of a right angle lever of equal arms shown in Figure (30.2). Let the lever be at rest in the $\sum$ frame and let it be in equilibrium in $\sum$ under the influence of the forces F_x and F_y as shown. For simplicity let the arms of the lever be of equal length ℓ, i.e. in equilibrium:

(30.34) $$F_x = F_y$$

Let us now consider the system from a frame $\sum'$. We would expect static equilibrium to be an invariant property since otherwise the frames would be distinguishable. Let us analyze the forces in $\sum'$. From the transformation equations (30.31) and (30.32) and the Lorentz contraction, it would appear

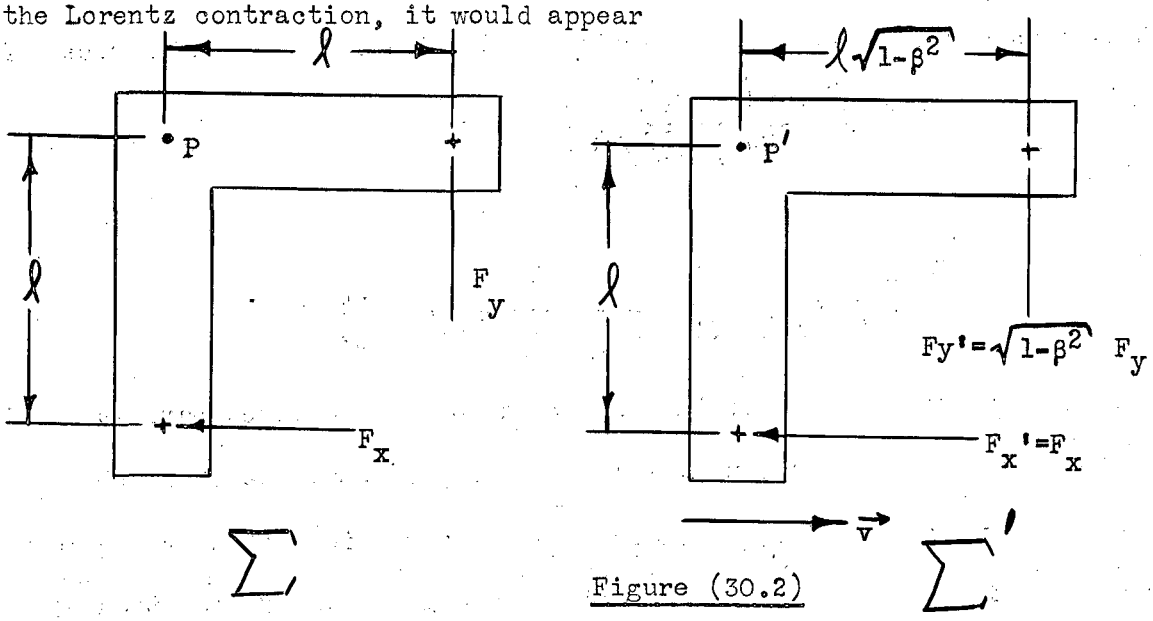


Figure (30.2)

that a net torque of magnitude:

(30.35) $$L' = F_x \ell \left[1-(1-\beta^2)\right] = F_x \ell \beta^2$$

were acting. However, this torque does not produce any rotation since, in $\sum'$, F_x' seems to be doing work on the lever at the rate $F_x v$; the angular momentum of the lever therefore increases at a rate:

(30.36) $$\frac{dM}{dt} = \ell\ (F_x v)\ \frac{v}{c^2} = F_x \ell \beta^2$$

Hence in $\sum'$ even though the torque and the angular momentum do not separately vanish, the torque exactly balances the gain in angular momentum. Equilibrium is thus preserved as an invariant property.*

---

* The above discussion omits entirely the mechanism by which forces are transmitted through the lever; obviously the laws of elasticity will also be profoundly modified by relativity; the lever cannot be treated as a rigid body, since the velocity of propagation of an impulse is limited. A more detailed discussion (See e.g. Tolman, "Relativity, Cosmology and Thermodynamics," Pg.79ff.)

does not alter the above conclusions.

- -

Equilibrium at a point evidently remains invariant since at a point all balancing forces transform by the same laws.

8) Let us consider the motion of a particle under the influence of forces. From:

(30.12) $$\vec{F} = \frac{d\vec{p}}{dt} = \frac{d}{dt}\left[\frac{m_o\vec{u}}{\sqrt{1-u^2/c^2}}\right]$$

We have:

(30.37) $$\vec{F} = \frac{m_o}{\sqrt{1-u^2/c^2}}\frac{d\vec{u}}{dt} + \frac{m_o u\,\vec{u}}{c^2(1-u^2/c^2)^{3/2}}\frac{du}{dt}$$

Acceleration of a moving particle thus requires not only a force parallel to the acceleration but also a component parallel to the velocity. We shall show in a following chapter that the "Lorentz force" $\vec{F}_\ell$ on a particle of charge q moving with velocity $\vec{u}$ in an electric field $\vec{E}$ and a magnetic field of induction $\vec{B}$:

(30.38) $$\vec{F}_\ell = q(\vec{E} + \vec{u} \times \vec{B})$$

has the same transformation character as the force $\vec{F}$ above and can thus be equated to (30.37). The resultant equation:

(30.39) $$\frac{d}{dt}\left[\frac{m_o\vec{u}}{\sqrt{1-u^2/c^2}}\right] = q(\vec{E} + \vec{u} \times \vec{B})$$

is the general equation of motion of a charged particle in an electromagnetic field and is in excellent agreement with observed particle behavior.

We have been led to the change of mass with velocity by two processes: first, by the formal approach of attempting to formulate the conservation of momentum covariantly; and second, have also indicated the possible deduction of the principle by the consideration of the interaction of "free" electromagnetic radiation with an absorber. We shall now show directly without reference to tensor methods, how this principle can be deduced by use of the Lorentz transformation and the requirement that in a two-particle collision momentum be conserved.

Consider two particles of equal properties colliding head on. Let us choose a frame $\sum'$ such as the center of mass frame in which the collision is symmetrical, i.e., they appear to approach with equal velocities as in Figure (30.3). After a short period of coalescence, they will then depart with opposite velocities.

Let us make the same postulates as we did before in the more formal approach, i.e. let us conserve momentum and mass in any frame. Hence, in $\sum$:

(30.40) $$m_1 + m_2 = M$$

(30.41) $$m_1u_1 + m_2u_2 = Mv$$

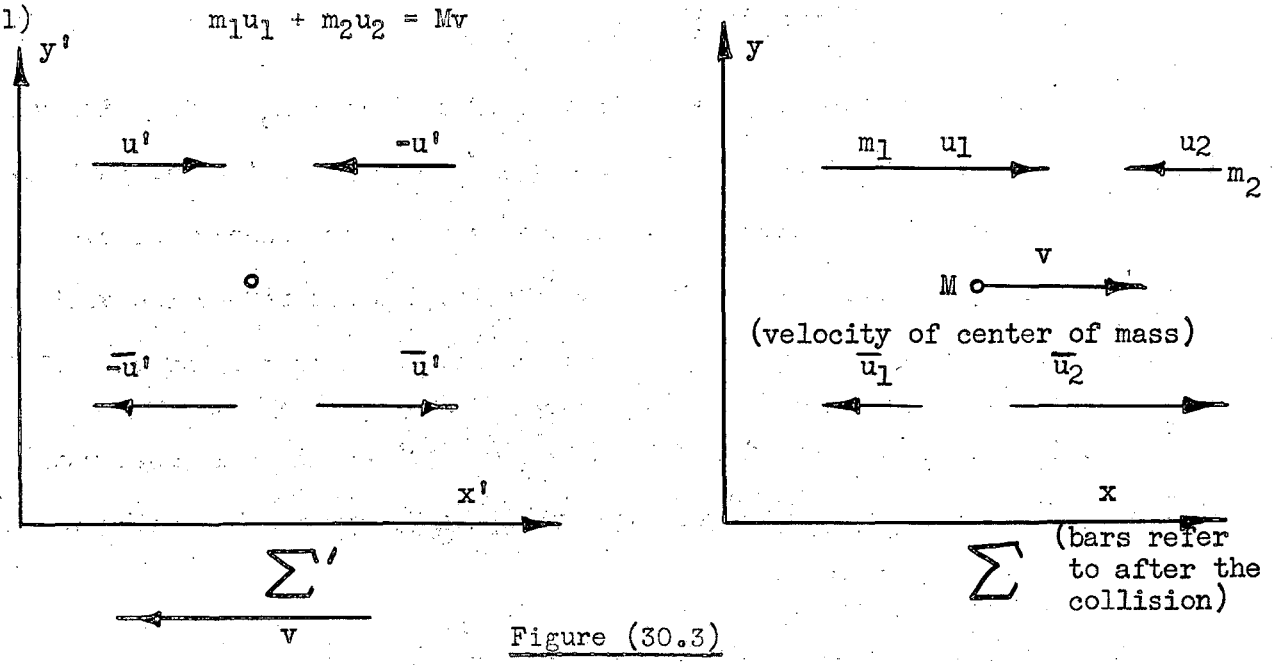


Figure (30.3)

where M is the combined mass during collision. But from the longitudinal velocity addition relations, we have:

(30.42) $$u_1 = \frac{u' + v}{1 + u'v/c^2} \qquad u_2 = \frac{-u' + v}{1 - u'v/c^2}$$

Eliminating M and reducing algebraically we obtain:

(30.43) $$\frac{m_1}{m_2} = \frac{1 + u'v/c^2}{1 - u'v/c^2} = \frac{\sqrt{1 - u_2^2/c^2}}{\sqrt{1 - u_1^2/c^2}}$$

Hence, in order to preserve the conservation laws (30.40) and (30.41) in all frames, we must have:

(30.44) $$m = \frac{m_o}{\sqrt{1 - u^2/c^2}}$$

where m_o is the mass in a proper frame. This is in agreement with (30.8) and the further deductions follow as above.

This description is of particular interest since it enables us to obtain from (30.42) the mass M during the collision:

$$M = m_o\left[\frac{1}{\sqrt{1 - u_1^2/c^2}} + \frac{1}{\sqrt{1 - u_2^2/c^2}}\right] = \frac{2m_o}{\sqrt{1 - v^2/c^2}} \frac{1}{\sqrt{1 - u'^2/c^2}} \tag{30.45}$$

This is larger than:

$$\frac{2m_o}{\sqrt{1 - v^2/c^2}} \tag{30.46}$$

which would be the mass of the two masses of rest mass m_o moving with velocity v. This increased mass represents the increase in energy of the two particles during collision owing to the stored elastic energy or to the energy increase in case not all the energy is released again. The distinction between an inelastic and elastic collision therefore essentially disappears in the first part of the collision.

Calculations of this "elastic mass increase" are usually most useful in practical cases, e.g. in calculating the available energy in the center of mass system in nuclear collisions.

CHAPTER 31

COVARIANT FORMULATION OF VACUUM ELECTRODYNAMICS

Maxwell's equations and their consequences lend themselves very simply to covariant description. This follows from the fact that no modifications are necessary at all in the laws of electrodynamics to make them agree with the requirements of relativity. The covariant formulation of the space-time coordinates in the equations automatically puts the rest of the equation in covariant form.

Recall that we introduced the Lorentz transformation by considering the covariant formulation of the propagation of a plane electromagnetic wave. Actually the form of the equations governing the propagation of any electromagnetic wave already agrees with the laws of relativity, since the D'Alembertian operation:

$$\Box = \frac{\partial}{\partial x^i}\,\frac{\partial}{\partial x_i} \tag{31.1}$$

is an invariant.

