



Electrodynamics

1st Edition

Cenalo Vaz

University of Cincinnati

Contents

1	Vectors	1
1.1	Displacements	1
1.2	Linear Coordinate Transformations	6
1.3	Vectors and Scalars	8
1.4	Rotations in two dimensions	9
1.5	Rotations in three dimensions	11
1.6	Algebraic Operations on Vectors	14
1.6.1	The scalar product	14
1.6.2	The vector product	15
1.7	Vector Spaces	17
1.8	Some Algebraic Identities	18
1.9	Differentiation of Vectors	20
1.9.1	Time derivatives	20
1.9.2	The Gradient Operator	20
1.10	Some Differential Identities	23
1.11	Vector Integration	25
1.11.1	Line Integrals	25
1.11.2	Surface integrals	26
1.11.3	Volume Integrals	26
1.12	Integral Theorems	27
1.12.1	Corollaries of Stokes' Theorem	27
1.12.2	Corollaries of Gauss' theorem	28
2	Electrostatics: Introduction	30
2.1	Coulomb's Force Law	30
2.2	The Electric Field	32
2.3	Two Properties of the Electrostatic Field	34
2.4	Simple Applications of Gauss' Law	38
2.4.1	Point charge	39

2.4.2	Infinite line of charge of constant linear charge density:	39
2.4.3	Infinite sheet of charge of constant areal charge density:	40
2.4.4	Electric field of a conductor of arbitrary shape:	40
2.5	Poisson's Equation and Laplace's Equation	42
3	Boundary Value Problems: Laplace's Equation	45
3.1	Preliminaries	45
3.2	Boundary Conditions	45
3.3	Symmetries	49
3.3.1	Spherical Coordinates	49
3.3.2	Cylindrical coordinates	51
3.4	General Solutions of Laplace's Equation with symmetries	53
3.4.1	One dimensional solutions	53
3.4.2	Two dimensional solutions	54
3.4.3	Three dimensional solutions	54
3.5	Examples in three dimensions	56
3.5.1	Problems with Rectangular Symmetry	56
3.5.2	Problems with Cylindrical Symmetry	58
3.5.3	Problems with Spherical Symmetry	63
4	Boundary Value Problems: Poisson's Equation	68
4.1	The Method of Images	68
4.1.1	Point charge near infinite conducting planes.	69
4.1.2	Point charge outside a grounded, conducting sphere.	71
4.1.3	Point charge outside an insulated, conducting sphere.	74
4.1.4	Conducting sphere in a uniform external electric field.	74
4.2	The Green Function	75
4.3	Expansions of the Green function	79
4.3.1	Spherical Symmetry	79
4.3.2	Cylindrical Symmetry	82
4.4	Eigenfunction expansion of the Green function	84
5	Dielectric Media	87
5.1	Polarization	87
5.2	Electric field outside a Dielectric	89
5.3	Electric field inside a Dielectric	90
5.4	Gauss' Law inside Dielectrics	92
5.5	Boundary Conditions	94
5.6	Examples	97
5.6.1	Dielectric sphere in a uniform electric field	97

5.6.2	Images: Two dielectric media	99
5.7	Electrostatic Energy of a Charge distribution	101
5.7.1	Electrostatic self-energy of a discrete distribution	102
5.7.2	Electrostatic energy of a system of conductors	103
5.7.3	Electrostatic energy of a continuous distribution	106
5.8	Energy of a linear dielectric in an external field	108
5.9	Multipoles: The multipole expansion	110
6	Currents and the Magnetic Field	114
6.1	Current and Current Density	114
6.2	The Magnetic Field	116
6.3	Force and torque on a current carrying conductor	117
6.4	Magnetostatics: The Biot-Savart Law	119
6.5	Simple Applications	123
6.5.1	An infinitely long current carrying wire	123
6.5.2	A current loop	124
6.5.3	A solenoid	125
6.5.4	A Distant Circuit	126
6.6	Faraday's Law	127
6.7	Inductance	131
6.7.1	The toroidal coil	132
6.7.2	The Neumann Formula	133
7	Magnetic Media	135
7.1	Magnetization	136
7.2	Boundary Conditions	142
7.3	Examples	143
7.3.1	Magnetic sphere in a uniform magnetic field	143
7.3.2	Uniformly magnetized sphere	147
7.3.3	Magnetic shielding	148
7.4	Magnetic Energy	151
7.5	Maxwell's equations	153
8	Maxwell's Equations	156
8.1	A brief review	156
8.2	An alternative description	158
8.3	Gauge invariance	159
8.4	Choice of gauge	160
8.4.1	The Lorentz gauge	160
8.4.2	The Coulomb gauge	162

8.5	The homogeneous wave equation: a first look	164
9	Space-time Symmetries	168
9.1	Introduction	168
9.2	Lorentz Transformations	170
9.3	Tensors on the fly	174
9.4	Covariance of Maxwell's equations	182
9.5	Discrete Symmetries	188
9.5.1	Parity	188
9.5.2	Time reversal	190
9.5.3	Charge Conjugation	190
9.6	Electromagnetic Duality	190
10	More general coordinate systems*	193
10.1	Introduction	193
10.2	Vectors and Tensors	195
10.3	Differentiation	198
10.3.1	Lie Derivative	198
10.3.2	Covariant Derivative: the Connection	200
10.3.3	Absolute Derivative: parallel transport	204
10.3.4	The Laplacian	205
10.4	Examples	206
10.5	Integration: The Volume Element	211
11	Solutions of the Wave Equation	213
11.1	Green's functions	213
11.1.1	Spatial Green's function	213
11.1.2	Invariant Green's function	216
11.2	Mode Expansions	221
11.3	Boundary conditions	221
11.4	Poynting's theorem: energy and momentum density	224
12	Lagrangians and Hamiltonians*	228
12.1	Lagrangian description of mechanics	228
12.2	Noether's theorem in Mechanics	231
12.3	Hamiltonian description of Mechanics	234
12.4	Lagrangian description of fields	236
12.5	Noether's theorem for fields	239
12.6	Hamiltonian description of fields	243

13 The Electromagnetic Field and Sources*	247
13.1 Action for massive, free particles	247
13.2 Action for the Electromagnetic field	251
13.3 Inclusion of Sources	253
13.4 Conservation Laws	257
13.4.1 Energy and Momentum	257
13.4.2 Orbital Angular Momentum	259
13.4.3 Intrinsic Angular Momentum	260
13.5 Energy and Momentum with sources	261
14 The Homogeneous Wave Equation	264
14.1 Isotropic, linear, non-dispersive media	264
14.2 Non-conducting medium	266
14.2.1 Energy and momentum density	267
14.2.2 Polarization	268
14.3 Conducting medium	272
15 Interfaces between media	279
15.1 Introduction	279
15.2 Interfaces: non-conducting media	279
15.3 Interfaces: non-conducting/conducting medium	289
16 Wave Guides and Resonant Cavities	294
16.1 Introduction	294
16.2 Wave Guides	295
16.3 Resonant Cavities	304
17 Time Varying Charge Distributions	307
17.1 General Considerations	307
17.2 Current densities	310
17.3 Fields in the far zone: Spectral Distribution	312
17.4 Examples	314
17.4.1 Point Charge in Uniform motion	314
17.4.2 Point Charge in uniform circular motion: Synchrotron Radiation	314
17.4.3 Radiation from an oscillating Dipole	318
17.4.4 The Half-Wave Antenna	320
17.5 More General Distributions: Further Approximations	320
17.5.1 The Electric Dipole Approximation: Fields and Radiation	321
17.5.2 The Magnetic Dipole and Electric Quadrupole Approximation	322
17.6 The Multipole Expansion	324

18 The Liénard Wiechert Potentials	325
18.1 Potentials for a single point charge	325
18.2 The Field Strength Tensor	329
18.3 The Fields	331
18.3.1 Example: Larmor's formula	332
A The δ-function	i
A.1 Introduction	i
A.2 The δ -function as a distribution	ii
A.2.1 An example	iii
A.2.2 Another example	iv
A.2.3 Properties	v
A.3 The δ -function in curvilinear coordinates	vii

Chapter 1

Vectors

1.1 Displacements

Even though motion in mechanics is best described in terms of vectors, the formal study of vectors began only after the development of electromagnetic theory, when it was realized that they were essential to the problem of describing the electric and magnetic fields. However, vector analysis assumes an even more powerful role when it is used to implement a powerful principle of physics called the principle of **covariance**. This principle was first explicitly stated by Einstein as a fundamental postulate of the special theory of relativity. It requires the laws of physics to be independent of the features of any particular coordinate system, thereby lending a certain depth to the fundamental laws of physics and giving us a way to compare observations of physical phenomena by different observers using different coordinate frames. The great value of vector analysis lies in the fact that it clarifies the meaning of coordinate independence.