Since time and space coordinates are no longer independent, it is clear that charge density and current flow are simply different aspects of the same thing. If we have a "proper" charge density ρ_0 in a frame where such charges are at rest, then the contravariant vector:

$$j^i = \rho_0 \frac{dx^i}{ds} \tag{31.2}$$

has the components:

$$j^i = \left(\rho \frac{\vec{u}}{c}, \rho\right) \tag{31.3}$$

where:

$$\rho = \frac{\rho_0}{\sqrt{1 - u^2/c^2}} \tag{31.4}$$

Hence the transformation equations of charge and current densities follow automatically. Since current and charge densities are components of a single four-vector j^i, we are led to combine the inhomogeneous wave equations (18.18)

and (18.19) expressed as:

$$(31.5) \qquad \Box \phi = - \frac{\rho}{k_o}$$

$$(31.6) \qquad \Box (c\vec{A}) = - \frac{\rho \vec{u}}{c k_o}$$

into a single equation:

$$(31.7) \qquad \Box \phi^i = - \frac{j^i}{k_o}$$

where:

$$(31.8) \qquad \phi^i = (c\vec{A}, \phi)$$

We therefore find that electric and magnetic fields are no longer quantities permitting independent description; they are different aspects of the same thing.

The equation of continuity takes the simple covariant form:

$$(31.9) \qquad \frac{\partial j^i}{\partial x^i} = 0$$

and the Lorentz condition (18.15) becomes its counterpart:

$$(31.10) \qquad \frac{\partial \phi^i}{\partial x^i} = 0$$

The gauge transformations (18.22) and (18.23) combine into the form:

$$(31.11)^* \qquad \phi_i' = \phi_i + \frac{\partial}{\partial x^i} (c\Psi)$$

---

* Note that covariant components of ϕ^i are used, giving both the correct transformation character and the correct sign of (18.22) and (18.23).

The derivation of the field from the ϕ^i and any law of physical consequence must not depend on the choice of the scalar function Ψ.

Note that Eq. (31.4) which gives the transformation from a charge density at rest to a charge density in a non-proper frame, is in agreement with an invariant charge. A (spatial!) volume element dV is related to a proper volume

element dV_o by:

$$(31.12) \qquad dV = dV_o \sqrt{1-u^2/c^2}$$

since only one dimension is Lorentz contracted. Hence:

$$(31.13) \qquad \rho\, dV = \rho_o dV_o$$

and the charge within a given boudary remains invariant. Therefore, the electronic charge e remains a universal constant. Since no charges have been found in nature which are not integral multiples of e, total charge could be measured by a counting operation which should be invariant. These facts are all in agreement with experiment. Note that the invariance of total charge is also a direct consequence of theorem (29.23).

Let us obtain the integral of the inhomogeneous wave equation (31.7) corresponding to an (invariant) point charge e at a point P. This integral should correspond to the Wiechert-Liénard potentials (20.11) and (20.12). We know that in a proper frame, as in Figure (31.1), the integral of (31.7) is simply the Coulomb potential:

$$(31.14) \qquad \phi_o{}^i(x_Q{}^j) = \left[\vec{0},\ \frac{e}{4\pi k_o}\ \frac{1}{r_o(x_P{}^j,\ x_Q{}^j)}\right]$$

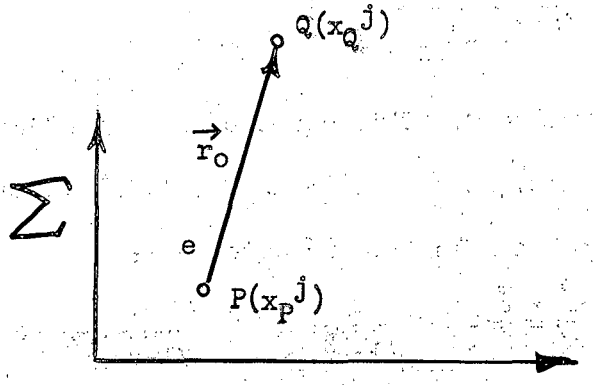


Figure (31.1)

Here $\vec{r}_o$ is the proper vector distance PQ. The potential signal at Q is to be measured at the time corresponding to the retardation condition:

$$(31.15) \qquad R^j R_j = 0 = -r^2 + c^2 t^2$$

where:

(31.16) $$R^j = x_P{}^j - x_Q{}^j \qquad \text{i.e. } R^j = (\vec{r}, ct) = (\vec{r}, r)$$

To make (31.14) valid in any frame, we seek to write it in tensor form such that it reduces to (31.14) if $\vec{u} = 0$. Consider the four-velocity:

(30.2) $$u^i = \frac{dx^i}{ds}$$

In a proper frame it has the components:

(31.17) $$u_o{}^i = (\vec{0}, 1)$$

Also the invariant:

(31.18) $$u^i R_i$$

can be evaluated in the proper frame where:

(31.19) $$R_o{}^j = (\vec{r}_o, r_o)$$

using the condition (31.15) which defines the time of propagation of the signal, we have:

(31.20) $$u^i R_i = r_o$$

and (31.14) can be written in tensor form:

(31.21) $$\phi^i = \frac{e}{4\pi k_o} \frac{u^i}{u^j R_j}$$

subject to the condition $R_j R^j = 0$. This equation is therefore now valid in any frame, whether proper or not. Let us show that (31.21) actually does correspond to the Wiechert-Liénard potentials. If we note that:

(31.22) $$u^j R_j = \frac{-\vec{r} \cdot \vec{u}}{c\sqrt{1-u^2/c^2}} + \frac{r}{\sqrt{1-u^2/c^2}} = \frac{1}{\sqrt{1-u^2/c^2}} \left(r - \frac{\vec{r} \cdot \vec{u}}{c}\right)$$

since:

(31.23) $$u^j = \left(\frac{\vec{u}}{c\sqrt{1-u^2/c^2}}, \frac{1}{\sqrt{1-u^2/c^2}} \right)$$

then (31.21) has the components:

(31.24) $$\phi^i = \frac{e}{4\pi k_o} \left(\frac{\vec{u}}{cs}, \frac{1}{s} \right)$$

where:

$$s = r - \frac{\vec{r} \cdot \vec{u}}{c} \tag{31.25}$$

Equation (31.24) is equivalent to Eqs. (20.17) and (20.18). Note however that the significance of the velocity has changed profoundly. In Chapter 20 the velocity was measured relative to a special frame in which the wave equation was valid; in (31.24) $\vec{u}$ is the velocity of the observer relative to the frame in which the charge was at rest at the time of "emission" of the signal. Hence all our detailed calculations of the fields of charges in various forms of motion remain correct provided the velocity is re-interpreted, in this manner.

Thus far, we have discussed the field equations entirely in terms of the electrodynamic potentials. Let us now derive the fields from the potentials covariantly. The field vectors $\vec{E}$ and $\vec{B}$ have different transformation character in three dimensional space. $\vec{E}$ is a "polar" vector, i.e. a vector whose components behave as do co-ordinate differences of a distance vector under change from a left-handed to right-handed co-ordinate system. $\vec{B}$ is an "axial" vector, i.e. a vector related to two polar vectors by a cross product relation. All correct physical vector equations are equations only between one kind of vector. Axial vectors cannot form the spatial components of a four-vector. On the other hand, the equation $\vec{C} = \vec{A} \times \vec{B}$ can be expressed in a "coordinate-sense" insensitive manner by describing $\vec{C}$ by a quantity:

$$C_{ij} = A_i B_j - A_j B_i = - C_{ji} \tag{31.26}$$

where:

$$C_{12} = \pm C_z \qquad C_{23} = \pm C_x \qquad C_{31} = \pm C_y \tag{31.27}$$

the + sign referring to a right-handed Cartesian system and the - sign to a left-handed system. For example, the equation $\vec{\nabla} \times \vec{E} = - \dot{\vec{B}}$ is written in this notation as:

$$\frac{\partial E_j}{\partial x_i} - \frac{\partial E_i}{\partial x_j} = -\dot{B}_{ij} \tag{31.28}$$

and $\vec{\nabla} \times \vec{A} = \vec{B}$ as:

$$\frac{\partial A_j}{\partial x_i} - \frac{\partial A_i}{\partial x_j} = B_{ij} \tag{31.29}$$

$\vec{E}$, on the other hand, is polar; it can be derived from the potentials by:

$$E_i = -\frac{\partial \phi}{\partial x_i} - \frac{\partial A_i}{\partial t} \qquad (i = 1,2,3) \tag{31.30}$$

This leads us to the introduction of a four-dimensional antisymmetric field tensor F_{ij} which, as a function of:

$$(\phi_i) = (-c\vec{A}, \phi) \tag{31.31}$$

is:

$$F_{ij} = \frac{\partial \phi_j}{\partial x^i} - \frac{\partial \phi_i}{\partial x^j} \; ; \qquad F^{ij} = \frac{\partial \phi^j}{\partial x_i} - \frac{\partial \phi^i}{\partial x_j} \tag{31.32}$$

The components of F_{ij} are, in conformity with (31.29) and (31.30):

$$F_{ij} = \begin{matrix} \\ i \downarrow \end{matrix} \begin{matrix} \quad j \rightarrow \\ \end{matrix} \begin{pmatrix} 0 & -cB_z & +cB_y & -E_x \\ +cB_z & 0 & -cB_x & -E_y \\ -cB_y & +cB_x & 0 & -E_z \\ +E_x & +E_y & +E_z & 0 \end{pmatrix} \tag{31.33}$$

and the components of the contravariant tensor F^{ij} are:

$$F^{ij} = \begin{matrix} \\ i \downarrow \end{matrix} \begin{matrix} \quad j \rightarrow \\ \end{matrix} \begin{pmatrix} 0 & -cB_z & +cB_y & +E_x \\ +cB_z & 0 & -cB_x & +E_y \\ -cB_y & +cB_x & 0 & +E_z \\ -E_x & -E_y & -E_z & 0 \end{pmatrix} = g^{in} g^{jm} F_{nm} \tag{31.34}$$

Note that (31.31) is in accordance with the requirement of gauge invariance (31.11), since:

$$\frac{\partial^2}{\partial x^i \partial x^j} = \frac{\partial^2}{\partial x^j \partial x^i}$$

Since it was possible to write the field equations covariantly in terms of the potentials, it should be possible to do so in terms of the fields. We can verify easily that the source equations:

(31.35) $$\vec{\nabla} \cdot \vec{E} = \frac{\rho}{k_o} \; ; \; \vec{\nabla} \times \vec{B} = \mu_o(\rho \vec{u} + k_o \frac{\partial \vec{E}}{\partial t})$$

correspond to:

(31.36) $$\frac{\partial F^{ji}}{\partial x^j} = \frac{j^i}{k_o}$$

and the equations:

(31.37) $$\vec{\nabla} \cdot \vec{B} = 0; \qquad \vec{\nabla} \times \vec{E} = - \frac{\partial \vec{B}}{\partial t}$$

correspond to:

(31.38) $$\frac{\partial F_{ij}}{\partial x^k} + \frac{\partial F_{jk}}{\partial x^i} + \frac{\partial F_{ki}}{\partial x^j} = 0$$

Note that the agreement between (31.36) and the equation of continuity (31.9) is obvious from the antisymmetric character of F^{ij}. Also note that despite the three free indices in Eq. (31.38), this equation represents only four non-vanishing equations, since the equation vanishes identically unless $i \neq j \neq k$. Eq. (31.38) can be written in a form resembling (31.36) by the introduction of the "dual" of F_{ij} by the relation:

(31.39) $$G^{ij} = P^{ijkl} F_{kl}$$

where P^{ijkl} is a tensor* defined to be zero unless $i \neq j \neq k \neq l$ and equal to +1 if ijkl are in cyclic order or have been permuted an even number of times from cyclic order and equal to 1 if ijkl are permuted an odd number of times from cyclic order. (31.38) is then equivalent to:

(31.40) $$\frac{\partial G^{ij}}{\partial x^i} = 0$$

This equation is written such as to permit the introduction of magnetic poles and pole-currents if they existed.

*P^{ijkl} strictly is not a tensor but a "pseudo-tensor"; it is not invariant to the "screw-sense" of the coordinate system. G^{ij} is thus a pseudo-tensor and any hypothetical pole 4-vector to be added to Eq. (31.40) is a pseudo-vector.

The tensor expression for the fields immediately permits a derivation of the transformation relations of the fields. Since from Equation (29.12):

$$F'^{ij} = Q^i_k \, Q^j_\ell \, F^{k\ell} \tag{31.41}$$

we can easily derive the relations:

$$E'_{\parallel} = E_{\parallel} \tag{31.42}$$

$$B'_{\parallel} = B_{\parallel} \tag{31.43}$$

$$\vec{E}'_{\perp} = \gamma(\vec{E} + \vec{v} \times \vec{B}) \tag{31.44}$$

$$\vec{B}'_{\perp} = \gamma(\vec{B} - \vec{v}/c^2 \times \vec{E})_{\perp} \tag{31.45}$$

where $\gamma = 1/\sqrt{1-\beta^2}$ and $E_{\parallel}$, $B_{\parallel}$ and $\vec{E}_{\perp}$, $\vec{B}_{\perp}$ are the components of $\vec{E}$ and $\vec{B}$ parallel and normal to $\vec{v}$, respectively.

Equations (31.44) and (31.45) can be interpreted fairly easily physically: the terms, other than the factor γ, being linear in $\vec{v}$ should be essentially classical, i.e. describable by Maxwell's equations without explicitly using relativistic arguments. Equation (31.44) corresponds to the fact that a particle moving relative to a magnetic field, to the order β^2, experiences an electric field $\vec{E} + \vec{v} \times \vec{B}$, as has been discussed in detail in Ch. 14. To interpret Eq. (31.45) consider a finite region containing lines of electric field as shown in Figure (31.2):

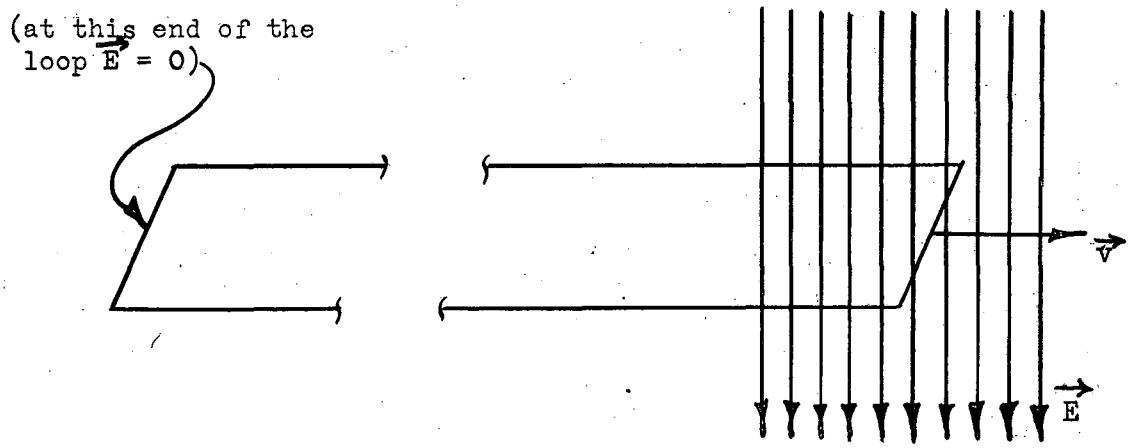


Figure (31.2)

Consider the line integral of $\vec{B}$ around the rectangle indicated. To an observer moving relative to $\vec{E}$, the flux of $\vec{E}$ through the rectangle is changing at the rate:

$$\int \vec{E} \cdot (d\vec{\ell} \times \vec{v}) = \oint (\vec{v} \times \vec{E}) \cdot d\vec{\ell} \tag{31.46}$$

and hence since the circulation of $\vec{B}$ is produced by the displacement current and the real current by the relation:

$$(31.47) \qquad \oint \vec{B} \cdot d\vec{\ell} = \frac{1}{c^2} \frac{d}{dt} \int_S \vec{E} \cdot d\vec{S} + \mu_0 I = \frac{1}{c^2} \oint (\vec{v} \times \vec{E}) \cdot d\vec{\ell} + \mu_0 I$$

the moving observer will observe an effective magnetic field:

$$(31.48) \qquad \vec{B}_{eff} = \vec{B} - \vec{v}/c^2 \times \vec{E}$$

in agreement with (31.45) for small $\vec{\beta}$.

The transformation equations for the fields are of considerable value in the solution of practical problems involving the motion of electrons and ions in electromagnetic fields. It is frequently possible to transform either the electric or the magnetic field away by choosing a suitable Lorentz frame.

The Lorentz force per unit volume:

$$(31.49) \qquad \vec{f} = \rho (\vec{E} + \vec{u} \times \vec{B})$$

is the space component for a four-vector:

$$(31.50) \qquad f^i = F^{ik} j_k = (\vec{f}, \frac{\vec{u} \cdot \vec{f}}{c})$$

The fourth component is 1/c times the power expended by the electric field/unit volume. The total force acting on a charge:

$$(31.51) \qquad \delta q = \rho \delta V$$

is not the space component of a four-vector, but is given by:

$$(31.52) \qquad \vec{F} = \vec{f} \delta V = \delta q (\vec{E} + \vec{u} \times \vec{B})$$

In a proper frame, since δq is an invariant (cf. 31.13):

$$(31.53) \qquad F_o = \delta q E_o$$

and hence in general, from (31.42) and (31.44):

$$(31.54) \qquad F_{\parallel} = F_{o\parallel}$$

$$(31.55) \qquad F_{\perp} = \gamma^{-1} F_{o\perp}$$

This is in agreement with the mechanical force transformations (30.31) ff. Hence equilibrium between mechanical forces and electrical forces is invariant to the choice of frame; the nature of the force does not affect its transformation

character. Also, we are justified in using the relation $\vec{F} = d\vec{p}/dt$ which we introduced as the definition of force, if for $\vec{F}$ we use the Lorentz force (31.52); Eq. (30.39) is thus the relativistically correct expression for the equation of motion of a charged particle in an electromagnetic field.

CHAPTER 32

COVARIANT FORMULATION OF THE ELECTROMAGNETIC FIELD EQUATIONS IN MATERIAL MEDIA

From the electron theory viewpoint, the ordinary Maxwell equations are the result of an averaging process of electrical quantities over regions large enough to permit macroscopic observation. We showed in the previous chapter that vacuum electrodynamics may be described in a simple covariant manner. It seems reasonable then that electrodynamics in material media may also be described in a single covariant form.

The principal element entering into macroscopic electrodynamics that is new is the fact that the current four-vector j^i will in general have all four components non-zero even in a frame in which the medium is at rest. In such a frame j^i will now have the components:

$$(j^i) = \left(\frac{\vec{j}^{\,o}}{c}, \rho^o\right) \tag{32.1}$$

where $\vec{j}^{\,o}$ is the current density in the proper frame. This would correspond to the vacuum definition. Let us now, however, consider the form of the components of j^i in a non-proper frame. In order to correspond to the vacuum case, we must assume that j^i retains the components:

$$(j^i) = \left(\frac{\vec{j}}{c}, \rho\right) \tag{32.2}$$

in any frame, and hence:

$$j_x = \gamma(j_x^o - \rho^o v) \tag{32.3}$$

$$\rho = \gamma\left(\rho^o - \frac{j_x^o v}{c^2}\right) \tag{32.4}$$

(32.3) is physically clear; it contains the convective current due to transport of charge and the contraction factor which assures the invariance of the charge. (32.4) is physically less obvious; it says that a substance which carries current but is electrically neutral ($\rho^o = 0$) in a proper frame does not necessarily remain so when observed from another inertial frame. This effect can be

understood in terms of the kinematics of the moving charges. Let us consider positive and negative charges in the proper frame $\sum^{o}$. For simplicity, let the positive charges be at rest (though this is not essential to the argument) and the negative charges be in motion. Let there be equal numbers of + and - charges in $\sum^{o}$. The world lines of the + charges (shown as dashes) and of the - charges (shown solid) as observed in $\sum^{o}$ and in an arbitrary frame $\sum$ are shown in the Minkowski diagram in Figure (32.1):

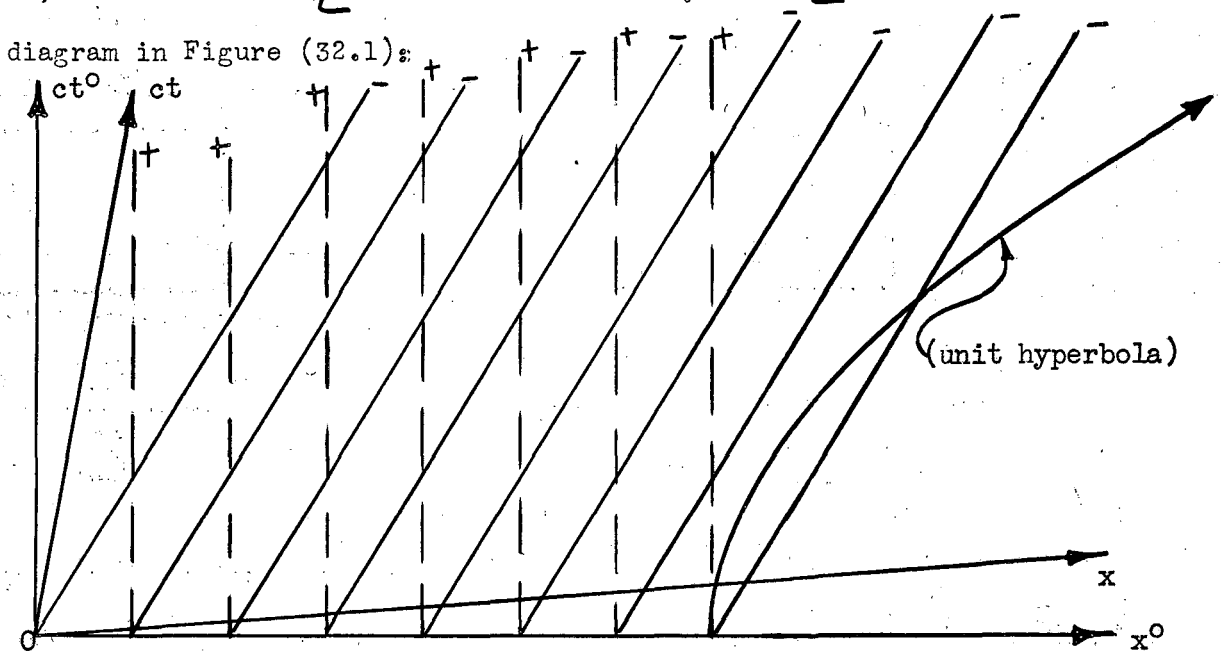


Figure (32.1)

The charge density in $\sum^{o}$ is measured by counting + and - charges "simultaneously" in $\sum^{o}$, i.e. on the x^{o} axis; the charge density in $\sum$ is measured by counting + and - charges along the x axis (i.e. averaging them simultaneously in $\sum$). Note that the density of - charges along the x-axis is decreased relative to the density along the x_{o}-axis, while the density of + charges has changed less.* Therefore a net positive charge is found in $\sum$,

---

* Note however that density of charge is measured by counting charges per unit length defined by the intercept of the unit hyperbola with the respective axes, as shown in Figure (32.1).

---

corresponding to the negative current in the neutral proper frame, in agreement with Eq. (32.4).