The concepts involved are most deeply rooted in mechanics and it is best to begin there. Motion in space involves dislodgements of a body from one point in space to another. Such dislodgements may occur smoothly along any curve passing through the two points, but the net effect of the motion is described by a directed line segment beginning at the initial position of the moving body and terminating at its final position, as shown in figure (1.1). If a body moves from the point labeled “ i ” in figure (1.1) to the point “ f ” then, no matter what the actual path traced by the body in going from i to f , we define its **displacement** as the directed straight line segment from i to f as shown. This directed line segment has both magnitude (its length) and direction (the arrow pointing from the initial position to the final position) and will be our prototypical vector. Thus, roughly speaking, a vector is a physical quantity that has both magnitude and direction in space and it may graphically be represented by a directed line segment. It is important to bear in mind that what defines a displacement is its magnitude and direction, not the actual

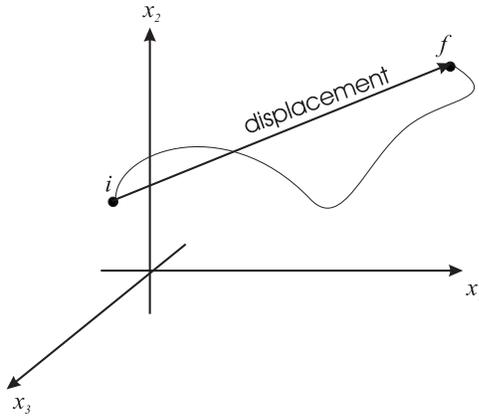


Figure 1.1: Displacement vector

initial and final points. Two displacements with the same magnitude and direction are identical, regardless of the initial and final points. So also, what defines a vector is its magnitude and direction and not its location in space.

We must now consider how displacements in particular and vectors in general are represented. In a two dimensional plane, we introduce two mutually perpendicular axes intersecting at some point O , the origin, order them in some way calling one the x axis and the other the y axis, and label points by an ordered pair, the coordinates (x, y) , where x represents the projection of the point on the x axis and y its projection on the y axis. A more fruitful way to think about this Cartesian coordinate system is to imagine that we have two mutually perpendicular and space filling one parameter families of parallel straight lines in the plane (see figure (1.2)). Because the families are space filling, every point will lie on the intersection of one vertical and one horizontal line. Label a point by the parameter values of the straight lines it lies on. Why is this a better way to think about coordinates? Because it is now easily generalized. Straight lines are not the only curves possible. We could also consider circles of radius r about an origin together with radial lines from the origin, each making an angle θ with some chosen direction [see figure (1.3)]. Every point in the plane lies on the intersection of a circle with some radial line and could therefore be labeled by the pair (r, θ) . These, of course, are the familiar polar coordinates. The system is ill defined at the origin.

The situation is similar in three dimensions, except that the curves are now replaced by surfaces. A coordinate system in three dimensions is a set of three independent, space filling, one parameter families of surfaces relative to which points are labeled. In the Cartesian system this set consists of three mutually perpendicular one parameter families of parallel planes. All points in \mathbb{R}^3 will lie on the intersection of a unique set of three

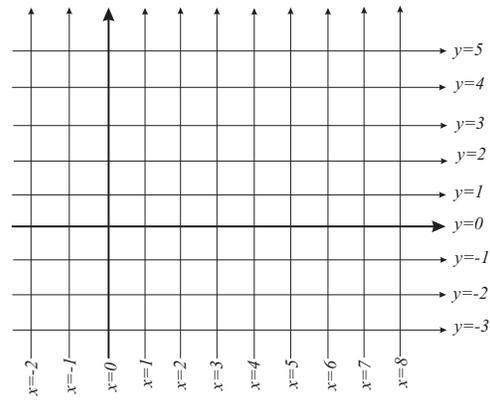


Figure 1.2: Cartesian coordinates in the plane.