One consequence of this effect, resulting directly from the difference in the simultaneity measurement of $\sum^o$ and $\sum$, is the fact that a neutral stationary current loop in $\sum^o$ acquires an electric moment when observed in $\sum$. Consider a rectangular current loop in Figure (32.2) carrying a current I:

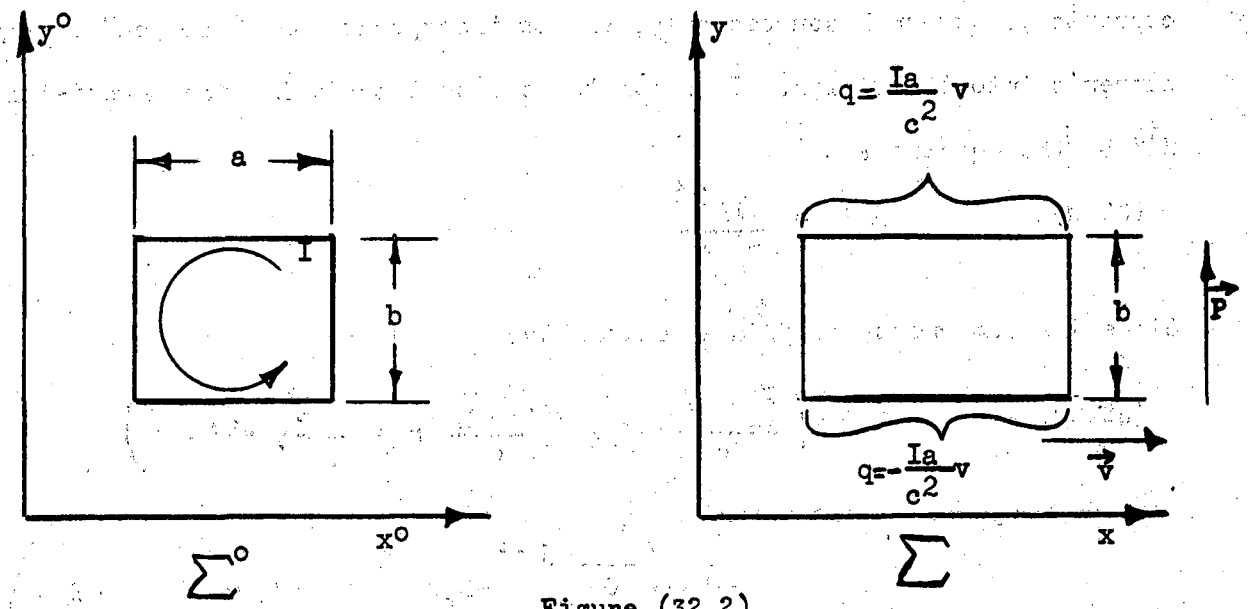


Figure (32.2)

From the point of view of $\sum$ the legs parallel to x^o will carry charges of $\pm \frac{Iav}{c^2}$ respectively; thus the system has an electric moment:

$$P = \frac{vabI}{c^2} = -\frac{|\vec{v} \times \vec{M}|}{c^2} \tag{32.5}$$

where $\vec{M} = Iab\vec{S}/S$ is the "magnetic moment" of the loop.* Quadratic effects

* $\vec{S}$ is the vector area of the loop, $|\vec{S}| = ab$.

in v have been ignored thus far. Shortly we shall obtain the exact transformation equation of the moments.

Let us now write the field equations in matter. If we divide the charge-current four-vector into "true" components j^i and magnetization-polarization components $j_M{}^i$, then Maxwell's equations (31.36) and (31.38) become:

$$\frac{\partial F^{il}}{\partial x^i} = \frac{j^l + j_M{}^l}{k_o} \tag{32.6}$$

and:

$$(32.7)\qquad \frac{\partial F_{ij}}{\partial x^k} + \frac{\partial F_{jk}}{\partial x^i} + \frac{\partial F_{ki}}{\partial x^j} = 0$$

where, as before, the F^{ij} are given by the matrix (31.34) in terms of $\vec{B}$ and $\vec{E}$. It is desirable, as before, to write (32.6) in terms of the true charge and currents as external sources only, and to incorporate the "induced" charge-currents into the fields. This can be done by introducing the moment-tensor M^{ij} by the equation:

$$(32.8)\qquad j_M{}^{l} = \frac{\partial M^{il}}{\partial x^i}$$

Since the components of $j_M^{\,l}$ are presumably:

$$(32.9)\qquad j_M{}^{l} = \left(\frac{\vec{j}_M + \vec{j}_P}{c}, \rho_P\right) = \left(\frac{\vec{\nabla} \times \vec{M}}{c} + \frac{1}{c}\frac{\partial \vec{P}}{\partial t}, -\vec{\nabla}\cdot\vec{P}\right)$$

this corresponds to the form:

$$(32.10)\qquad (M^{ij}) = \begin{array}{c} \\ i \downarrow \end{array} \begin{array}{c} \begin{array}{cccc} \multicolumn{4}{c}{j \longrightarrow} \end{array} \\ \begin{pmatrix} 0 & -M_z/c & +M_y/c & -P_x \\ +M_z/c & 0 & -M_x/c & -P_y \\ -M_y/c & +M_x/c & 0 & -P_z \\ +P_x & +P_y & +P_z & 0 \end{pmatrix} \end{array}$$

for the components of the moment tensor M^{ij}.

If we introduce a new field H^{ij} by:

$$(32.11)\qquad \frac{H^{ij}}{k_o} = F^{ij} - \frac{M^{ij}}{k_o}$$

corresponding to the three dimensional relations:

$$(32.12)^*\qquad \frac{\vec{H}}{ck_o} = cB - \frac{\vec{M}}{ck_o}$$

---

* I.e. $\vec{H} = \dfrac{\vec{B}}{\mu_o} - \vec{M}$

$$(32.13)\qquad \frac{\vec{D}}{k_o} = \vec{E} + \frac{\vec{P}}{k_o}$$

then the source equation (32.6) becomes simply:

$$\frac{\partial H^{ij}}{\partial x^i} = j^j \tag{32.14}$$

where:

$$(H^{ij}) = \begin{array}{c} i \downarrow \\ \\ \\ \\ \end{array} \overset{j \rightarrow}{\begin{pmatrix} 0 & -H_z/c & +H_y/c & +D_x \\ +H_z/c & 0 & -H_x/c & +D_y \\ -H_y/c & +H_x/c & 0 & +D_z \\ -D_x & -D_y & -D_z & 0 \end{pmatrix}} \tag{32.15}$$

Note the way in which the equivalent currents and charges are derived from the moments. The sign in the defining equations of the auxiliary fields also follows automotically as a result of this formulation.

The transformation properties of the moments follow directly from (32.10). We obtain:

$$P'_{\parallel} = P_{\parallel} \tag{32.16}$$

$$M'_{\parallel} = M_{\parallel} \tag{32.17}$$

$$\vec{P}'_{\perp} = \gamma\left(\vec{P} - \frac{\vec{v} \times \vec{M}}{c^2}\right)_{\perp} \tag{32.18}$$

$$\vec{M}'_{\perp} = \gamma(\vec{M} + \vec{v} \times \vec{P})_{\perp} \tag{32.19}$$

(32.16) is to be expected since $\vec{P}$ is the product of an (invariant) charge and a distance divided by a volume, if $\vec{P}$ is parallel to $\vec{v}$, then the distance and the volume contract in the same ratio. A similar argument applies to (32.17). The term $\vec{v} \times \vec{P}$ in Eq. (32.19) is a purely non-relativistic term arising from the fact that convection of a polarized medium corresponds to a net circulation of charge. We have met this effect before in Chapter 14, in the discussion of Maxwell's equations in moving media from a non-relativistic point of view. Consider an infinite polarized slab shown in Figure (32.3).

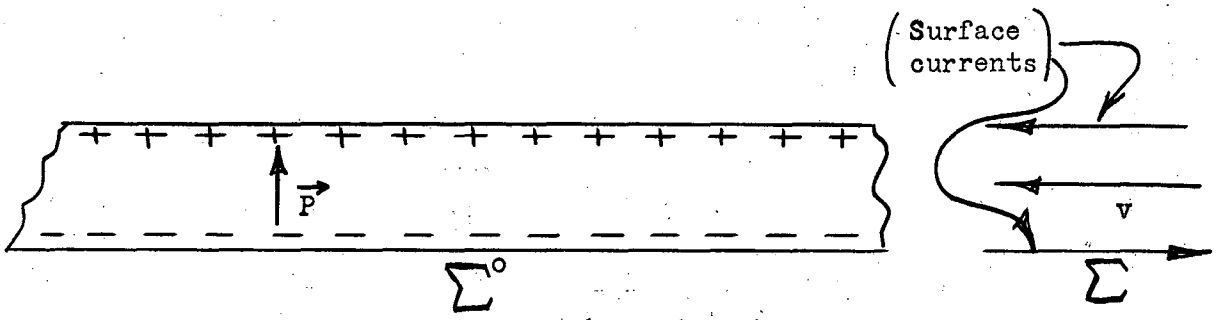


Figure (32.3)

From Σ, a moving frame, this slab possesses opposing surface currents corresponding to a uniformly magnetized medium.

Eq. (32.18) has no non-relativistic counterpart. It represents, as already shown, the effect of the net charges when a current in a neutral stationary conductor is viewed from a non-proper frame. The extra electric moment is precisely that predicted by Eq. (32.5) and arises from the relativistic definition of simultaneity. This equivalent electric moment resolves the apparent paradox of the uni-polar induction generator discussed briefly in Chapter 14. We concluded there that a current would flow when a conductor, in contact with take-off brushes, moves transverse to a magnetic field. No difficulties arise as long as the source of $\vec{B}$ is external; the effect should, however, persist if $\vec{B}$ is due to a permanent magnetization of the bar itself. The external description remains the same, but since the permanent magnetization is to be describable in terms of equivalent Amperian currents alone, the question arises how such currents could produce an electrostatic effect when viewed from a moving frame. This description leads to no difficulties, since we can now interpret the electrostatic field as due to the equivalent electric moment $\vec{v} \times \vec{M}/c^2$. Inasmuch as this equivalent moment is only a consequence of the relativistic redefinition of simultaneity, unipolar induction is fundamentally a relativistic effect.

CHAPTER 33

COVARIANT FORMULATION OF THE CONSERVATION LAWS OF ELECTRODYNAMICS

In Chapter 30, we discussed the conservation laws as applied to point mechanics. We found that the law of conservation of momentum and the law of conservation of energy are no longer independent concepts but are in fact parts of the same law. Now we would like to extend these concepts to vacuum electrodynamics.

The conservation laws of electrodynamics have the form (see Chapter 16) written in three-dimensional tensor language:

$$\frac{\partial N_i}{\partial x_i} + \rho E_i u_i = - k_o \frac{\partial}{\partial t}\left[\frac{(E^2 + c^2B^2)}{2}\right] \quad \text{(Conservation of energy)} \tag{33.1}$$

$$\frac{\partial}{\partial x_j}(\overset{\circ}{T}_{ij}) + \rho\left[E_i + (\vec{u} \times \vec{B})_i\right] = -k_o \frac{\partial}{\partial t}\left[(\vec{E} \times \vec{cB})_i\right] \quad \text{(Conservation of momentum)} \tag{33.2}$$

(Surface term) (Mechanical volume term) (Electromagnetic field volume term)

Here $\vec{N} = \vec{E} \times \vec{H}$ is the Poynting vector and $\overset{\circ}{T}_{ij} = E_iD_j + H_iB_j - \frac{\delta_{ij}}{2}(E_kD_k + H_kB_k)$ is the Maxwell stress tensor. Equation (33.1) balances energy between radiation loss over a given surface enclosing an arbitrary volume with the rate of mechanical and thermal work and the time rate of change of the electromagnetic field energy within that volume. Equation (33.2) balances the forces transmitted over a bounding surface by the electromagnetic field with the rate of increase of the mechanical and electromagnetic momentum of the system bounded by the volume.

These two equations can be combined into a single relation by introducing the energy-momentum tensor of the electromagnetic field defined by the symmetric matrix:

$$(33.3) \qquad (T^{ij}) = \begin{pmatrix} \overset{\circ}{T}_{11} & \overset{\circ}{T}_{12} & \overset{\circ}{T}_{13} & +cG_x \\ \overset{\circ}{T}_{12} & \overset{\circ}{T}_{22} & \overset{\circ}{T}_{23} & +cG_y \\ \overset{\circ}{T}_{13} & \overset{\circ}{T}_{23} & \overset{\circ}{T}_{33} & +cG_z \\ +cG_x & +cG_y & +cG_z & +W \end{pmatrix}$$

where $\vec{G} = \vec{N}/c^2$ is the momentum density of the field and $W = (k_o E^2 + \mu_o H^2)/2$ is the energy density. The conservation laws (33.1) and (33.2) are then equivalent to the simple covariant relation:

$$(33.4) \qquad \frac{\partial T^{ij}}{\partial x^i} = -f^j$$

where:

$$(33.5) \qquad (f^j) = \rho\left(\vec{E} + \vec{u} \times \vec{B}, \frac{\vec{E} \cdot \vec{u}}{c}\right)$$

is the four-vector representing the Lorentz force/unit volume and rate of work per unit volume of the electromagnetic field on material media.

To show that (33.4) is actually a tensor equation, we must not only show that it is correct algebraically but also that T^{ij} is a tensor. To show the latter, we note that T^{ij} can be generated from the tensor F^{ij} by the tensor operation:

$$(33.6) \qquad T^i{}_j = k_o\left[F^{ik}F_{jk} - \frac{1}{4}\,\delta^i_j F^{k\ell}\,F_{k\ell}\right]$$

Note that the second term is simply the invariant "trace" of $F^{k\ell}$:

$$(33.7)^* \qquad F^{k\ell}\,F_{k\ell} = 2(c^2B^2 - E^2)$$

---

* The invariance of $c^2B^2 - E^2$ shows that the ratio between the electric and magnetic fields in a plane electromagnetic wave is an invariant.

Therefore $T^i{}_j$ is a tensor and so is $T^{ij} = g^{jk}\,T^i_k$. The proof of the correctness of Equation (33.4) then remains an algebraic computation which can easily be done. We can draw a number of important conclusions from the form of the energy momentum tensor of the field.

1) Though the classical electromagnetic conservation laws are in agreement with

relativity without further modification, they are intimately connected with one another.

2) The transformation laws applied to (33.3) show again the equivalence of energy transport and momentum flow.

3) Let us consider a volume V containing totally a quantity of free electromagnetic radiation, but no charges or currents. The energy tensor thus obeys a "conservation law."

$$\frac{\partial T^{ij}}{\partial x^i} = 0 \tag{29.27}$$

and hence according to the theorem (29.28),

$$G^i = \left(\iiint c\,\vec{G}\,dv,\ \iiint W dv \right) = \iiint T^{4i}\,dv \tag{33.9}$$

is a contravariant four-vector also. Hence the momentum and energy of a radiation pulse totally contained within a finite volume has the same transformation properties as a material point particle. Note that this is not true for the total field of a charge, since (29.27) is not satisfied. We shall study this point in detail shortly. The invariant:

$$G^i G_i = W^2 - c^2 G^2 \tag{33.10}$$

is zero for a plane electromagnetic wave; hence the equivalent particle properties of such a wave correspond to zero rest mass. This is in agreement with the fact that, since the radiation is propagated with velocity c, it could obey the particle transformation laws and still yield finite momenta and energy only if its rest-mass is zero. All these facts are in agreement with the "light quantum" concept.

4) Since "state of equilibrium" is to be an invariant property of the system, we conclude that not only electromagnetic forces, energy, momenta, etc. must be describable by a tensor relation of the form $f^j = -\partial T^{ij}/\partial x^i$ but all such mechanical quantities when applied to a continuum must obey an equation of this form. Hence (33.4) will be a valid equation, e.g. in a medium under elastic

stress where W is the mass density (including elastic energy!), $\vec{G}$ is the mechanical momentum density, and $\overset{o}{T}_{ij}$ are the elastic stresses. This yields the transformation equations for all quantities entering into the mechanics of continuua. Note particularly that the mass density of a continuous medium cannot be treated as a scalar or even a component of a four-vector (as is the electrical charge density!) but is the (4,4) component of the mechanical energy-momentum tensor. The quantity $m = \iiint W dV$ has the transformation character of a mass as before. These facts are in agreement with the existence of a fundamental unit of charge and the apparent lack of existence of a fundamental unit of mass. The tensor component character of the mass density is of importance in the formulation of the gravitational action of matter in the general theory of relativity.

5) The "phase" ϕ of an electromagnetic wave is defined by the relation:

$$\vec{E} = \vec{E}_o e^{i\phi} = \vec{E}_o e^{i(\vec{k}\cdot\vec{r} - \omega t)} \qquad (33.11)$$

$|\vec{k}| = |\text{wave propagation vector}| = \lambda^{-1}$

ω = angular frequency

The zero point of a field must be an invariant physical fact and hence one would expect ϕ to be an invariant. Therefore, we can write:

$$\phi = k_i x^i \qquad (33.12)$$

where:

$$k_i = (\vec{k}, -\omega/c) \qquad (33.13)$$

$$k^i = -(\vec{k}, \omega/c) \qquad (33.14)$$

Since $k_i k^i = 0$; k^i transforms exactly like an energy momentum vector of a particle of zero rest mass. This is in agreement with making the momentum and energy of a light-quantum proportional to $\vec{k}$ and ω respectively; we saw in Conclusion 3 above that the momentum and energy of a electromagnetic wave also transforms like a particle of zero rest mass.

Equation (33.14) which defines the transformation character of $\vec{k}$ and ω provides a simple method to obtain the relativistically correct expressions for the

Doppler shift and for the aberration of star light.

Consider a source at rest in a frame $\sum^{o}$ radiating in the x direction such that $k_x^o = \omega^o/c$. In a $\sum$ frame we obtain from the Lorentz transformation applied to k^4:

$$\omega = c\gamma\left(\frac{\omega^o}{c} - \beta k_x^o\right) = \frac{\omega^o(1-\beta)}{\sqrt{1-\beta^2}} = \omega^o\sqrt{\frac{1-\beta}{1+\beta}} \tag{33.15}$$

This is the expression for the relativistic Doppler shift. The expression for aberration can be obtained similarly from the spatial components of k^i and gives the relation:

$$\sin\theta = \beta \tag{33.16}$$

for the aberration angle θ, in agreement with Equation (28.49).

6) It is to be noted that only the electromagnetic momentum and energy of a free wave have the transformation character of a particle of zero rest mass. For example, consider the electromagnetic field of a point particle moving with velocity $\vec{u}$ in the x direction. In the Lorentz frame in which the charge is at rest, the energy momentum tensor has only the components

$$T^{ij} = \begin{pmatrix} 0 & 0 \\ 0 & W_o \end{pmatrix} \qquad \alpha,\beta = 1,2,3 \tag{33.16}$$

(true only when integrated over a three-dimensional volume)

when W_o is the electrostatic field energy density of the charge. In the general frame it follows from the Lorentz transformation that

$$T^{14} = \frac{\beta(W_o - T_{11}^o)}{1-\beta^2} \tag{33.17}$$

$$T^{44} = \frac{W_o - \beta^2 T_{11}^o}{1-\beta^2} \tag{33.18}$$

Hence the momenta and energies of the field become:

$$cG_x = \int T^{14}dV = \int \sqrt{1-\beta^2}\, dV_o T^{14} = \frac{\beta}{\sqrt{1-\beta^2}}\int (W_o - T_{11}^o)dV_o \tag{33.19}$$

$$U = \int T^{44}dV = \int \sqrt{1-\beta^2}\, dV_o T^{44} = \frac{1}{\sqrt{1-\beta^2}}\int (W_o - \beta^2 T_{11}^o)dV_o \tag{33.20}$$

where dV_o is the volume element in the rest frame. But if the charge is spherically symmetrical in the rest frame, we have: (33.21) $\int W_o dV_o = U_o$ and:

$$(33.22) \qquad \int T_{11}^{o} dV_o = K_o \int (E_{xo}^{2} - \frac{1}{2} E_o^{2}) dV_o = -\frac{K_o}{6} \int E_o^{2} dV_o = -\frac{U_o}{3}$$

where U_o is the electrostatic energy of the charge, and where we have used the relation, applicable to a spherically symmetrical charge, that $\overline{E_x^2} = \frac{1}{3} E^2$. Hence we have:

$$(24.11) \qquad \vec{G} = \frac{4}{3} \frac{\vec{u}}{c^2} \frac{U_o}{\sqrt{1-\beta^2}}$$

$$(33.23) \qquad U = \frac{U_o}{\sqrt{1-\beta^2}} \left(1 + \frac{\beta^2}{3}\right)$$

These relations obviously do not transform as components of a 4-vector and therefore we obtain an additional argument for the fact, discussed at length in Ch. 24, that the electromagnetic mass of a particle cannot constitute the total mass of a charged particle.

It is of interest to write the equations of an electron in an external field in Hamiltonian form. These equations must of course be equivalent to the force equation:

$$(31.50) \qquad f^i = F^{ik} j_k$$

Let us describe the motion of the particle in 4-space as a function of the "proper time" s; i.e., let $x^i(s)$ and $p^i(s)$ be the positions and momenta of the particle. Let p^i be the momentum conjugate to x^i and not necessarily be equal to the p^i defined before. We are therefore looking for a function:

$$(33.24) \qquad \mathcal{H}(x^i, p^i; s)$$

which will yield an equation of motion equivalent to (31.50) of the form:

$$(33.25) \qquad \frac{dp_i}{ds} = -\frac{\partial \mathcal{H}}{\partial x^i}$$

$$(33.26) \qquad \frac{dx_i}{ds} = \frac{\partial \mathcal{H}}{\partial p^i} = u_i$$

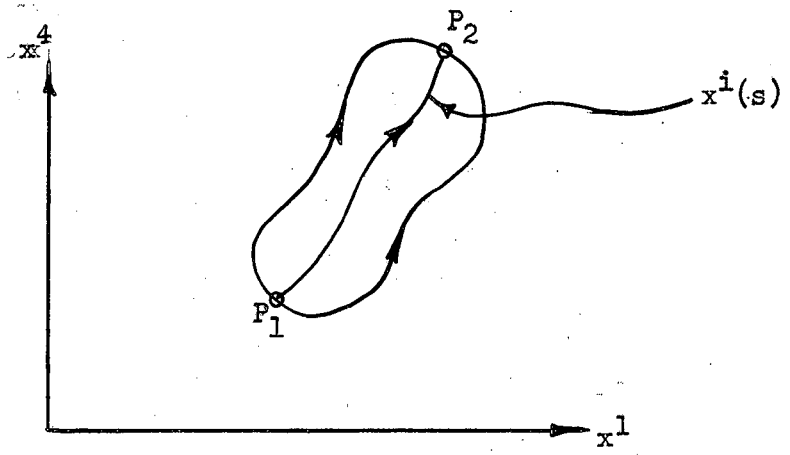
Let us derive this Hamiltonian from a Lagrangian L by the relation:

$$(33.27) \qquad \mathcal{H} = - L + u^i p_i$$

where the Lagrangian obeys a variational relation:

$$(33.28) \qquad \delta \int_{P_1}^{P_2} L ds = 0$$

The variation is to be taken over neighboring worldlines connecting two events P_1 and P_2 (Fig. 33.1).