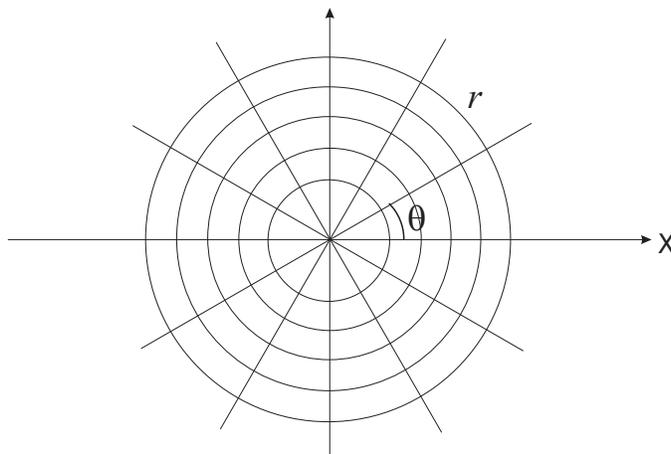


Figure 1.3: Polar coordinates in the plane.

planes and can be represented by an ordered triplet $\vec{r} = (x_1, x_2, x_3)$ consisting of the parameter values of the planes that intersect at the point in question. This is equivalent to the traditional way of viewing the Cartesian system as consisting of three mutually perpendicular straight lines, called coordinate axes, which intersect at a fixed point called the origin. The axes really represent the normals to the families of planes. They are ordered in some way and all points are represented by a correspondingly ordered set of three numbers, an ordered triplet $\vec{r} = (x_1, x_2, x_3)$, each number measuring the distance along the direction specified by one of the axes from the origin to the point in question.

Although we could consider finite displacements in \mathbb{R}^3 , it is sufficient and beneficial in the long run to restrict our attention to infinitesimal displacements. Introduce a Cartesian coordinate system and consider two points, i and f that are infinitesimally separated, with coordinates $\vec{r}_i = (x_1, x_2, x_3)$ and $\vec{r}_f = (x_1 + dx_1, x_2 + dx_2, x_3 + dx_3)$ respectively. The quantities dx_i represent displacements from i to f along the three (independent) coordinate directions. Let us represent the displacement, $d\vec{r}$, of a body moving from i to f as the ordered collection of these displacements (an ordered triplet),¹

$$d\vec{r} = (dx_1, dx_2, dx_3). \quad (1.1.1)$$

The numbers dx_i that form the triplet are called the **components** of $d\vec{r}$. The magnitude of $d\vec{r}$, denoted by $|d\vec{r}|$, is given by Pythagoras' theorem as

$$|d\vec{r}| = \sqrt{dx_1^2 + dx_2^2 + dx_3^2}. \quad (1.1.2)$$

Its direction can be specified by the angles that $d\vec{r}$ makes with the coordinate axes. Calling these angles α_1 , α_2 and α_3 respectively and applying Pythagoras' theorem again we find

$$dx_i = |d\vec{r}| \cos \alpha_i, \quad i \in \{1, 2, 3\}, \quad (1.1.3)$$

The cosines are called the **direction cosines** of the displacement. They are not all independent. Indeed, by substituting (1.1.3) into (1.1.2), one sees that they must satisfy the constraint

$$\sum_i \cos^2 \alpha_i = 1, \quad (1.1.4)$$

showing that one of the three angles is determined by the other two.

We will sometimes denote the i^{th} component, dx_i , of $d\vec{r}$ by $[d\vec{r}]_i$. The following definitions are natural:

- Two displacements are equal if they have the same magnitude and direction:

$$d\vec{r}_1 = d\vec{r}_2 \Leftrightarrow [d\vec{r}_1]_i = [d\vec{r}_2]_i \quad (1.1.5)$$

¹The arrow over the \vec{r} indicates that $d\vec{r}$ is not a number but an ordered triplet.