Figure(33.1)

Consider a Lagrangian of the form

$$(33.29) \qquad L = + \frac{1}{2} m_o c^2 u^i u_i + e u^i \phi_i$$

Let us derive the Eulerian equation corresponding to (33.28):

$$(33.30) \qquad \delta \int_{P_1}^{P_2} L ds = \delta \int_{P_1}^{P_2} \left\{ \frac{1}{2} m_o c^2 u^i u_i + e u^i \phi_i \right\} ds$$

$$= \int \left\{ \left[m_o c^2 u_i + e \phi_i \right] \delta u^i + e u^j \frac{\partial \phi_j}{\partial x^i} \delta x^i \right\} ds$$

Since $\delta u^i = \delta \frac{dx^i}{ds} = \frac{d}{ds} \delta x^i$ and since the x^i are functions of s, we can write:

$$(33.31) \qquad \frac{d\phi_i}{ds} = \frac{\partial \phi_i}{\partial x^j} \frac{dx^j}{ds} = \frac{\partial \phi_i}{\partial x^j} u^j$$

We can integrate the first term in (33.30) by parts (noting that δx^j vanishes at the limits), giving:

$$(33.32) \qquad \int_{P_1}^{P_2} \left\{ \frac{d}{ds} \left[m_o c^2 u_i + e\phi_i \right] - eu^j \frac{\partial \phi_j}{\partial x^i} \right\} \delta x^i ds$$

Since this is to vanish for arbitrary δx^i, we obtain the Eulerian equations:

$$(33.33) \qquad m_o c^2 \frac{du_i}{ds} = -e \left\{ \frac{d\phi_i}{ds} - u^j \frac{\partial \phi_j}{\partial x^i} \right\} = eu^j \left(\frac{\partial \phi_j}{\partial x^i} - \frac{\partial \phi_i}{\partial x^j} \right)$$

Hence:

$$(33.34) \qquad m_o c^2 \frac{du^i}{ds} = eu_j F^{ij} = F^{ij}{}_{j}$$

in accordance with (31.50). Eq. (33.29) is thus a correct Lagrangian and

$$(33.35) \qquad P_i = \frac{\partial L}{\partial u^i} = m_o c^2 u_i + e\phi_i$$

are the momenta conjugate to x^i.

The Hamiltonian is:

$$(33.36) \qquad \mathcal{H} = -L + u^i p_i = \frac{m_o c^2}{2} u_i u^i = \frac{1}{2m_o c^2} \left\{ (p^i - e\phi^i)(p_i - e\phi_i) \right\}$$

This Hamiltonian again gives the correct equations of motion. From (33.25), (33.26), and (33.36), we have:

$$\frac{dx^i}{ds} = u^i = \frac{\partial \mathcal{H}}{\partial p_i} = \frac{1}{m_o c^2} (p^i - e\phi^i)$$

in agreement with (33.35). Also:

$$\frac{dp_i}{ds} = -\frac{\partial \mathcal{H}}{\partial x^i} = \frac{e}{m_o c^2} (p_j - e\phi_j) \frac{\partial \phi^j}{\partial x^i}$$

or, using (33.31):

$$m_o c^2 \frac{du_i}{ds} + eu^j \frac{\partial \phi_i}{\partial x^j} = eu^j \frac{\partial \phi_j}{\partial x^i}$$

$$m_o c^2 \frac{du^i}{ds} = eu_j F^{ij} \tag{33.37}$$

as before.

The Hamiltonian and Lagrangian functions discussed above are not the ordinary functions which use the time as an independent variable. We can, however, easily obtain the "conventional" Hamiltonian:

$$H = H(\vec{p}, \vec{x}, t) \tag{33.38}$$

in relativistically correct form from the "world" Hamiltonian.

$$\mathcal{H} = \mathcal{H}(p^i, x^i) = \mathcal{H}(\vec{p}, \vec{x}, p^4, ct) \tag{33.39}$$

by a simple calculation. We are looking for the function H which obeys the relations (3-dimensional)

$$-\frac{dx^i}{dx^4} = \frac{\partial H}{\partial p_i} \; ; \; \frac{dp_i}{dx^4} = \frac{\partial H}{\partial x^i} \; ; \quad i = 1, 2, 3 \tag{33.40}$$

(Note signs, since $p_i = (-c\vec{p}, p_4)$, $x^i = (\vec{x}, ct)$.)

We shall show now that if we solve the equation:

$$\mathcal{H}(\vec{p}, \vec{x}, p_4, ct) = \frac{1}{2} m_o c^2 u^i u_i = -\frac{1}{2} m_o c^2 \tag{33.41}$$

(since $u^i u_i = -1$), for $p_4 = p_4(\vec{p}, \vec{x}, ct)$, then p_4 has precisely the properties (33.40). From (33.25) and (33.26) we have:

$$\frac{dx^i}{dx^4} = -\frac{\partial H}{\partial p_i} = \frac{dx_i}{ds} \frac{ds}{dx^4} = \frac{dx_i}{ds} \frac{1}{\left(\frac{\partial \mathcal{H}}{\partial p_4}\right)} = \frac{-\frac{\partial \mathcal{H}}{\partial p_i}}{\frac{\partial \mathcal{H}}{\partial p_4}} \tag{33.42}$$

$$(33.43)\qquad \frac{dp_i}{dx^4} = \frac{\partial H}{\partial x^i} = \frac{dp_i}{ds}\,\frac{ds}{dx^4} = \frac{dp_i}{ds}\,\frac{1}{\left(\frac{\partial \mathcal{H}}{\partial p_4}\right)} = -\,\frac{\frac{\partial \mathcal{H}}{\partial x^i}}{\frac{\partial \mathcal{H}}{\partial p_4}}$$

Hence since $p_i = (-c\vec{p},\ p^4)$ and since from (33.41), considered solved for p^4,

$$\frac{\partial \mathcal{H}}{\partial p_i} + \frac{\partial \mathcal{H}}{\partial p_4}\,\frac{\partial p_4}{\partial p_i} = 0 \qquad i = 1, 2, 3$$

$$\frac{\partial \mathcal{H}}{\partial x_i} + \frac{\partial \mathcal{H}}{\partial p_4}\,\frac{\partial p_4}{\partial x_i} = 0$$

Equations (33.42) and (33.43) become:

$$(33.44)\qquad \frac{dx^i}{dt} = -\frac{\partial p_4}{\partial p_i} \qquad i = 1, 2, 3$$

$$\frac{dp_i}{dt} = \frac{\partial p_4}{\partial x^i}$$

Hence $p_4(\vec{x},\ \vec{p},\ t)$ is identical with $H(\vec{x},\ \vec{p},\ t)$. Hence:

$$(33.45)\qquad H(\vec{x},\ \vec{p},\ t) = e\phi + c\sqrt{(\vec{p} - e\vec{A})^2 + (m_o c)^2}$$

is the relativistically correct "ordinary" Hamiltonian describing the motion in ordinary 3-space as a function of time. In the absence of an electromagnetic field, this takes the simple form

$$(33.46)\qquad H(\vec{x},\ \vec{p},\ t) = \sqrt{(c\vec{p})^2 + (m_o c^2)^2}$$

in agreement with the expression for the total energy. If $m_o c \gg |\vec{p} - e\vec{A}|$, Eq. (33.45) takes the non-relativistic form:

$$(33.47)\qquad H(\vec{x},\ \vec{p},\ t) - m_o c^2 = e\phi + \frac{(\vec{p} - e\vec{A})^2}{2m_o}$$

Note that the relativistic Hamiltonian is given as a square root expression with the roots of either sign formally permitted. These negative roots give rise to

negative energy states when this Hamiltonian is used in quantum mechanics.

An interesting consequence of the general form of the Hamiltonian is the fact that classically the average value of a function depending on the coordinates only of a system in thermal equilibrium is independent of an external magnetic field. The mean value of such a function $f(x)$ is given by:

$$\bar{f} = \frac{\int f(\vec{x}) e^{-\frac{H(\vec{x},\vec{p})}{kT}} d\vec{x}d\vec{p}}{\int e^{-\frac{H(\vec{x},\vec{p})}{kT}} d\vec{x}d\vec{p}} \tag{33.48}$$

If we take the Hamiltonian in the form (33.45) or (33.47) with $\phi = 0$, we can make a transformation

$$\vec{p}' = \vec{p} - e\vec{A} \tag{33.49}$$

The Jacobian of this transformation is unity and hence, since the integration limits in (33.48) are infinite, the value of (33.48) becomes independent of $\vec{A}$ and hence $\vec{B}$. The existence of a magnetic susceptibility therefore implies the existence of states governed by other than the classical equations of motion.

CHAPTER 34 HAMILTONIAN FORMULATION OF MAXWELL'S EQUATIONS

In the last chapter we have formulated the equation of motion of a point particle in an external field in a covariant manner, using a Hamiltonian formulation. We now aim to formulate the equations of motion of the field, i.e. Maxwell's equations, in terms of such a system. In trying to do so, a new consideration affects the discussion, namely the fact that the number of degrees of freedom, in the mechanical sense, of a field is infinite. Before attempting such a formulation let us therefore study the transition, in Hamiltonian formulation, of a system of a N degrees of freedom to a system of an infinite number of degrees.

Let us consider a set of N point particles of equal mass m connected by a set of springs of lengths a and force constant k. Let η_i be the displacement from equilibrium of the i^{th} mass. (See Figure 34.1.)

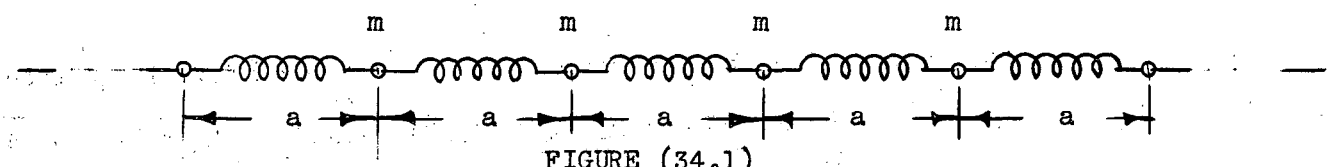


FIGURE (34.1)

The solution of this problem rests on finding a suitable Lagrangian L such that the equation:

$$(34.1)\quad \delta \int L(\eta_i, \dot{\eta}_i, t)\, dt = 0$$

represents the correct equation of motion. For a problem in classical mechanics we know that:

$$(34.2)\quad L = T - V$$

where T and V are the kinetic and potential energies respectively. By the geometry of the problem we thus have:

$$(34.3)\quad L = \frac{1}{2}\sum_{i=1}^{N} \left[m\, \dot{\eta}_i^2 - k(\eta_{i+1} - \eta_i)^2 \right]$$

which can be written in the form:

$$(34.3)\quad L = \sum_{i=1}^{N} a\, \mathcal{L}_i \quad , \text{ where:}$$

$$(34.5)\quad \mathcal{L}_i = \frac{1}{2}\left\{\frac{m}{a}\dot{\eta}_i^2 - (ka)\left(\frac{\eta_{i+1}-\eta_i}{a}\right)^2\right\}$$

is a quantity which we might call the "linear Lagrangian density". Equation (34.1) for the Lagrangian (34.3) gives then the usual equation of motion of a set of coupled oscillators; orthogonalization of (34.3) yields the set of normal modes.