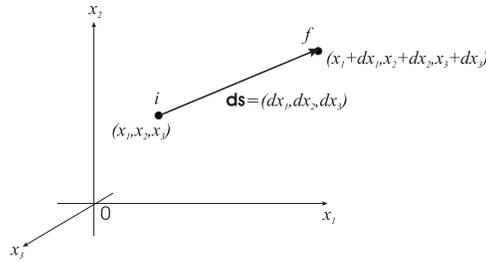


Figure 1.4: Representation of the displacement vector

- If a is a real number,

$$[a d\vec{r}]_i = a [d\vec{r}]_i \quad (1.1.6)$$

In particular, with $a = -1$, $[-d\vec{r}]_i = -[d\vec{r}]_i$.

- If $d\vec{r}_1$ and $d\vec{r}_2$ are two displacements then their sum is also a displacement given by

$$[d\vec{r}_1 + d\vec{r}_2]_i = [d\vec{r}]_{1,i} + [d\vec{r}]_{2,i} \quad (1.1.7)$$

(This definition can be understood as the algebraic equivalent of the familiar geometric parallelogram law of vector addition.)

Our implicit choice of coordinate system can be made explicit by assigning directions to the coordinate axes as follows: since every straight line is determined by two distinct points, on each axis choose two points one unit away from each other, in the direction of increasing coordinate value. There are only three corresponding displacements, which can be written as²

$$\hat{x}_1 = \hat{x} = (1, 0, 0), \quad \hat{x}_2 = \hat{y} = (0, 1, 0) \quad \text{and} \quad \hat{x}_3 = \hat{z} = (0, 0, 1) \quad (1.1.8)$$

and it is straightforward that, using the scalar multiplication rule (1.1.6) and the sum rule (1.1.7), any displacement $d\vec{r}$ could also be represented as

$$d\vec{r} = dx_1 \hat{x}_1 + dx_2 \hat{x}_2 + dx_3 \hat{x}_3 = \sum_i dx_i \hat{x}_i. \quad (1.1.9)$$

The \hat{x}_i represent the directions of the three axes of our chosen Cartesian system and the set $\{\hat{x}_i\}$ is called a **basis**.

²Carets, as opposed to arrows, are used to represent any displacement of unit magnitude.

In \mathbb{R}^3 , we could use the Cartesian coordinates of any point to represent its displacement from the origin. Displacements in \mathbb{R}^3 from the origin

$$\vec{r} = (x_1, x_2, x_3) = \sum_i x_i \hat{x}_i. \quad (1.1.10)$$

are called **position vectors**.

Because representations of position and displacement depend sensitively on the choice of coordinate axes whereas they themselves do not, we must distinguish between these objects and their representations. To see why this is important, we first examine how different Cartesian systems transform into one another.

1.2 Linear Coordinate Transformations

Two types of transformations exist between Cartesian frames, *viz.*, translations of the origin of coordinates and rotations of the axes. Translations are just constant shifts of the coordinate origin. If the origin, O , is shifted to the point O' whose coordinates are (x_O, y_O, z_O) , measured from O , the coordinates get likewise shifted, each by the corresponding constant,

$$x' = x - x_O, \quad y' = y - y_O, \quad z' = z - z_O \quad (1.2.1)$$

But since x_O , y_O and z_O are all constants, such a transformation does not change the representation of a displacement vector,

$$d\vec{r} = (dx, dy, dz) = (dx', dy', dz'). \quad (1.2.2)$$

Representations of displacement vectors are, however, affected by a rotation of the coordinate axes. Let us first consider rotations in two spatial dimensions [see figure (1.5)], where the primed axes are obtained from the original system by a rotation through some angle, θ . The coordinates (x_1, x_2) of a point P in the original system would be (x'_1, x'_2) in the rotated system. In particular, in terms of the length l of the hypotenuse OP of triangle AOP [figure (1.5)], we have

$$\begin{aligned} x_1 &= l \cos(\alpha + \theta) = (l \cos \alpha) \cos \theta - (l \sin \alpha) \sin \theta = x'_1 \cos \theta - x'_2 \sin \theta \\ x_2 &= l \sin(\alpha + \theta) = (l \sin \alpha) \cos \theta + (l \cos \alpha) \sin \theta = x'_2 \cos \theta + x'_1 \sin \theta. \end{aligned} \quad (1.2.3)$$

Inverting these relations gives

$$\begin{aligned} x'_1 &= x_1 \cos \theta + x_2 \sin \theta \\ x'_2 &= -x_1 \sin \theta + x_2 \cos \theta \end{aligned} \quad (1.2.4)$$

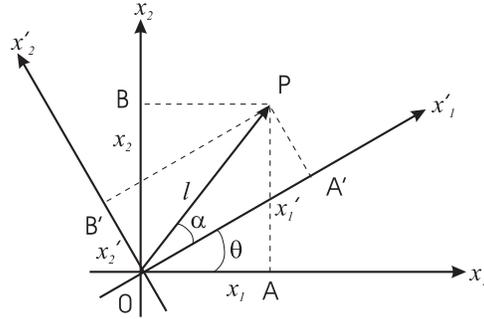


Figure 1.5: Rotations in two space dimensions

or, in terms of the components of an infinitesimal displacement from P ,

$$\begin{aligned} dx'_1 &= dx_1 \cos \theta + dx_2 \sin \theta \\ dx'_2 &= -dx_1 \sin \theta + dx_2 \cos \theta \end{aligned} \quad (1.2.5)$$

We could also exploit the representation given in (1.1.9) to obtain the same result. Let \hat{x}_1 and \hat{x}_2 designate the directions of the original x_1 and x_2 axes respectively and \hat{x}'_1 and \hat{x}'_2 the directions of the rotated axes. Then, because the displacement itself is independent of the coordinate system, we may write

$$d\vec{r} = dx_1 \hat{x}_1 + dx_2 \hat{x}_2 = dx'_1 \hat{x}'_1 + dx'_2 \hat{x}'_2 \quad (1.2.6)$$

Clearly, from figure (1.5)

$$\begin{aligned} \hat{x}'_1 &= \cos \theta \hat{x}_1 + \sin \theta \hat{x}_2 \\ \hat{x}'_2 &= -\sin \theta \hat{x}_1 + \cos \theta \hat{x}_2. \end{aligned} \quad (1.2.7)$$

Inserting these into the (1.2.6) we find

$$\begin{aligned} d\vec{r} = dx_1 \hat{x}_1 + dx_2 \hat{x}_2 &= dx'_1 (\cos \theta \hat{x}_1 + \sin \theta \hat{x}_2) + dx'_2 (-\sin \theta \hat{x}_1 + \cos \theta \hat{x}_2) \\ &= (dx'_1 \cos \theta - dx'_2 \sin \theta) \hat{x}_1 + (dx'_1 \sin \theta + dx'_2 \cos \theta) \hat{x}_2 \end{aligned} \quad (1.2.8)$$

A simple comparison now gives

$$dx_1 = dx'_1 \cos \theta - dx'_2 \sin \theta$$

$$dx_2 = dx'_1 \sin \theta + dx'_2 \cos \theta \quad (1.2.9)$$

or, upon inverting the relations,

$$\begin{aligned} dx'_1 &= dx_1 \cos \theta + dx_2 \sin \theta \\ dx'_2 &= -dx_1 \sin \theta + dx_2 \cos \theta. \end{aligned} \quad (1.2.10)$$

It is easy to see that these transformations can also be written in matrix form as

$$\begin{pmatrix} dx'_1 \\ dx'_2 \end{pmatrix} = \begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} dx_1 \\ dx_2 \end{pmatrix} \quad (1.2.11)$$

and

$$\begin{pmatrix} dx_1 \\ dx_2 \end{pmatrix} = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} dx'_1 \\ dx'_2 \end{pmatrix} \quad (1.2.12)$$

Other, more complicated but rigid transformations of the coordinate system can always be represented as combinations of rotations and translations.

1.3 Vectors and Scalars

Definition: A **vector** is a quantity that can be represented in a Cartesian system by an ordered triplet (A_1, A_2, A_3) of components, which transform as the components of an infinitesimal displacement under a rotation of the reference coordinate system.