Our principal interest here is to let the number of degrees of freedom tend to infinity; we shall do this by letting:

$$(34.6)\quad \begin{cases} a \longrightarrow dx \\ \dfrac{m}{a} \longrightarrow \mu & \text{(linear mass density)} \\ ka \longrightarrow Y & \text{(Young's modulus)} \\ \dfrac{\eta_{i+1}-\eta_i}{a} \longrightarrow \dfrac{\partial \eta}{\partial x} \end{cases}$$

We have thus replaced the discrete index $_i$ by the continuous variable x. The Lagrangian then becomes:

$$(34.7)\quad L = \frac{1}{2}\int\left[\mu\dot{\eta}^2 - Y\left(\frac{\partial\eta}{\partial x}\right)^2\right]dx = \int\mathcal{L}\,dx$$

where:

$$(34.8)\quad \mathcal{L}\left\{\eta, \dot{\eta}, \frac{\partial\eta}{\partial x}, t\right\} = \frac{1}{2}\left\{\mu\dot{\eta}^2 - Y\left(\frac{\partial\eta}{\partial x}\right)^2\right\}$$

The Lagrangian density acquires thus an explicit dependence on the spatial derivatives of the "field coordinate" η.

Let us derive the Eulerian equation of the variational principle (34.1). We obtain:

$$(34.9)\quad \delta\int L\,dt = \delta\iint\mathcal{L}\,dx\,dt$$

Substituting Eq. (34.8) and integrating by parts in the usual way, we obtain, ignoring explicit time dependence:

$$\delta\int L\,dt = \iint\left\{\frac{\partial\mathcal{L}}{\partial\eta}\delta\eta + \frac{\partial\mathcal{L}}{\partial\left[\frac{\partial\eta}{\partial x}\right]}\delta\left(\frac{\partial\eta}{\partial x}\right) + \frac{\partial\mathcal{L}}{\partial\left[\frac{\partial\eta}{\partial t}\right]}\delta\left(\frac{\partial\eta}{\partial t}\right)\right\}dx\,dt$$

$$(34.10)\qquad = \iint\left\{\frac{\partial\mathcal{L}}{\partial\eta} - \frac{\partial}{\partial x}\left[\frac{\partial\mathcal{L}}{\partial\left(\frac{\partial\eta}{\partial x}\right)}\right] - \frac{\partial}{\partial t}\left[\frac{\partial\mathcal{L}}{\partial\left(\frac{\partial\eta}{\partial t}\right)}\right]\right\}\delta\eta\,dx\,dt$$

where the integrated out part has been made to vanish by the condition that $\delta\eta = 0$ at the endpoints of integration over t. Since $\delta\eta$ is an arbitrary function of x, we thus obtain the partial differential equation:

$$(34.11)\quad \frac{\partial \mathcal{L}}{\partial \eta} - \frac{\partial}{\partial x}\left[\frac{\partial \mathcal{L}}{\partial\left(\frac{\partial \eta}{\partial x}\right)}\right] - \frac{\partial}{\partial t}\left[\frac{\partial \mathcal{L}}{\partial\left(\frac{\partial \eta}{\partial t}\right)}\right] = 0$$

which is often written in the form:

$$(34.12)\quad \frac{\delta \mathcal{L}}{\delta \eta} - \frac{\partial}{\partial t}\left[\frac{\partial \mathcal{L}}{\partial\left(\frac{\partial \eta}{\partial t}\right)}\right] = 0$$

where:

$$\frac{\delta \mathcal{L}}{\delta \eta} = \frac{\partial \mathcal{L}}{\partial \eta} - \frac{\partial}{\partial x}\left[\frac{\partial \mathcal{L}}{\partial\left(\frac{\partial \eta}{\partial x}\right)}\right]$$

is called the "variational" derivative. Note that in going to the limit of a continuous variable we have replaced a system of N ordinary differential Lagrangian equations by a partial differential equation.

Using the Lagrangian density (34.8), Eq. (34.12) yields immediately the wave equation:

$$(34.13)\quad \mu \frac{\partial^2 \eta}{\partial t^2} - Y \frac{\partial^2 \eta}{\partial x^2} = 0 \quad ,$$

corresponding to compressional waves traveling with velocity $\sqrt{Y/\mu}$.

Let us now generalize these considerations to a three dimensional field η, where η may be any covariant parameter. The action integral (34.1) then generalizes to:

$$(34.14)\quad \delta I = \delta \iint \mathcal{L}\, dv\, dt = 0$$

or:

$$(34.15)\quad \delta \iint \mathcal{L}\left\{\eta, \frac{\partial \eta}{\partial x^i}\right\} d^4x = 0$$

This formulation is evidently covariant. Let us again vary the functional dependence of $\mathcal{L}$ on η and $\frac{\partial \eta}{\partial x^i}$, but consider the x^i as fixed, independent, coordinates. Partial integration gives:

$$(34.16)\quad \int \left\{\frac{\partial \mathcal{L}}{\partial \eta} - \frac{\partial}{\partial x^i}\left[\frac{\partial \mathcal{L}}{\partial\left(\frac{\partial \eta}{\partial x^i}\right)}\right]\right\} d^4x\, \delta\eta = 0$$

leading to the covariant Lagrangian equations:

$$(34.17)\qquad \frac{\partial \mathcal{L}}{\partial \eta} = \frac{\partial}{\partial x^i}\left[\frac{\partial \mathcal{L}}{\partial\left(\frac{\partial \eta}{\partial x^i}\right)}\right] \qquad (i = 1, 2, 3, 4)$$

To obtain an explicit time dependence, these can be written in the form:

$$(34.18)\qquad \frac{\delta \mathcal{L}}{\delta \eta} = \frac{\partial}{\partial t}\left\{\frac{\partial \mathcal{L}}{\partial(\dot{\eta})}\right\} \qquad \left(\dot{\eta} = \frac{\partial \eta}{\partial t}\right)$$

where

$$(34.19)\qquad \frac{\delta \mathcal{L}}{\delta \eta} = \frac{\partial \mathcal{L}}{\partial \eta} - \frac{\partial}{\partial x^\alpha}\left[\frac{\partial \mathcal{L}}{\partial\left(\frac{\partial \eta}{\partial x^\alpha}\right)}\right] \qquad \alpha = 1, 2, 3.$$

Our program is thus to find an $\mathcal{L}$ such that Eq. (34.17) will lead to Maxwell's equations. Clearly this discussion will apply to any field theory.

It has been possible to state the action principle and write a Lagrangian partial differential equation in an evidently covariant form. To introduce a Hamiltonian, the time has to be singled out among the x^i as has been done in Eqs. (34.18) and (34.19). We can then define a "momentum density" conjugate to η:

$$(34.20)\qquad \pi(x,t) = \frac{\partial \mathcal{L}}{\partial\left(\frac{\partial \eta}{\partial t}\right)}$$

and a Hamiltonian density:

$$(34.21)\qquad \mathcal{H}(\eta, \pi; x, t) = \pi\left(\frac{\partial \eta}{\partial t}\right) - \mathcal{L}$$

The Hamiltonian equations follow in the usual way: Consider an increment dH of the total Hamiltonian $H = \int \mathcal{H}\, dv$:

$$(34.\ldots)\qquad dH = \int\left\{\dot{\eta}\, d\pi + \pi d\dot{\eta} - \frac{\partial \mathcal{L}}{\partial\left(\frac{\partial \eta}{\partial x^i}\right)} d\left(\frac{\partial \eta}{\partial x^i}\right) - \frac{\partial \mathcal{L}}{\partial \eta}\, d\eta\right\} dv$$

using (34.19) and (34.20) this becomes, on partial integration:

$$(34.22)\qquad dH = \int (\dot{\eta}\, d\pi - \dot{\pi}\, d\eta)\, dv$$

Since:

$$(34.23)\qquad H = \int \mathcal{H}\left\{\pi, \frac{\partial \pi}{\partial x^\alpha}, \eta, \frac{\partial \eta}{\partial x^\alpha}\right\} dv$$

$$(34.24)\qquad dH = \int\left\{\frac{\partial \mathcal{H}}{\partial \pi} d\pi + \frac{\partial \mathcal{H}}{\partial\left(\frac{\partial \pi}{\partial x^\alpha}\right)} d\left(\frac{\partial \pi}{\partial x^\alpha}\right) + \frac{\partial \mathcal{H}}{\partial \eta} d\eta + \frac{\partial \mathcal{H}}{\partial\left(\frac{\partial \eta}{\partial x^\alpha}\right)} d\left(\frac{\partial \eta}{\partial x^\alpha}\right)\right\} dv$$

Integrating by parts and using (34.23) and the notation used in (34.19), we obtain:

$$(34.25)\qquad \dot{\eta} = \frac{\delta \mathcal{H}}{\delta \pi} \qquad \dot{\pi} = -\frac{\delta \mathcal{H}}{\delta \eta}$$

as the new form of Hamilton's equations.

Hamilton's equations (34.25) lead to the usual meaning of the time rate of change in terms of Poisson brackets. If Λ is the density of a physical variable L, i.e. if $L = \int \Lambda \, dv$, then, using the process used in obtaining Eq. (34.25) we have:

$$(34.26)\qquad \frac{dL}{dt} = \int \left\{ \frac{\delta \Lambda}{\delta \eta} \dot{\eta} + \frac{\delta \Lambda}{\delta \pi} \dot{\pi} \right\} dv = \int \left\{ \frac{\delta \Lambda}{\delta \eta} \frac{\delta \mathcal{H}}{\delta \pi} - \frac{\delta \Lambda}{\delta \pi} \frac{\delta \mathcal{H}}{\delta \eta} \right\} dv$$

$$= \int [\Lambda, \mathcal{H}] dv$$

This is analogous to the usual Poisson bracket. These expressions lead to convenient starting points for quantization.

The discussion above leads to definite field equations if a Lagrangian density $\mathcal{L}$ is given. In order to lead to linear field equations, the Lagrangian must not contain powers of η or $\frac{\partial \eta}{\partial x^i}$ higher than the second. As the simplest example we might consider:

$$(34.27)\qquad \mathcal{L} = \frac{1}{2} \left\{ \frac{\partial \eta}{\partial x^i} \frac{\partial \eta}{\partial x_i} - \mu^2 \eta^2 \right\}$$

leading to the field equations (from (34.12)):

$$(34.28)\qquad \left[\Box - \mu^2 \right] \eta = 0$$

a momentum density:

$$(34.29)\qquad \pi = \frac{1}{c^2} \frac{\partial \eta}{\partial t}$$

and a positive definite Hamiltonian:

$$(34.30)\qquad \mathcal{H} = \frac{1}{2} \left\{ c^2 \pi^2 + (\vec{\nabla} \eta)^2 + \mu^2 \eta^2 \right\}$$

This is the scalar meson field of Yukawa, whose point solution is:

$$(34.31)\qquad \eta = e^{i(kr - \omega t)} \quad \frac{e^{-\mu r}}{r}$$

The electromagnetic field leads to a more complicated formulation. Presumably we are now dealing with a vector field, i.e.:

$$(34.32)\qquad \eta \rightarrow \phi^i$$

The Lagrangian will have 3 parts: 1) a mass term for the motion of material particles. 2) an interaction term between particle and field. 3) a field term corresponding to the field equation.

Terms (1) and (2) can be written from our former considerations: if we take

(34.33) $\mathcal{L}_1 = \frac{1}{2}\, g_o c^2 u_i u^i$ $\qquad g_o$ = proper mass density

(34.34) $\mathcal{L}_2 = j^i \phi_i$

we know that the correct motions of point particles in an external field result. In the choice of L_3 we might be guided by a classical analogy: let us choose L_3 such that 1) $\mathcal{L}_3$ is an invariant and 2) L_3 is a quantity analogous to the classical difference between potential and kinetic energy. In an electromagnetic oscillation energy oscillates between electric and magnetic energy, just as in a mechanical oscillation energy oscillates between kinetic and potential energy. We are thus led to take:

(34.35) $\mathcal{L}_3 = \frac{k_o(E^2 - c^2 B^2)}{2} = -\frac{k_o}{4}\, F^{ij} F_{ij}$

Note that the field equations:

(34.36) $\vec{\nabla} \cdot \vec{B} = 0 \qquad \vec{\nabla} \times \vec{E} = -\frac{\partial \vec{B}}{\partial t}$

are already implied in the connecting equations:

(34.37) $F^{ij} = \frac{\partial \phi^j}{\partial x_i} - \frac{\partial \phi^i}{\partial x_j}$ $\qquad$ corresponding to:

(34.38) $\vec{E} = -\vec{\nabla}\phi - \frac{\partial \vec{A}}{\partial t}$; $\vec{B} = \vec{\nabla} \times \vec{A}$

The total Lagrangian is thus:

(34.39) $\mathcal{L} = -\frac{k_o}{4} F^{ij} F_{ij} + e\, u^i \phi_i + \frac{1}{2} m_o c^2 u^i u_i$

A further motivation for a choice of the Lagrangian is the connection between Lagrangian density and the energy momentum tensor (Eq. (33.3)). If we consider the Hamiltonian density to retain the meaning of an energy density then we can put:

(34.40) $\frac{\partial \mathcal{H}}{\partial t} + \vec{\nabla} \cdot \vec{N} = 0$

$$\therefore \frac{\partial}{\partial t}(\mathcal{H}) = \frac{\partial}{\partial t}\left\{\dot{\eta}\frac{\partial \mathcal{L}}{\partial \dot{\eta}} - \mathcal{L}\right\} = \dot{\eta}\frac{\partial}{\partial t}\left(\frac{\partial \mathcal{L}}{\partial \dot{\eta}}\right) + \dot{\eta}\frac{\partial \mathcal{L}}{\partial \eta} - \frac{\partial \mathcal{L}}{\partial\left(\frac{\partial \eta}{\partial x^\alpha}\right)}\left(\frac{\partial^2 \eta}{\partial x^\alpha \partial t}\right)$$

(34.41)

$$= -\frac{\partial}{\partial x^\alpha}\left[\dot{\eta}\frac{\partial \mathcal{L}}{\partial\left(\frac{\partial \eta}{\partial x^\alpha}\right)}\right]$$

Hence from (34.40) we can put:

$$(34.42) \quad N^{\alpha} = c\, T^{4\alpha} = \dot{\eta}\, \frac{\partial \mathcal{L}}{\partial\left(\frac{\partial \eta}{\partial x^{\alpha}}\right)}$$

from which we can form the covariant generalization:

$$(34.43) \quad T^{\mu}_{\nu} = \frac{\partial \eta}{\partial x^{\nu}}\, \frac{\partial \mathcal{L}}{\partial\left(\frac{\partial \eta}{\partial x^{\mu}}\right)} - \mathcal{L}\delta^{\mu}_{\nu}$$

Note that the choice of Lagrangian fulfills both Eq. (34.42) and $T_4^{\ 4} = \mathcal{H}$ as required. The Lagrangian selected in Eq. (34.35) is thus in agreement with the stress tensor (33.6).

We can show easily that the Lagrangian $\mathcal{L}_2 + \mathcal{L}_3$ of Eq. (34.34)&(34.35) leads to Maxwell's equations. We have shown that $\mathcal{L}_1$ gives the correct motion of a particle in an external field (see Eqs. (33.28)ff); considering $\mathcal{L}_2 + \mathcal{L}_3$ separately is equivalent to ignoring the back reaction of the field on the motion of the particle. Treating the overall Lagrangian will lead to the difficulties already discussed previously in the delayed interaction calculations.

Note that this choice of $\mathcal{L}$ is not unique. Since the equations of motion will depend on the fields F^{ij} which in turn do not depend on the quantity $\frac{\partial \phi^i}{\partial x^i}$, any function of $\frac{\partial \phi^i}{\partial x^i}$ when added to $\mathcal{L}$ will still be a correct Lagrangian. It is customary but not mandatory to use the Lorentz condition $\frac{\partial \phi^i}{\partial x^i} = 0$.

From Eq. (34.17) we obtain, using $\mathcal{L} = \mathcal{L}_2 + \mathcal{L}_3$ and $\frac{\partial \mathcal{L}}{\partial \phi_j} = j^j$,

$$(34.44) \quad \frac{\partial}{\partial x^i}\, \frac{\partial \mathcal{L}}{\partial\left(\frac{\partial \phi_j}{\partial x^i}\right)} = \frac{k_0}{4}\, \frac{\partial}{\partial x^i} \left\{ \frac{\partial}{\partial\left(\frac{\partial \phi_j}{\partial x^i}\right)} \left[\left(\frac{\partial \phi^{\ell}}{\partial x_k} - \frac{\partial \phi^k}{\partial x_{\ell}}\right)\left(\frac{\partial \phi_{\ell}}{\partial x^k} - \frac{\partial \phi_r}{\partial x_{\ell}}\right) \right] \right\} = k_0 \frac{\partial}{\partial x^i} (F^{ij})$$

Hence:

$$(31.36) \quad \frac{\partial F^{ij}}{\partial x^i} = \frac{j^j}{k_0}$$

which are Maxwell's equations. This calculation can of course also be carried out using the three-dimensional form:

$$(34.45) \quad \mathcal{L} = \frac{k_0}{2} \left\{ (c^2 B^2 - E^2) - \vec{A} \cdot \vec{j} + \rho \phi \right\}$$

and Eq. (34.19).

In the three dimensional language we have:

(34.46) $\eta^{\alpha} = c\, A^{\alpha}$

(34.47) $\pi^{\alpha} = - k_o\, E^{\alpha}$

i.e., $\vec{A}$ and $\vec{E}$ are canonically conjugate. These variables give a Hamiltonian density:

$$(34.48)\quad \mathcal{H} = \frac{1}{2} k_o \left(\frac{\partial \vec{A}}{\partial t}\right)^2 + k_o\left(\vec{\nabla}\phi \cdot \frac{\partial \vec{A}}{\partial t}\right) + \frac{(\vec{\nabla}\times\vec{A})^2}{2\mu_o} - \frac{1}{2} k_o \left(\frac{\partial \vec{A}}{\partial t}\right)^2 - k_o\left(\vec{\nabla}\phi \cdot \frac{\partial \vec{A}}{\partial t}\right) - \frac{1}{2} k_o (\vec{\nabla}\phi)^2$$

$$= \frac{1}{2k_o} \vec{\pi}^2 + k_o (\vec{\nabla} \times \vec{\eta})^2 + (\vec{\pi} \cdot \vec{\nabla}\phi)$$

The last term can be made zero by a particular choice of gauge; at any rate since $\vec{\nabla}\cdot\vec{E} = 0$ the volume integral of the last term vanishes and thus does not contribute to the energy.

Hence:

$$(34.49)\quad H = \int \mathcal{H}\, dv = \frac{1}{2}\int \left[\frac{\vec{\pi}^2}{k_o} + k_o (\vec{\nabla}\times\vec{\eta})^2\right] dv = \frac{1}{2}\int \left(k_o \vec{E}^2 + \frac{\vec{B}^2}{\mu_o}\right) dv$$

as expected.

This method of defining canonical variables of the field clearly contains no physical information beyond the content of Maxwell's equation. The main reason for treating the subject is: a) the particle motions and the field equations are treated by the same procedure and b) transition to quantum electrodynamics is best accomplished through the canonical formulation of the field.

An alternative, and possibly physically more interesting, way of treating Maxwell's equations in Hamiltonian form is to make the number of degrees of freedom of the field finite by confining the field to a box of dimensions L. Let us take boundary conditions such that the field functions shall be periodic with period L in the three dimensions.

Let us take the Hamiltonian as:

$$(34.50)\quad H = \int\int\int \left\{\frac{k_o \vec{E}^2 + \mu_o \vec{H}^2}{2}\right\} dv = \int\int\int \left\{\frac{\vec{\pi}^2}{2k_o} + \frac{(\vec{\nabla}\times\vec{A})^2}{2\mu_o}\right\} dv$$

Since the solutions are periodic, they can be expanded in space by a Fourier

integral; i.e., if we put:

$$(34.51)\qquad \vec{u}_{k,\lambda} = \frac{1}{(L)^{3/2}}\ \vec{\mathcal{E}}_{k\lambda}\ e^{i(\vec{k}\cdot\vec{r})}$$

then any vector function can be expanded as a sum over $\vec{k} = \frac{2\pi}{L}(\ell\vec{i} + m\vec{j} + n\vec{i})$ and over the unit vectors $\vec{\mathcal{E}}_{k\lambda}$. The $\vec{\mathcal{E}}_{k\lambda}$ $(\lambda=1,2,3)$ permit an arbitrary choice of polarization.

As an example we can expand:

$$(34.52)\qquad \vec{A}(x^{\alpha},t) = \sqrt{\mu_o}\sum_{\lambda,k}\ \vec{u}_{k\lambda}\ (r)\ q_{k\lambda}\ (t)$$

$$(34.53)\qquad \vec{\pi}(x^{\alpha},t) = \sqrt{k_o}\sum_{\lambda,k}\ \vec{u}_{k\lambda}\ (r)\ p_{k\lambda}\ (t)$$

We can derive the auxiliary relations:

$$(34.54)\qquad \vec{\nabla}\cdot\vec{u}_{k\lambda} = \vec{k}\cdot\vec{\mathcal{E}}_{k\lambda}$$

$$\vec{\nabla}x\vec{u}_{k\lambda} = \vec{k}\ x\ \vec{\mathcal{E}}_{k\lambda}$$

Let us apply these relations to a pure, i.e. transverse radiation field. We can then choose a "gauge" $\phi = 0$ and thus take:

$$(34.55)\qquad \vec{\nabla}\cdot\vec{A} = 0$$

$$\vec{\nabla}\cdot\vec{E} = 0$$

In that case $\vec{k}\cdot\vec{\mathcal{E}}_{k\lambda} = 0$ and hence the summation over $\mathcal{E}_{k\lambda}$ will include only two components $\vec{\mathcal{E}}_{k1}$ and $\vec{\mathcal{E}}_{k2}$ for each $\vec{k}$ where $\vec{\mathcal{E}}_{k\lambda}$ is $\perp$ to $\vec{k}$. Hence:

$$(34.56)\qquad \vec{\nabla}x\ \vec{u}_{k\lambda_1} = i\ k\ \vec{u}_{k\lambda_2}$$

Hence:

$$(34.57)\qquad \int\pi^2\ dv = \frac{1}{L^3}\sum_{k}\sum_{k'} e^{i(\vec{k}+\vec{k}')\cdot\vec{r}}\ p_k\ p_{k'}\ dv$$

$$= \frac{1}{L^3}\int\sum_{k} p_k\ p_{-k}\ dv = \sum_{k} p_k\ p_{-k}$$

Since $\vec{\pi}$ is real, the sum (34.53) is equal to its complex conjugate; since $\vec{u}_{k\lambda} = \vec{u}^{*}_{-k\lambda}$ we must have: $p_k = p^{*}_{-k}$.

Hence:

$$(34.58)\qquad \int\pi^2\ dv = \sum_{k,\lambda}\ |\ p_{k\lambda}|^2$$

Also, from (34.56) and (34.52):

$$(34.59)\quad \int (\vec{\nabla} \times \vec{A})^2 \, dv = \sum_{k,\lambda} \left| i k q_{k\lambda} \right|^2 = \sum_{k,\lambda} k^2 \left| q_{k\lambda} \right|^2$$

Hence:

$$(34.60)\quad H = \frac{1}{2} \sum_{k,\lambda} \left\{ \left| p_{k\lambda}^2 \right|^2 + k^2 \left| q_{k\lambda} \right|^2 \right\}$$

This is the same as the Hamiltonian of a set of harmonic oscillators. Thus the equations of the electromagnetic field are equivalent to the equations of motion of a set of harmonic oscillators.