In two dimensions, a vector may be represented by two Cartesian components $\vec{A} = (A_1, A_2)$, which transform under a rotation of the Cartesian reference system as $(A_1, A_2) \rightarrow (A'_1, A'_2)$ such that

$$\begin{pmatrix} A'_1 \\ A'_2 \end{pmatrix} = \begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} A_1 \\ A_2 \end{pmatrix} \quad (1.3.1)$$

Definition: A **scalar** is any physical quantity that *does not* transform (stays **invariant**) under a rotation of the reference coordinate system.

A typical scalar quantity in Newtonian mechanics would be the mass of a particle. The magnitude of a vector is also a scalar quantity, as we shall soon see. It is of great interest to determine scalar quantities in physics because these quantities are not sensitive to particular choices of coordinate systems and are therefore the same for all observers. Other examples of scalars within the context of Newtonian mechanics are temperature and density.

In the Newtonian conception of space and time, time is also a scalar. Because time is a scalar all quantities constructed from the position vector of a particle moving in space by taking derivatives with respect to time are also vectors, therefore

- the velocity: $\vec{v} = \frac{d\vec{r}}{dt}$
- the acceleration: $\vec{a} = \frac{d\vec{v}}{dt}$
- the momentum: $\vec{p} = m\vec{v}$ and
- the force $\vec{F} = \frac{d\vec{p}}{dt}$

are all examples of vectors that arise naturally in mechanics. In electromagnetism, the electric and magnetic fields are vectors. As an example of a quantity that has the appearance of a vector but is not a vector, consider $\mathbf{A} = (x, -y)$. Under a rotation of the coordinate system by an angle θ ,

$$\begin{aligned} A'_1 &= A_1 \cos \theta - A_2 \sin \theta \\ A'_2 &= A_1 \sin \theta + A_2 \cos \theta \end{aligned} \quad (1.3.2)$$

which are not consistent with (1.3.1). The lesson is that the transformation properties must always be checked.

1.4 Rotations in two dimensions

Equation (1.3.1) can also be written as follows

$$A'_i = \sum_j \hat{R}_{ij} A_j \quad (1.4.1)$$

where

$$\hat{R}_{ij}(\theta) = \begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix} \quad (1.4.2)$$

is just the two dimensional “rotation” matrix. We easily verify that it satisfies the following very interesting properties:

1. If we perform two successive rotations on a vector \vec{A} , so that after the first rotation

$$A_i \rightarrow A'_i = \sum_j \hat{R}_{ij}(\theta_1) A_j \quad (1.4.3)$$

and after the second rotation

$$A'_i \rightarrow A''_i = \sum_k \hat{R}_{ik}(\theta_2) A'_k = \sum_k \hat{R}_{ik}(\theta_2) \hat{R}_{kj}(\theta_1) A_j \quad (1.4.4)$$

then by explicit calculation it follows that

$$\sum_k \hat{R}_{ik}(\theta_2) \hat{R}_{kj}(\theta_1) = \hat{R}_{ij}(\theta_1 + \theta_2) \quad (1.4.5)$$

so

$$A_i'' = \sum_j \hat{R}_{ij}(\theta_1 + \theta_2) A_j \quad (1.4.6)$$

i.e., the result of two rotations is another rotation. The set of rotation matrices is therefore “closed” under matrix multiplication.

2. The unit matrix, $\mathbf{1}$, is the rotation matrix $\hat{R}(0)$.
3. The transpose of the rotation matrix whose angle is θ is the rotation matrix whose angle is $-\theta$. This follows easily from,

$$\hat{R}(-\theta) = \begin{pmatrix} \cos(-\theta) & \sin(-\theta) \\ -\sin(-\theta) & \cos(-\theta) \end{pmatrix} = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} = \hat{R}^T(\theta) \quad (1.4.7)$$

Now, using the closure property,

$$R^T(\theta) \cdot R(\theta) = R(-\theta) \cdot R(\theta) = R(0) = \mathbf{1} \quad (1.4.8)$$

Therefore, for every rotation matrix $\hat{R}(\theta)$ there exists an inverse, $\hat{R}(-\theta) = \hat{R}^T$.

4. Matrix multiplication is associative.

The rotation matrices therefore form a group under matrix multiplication.³ The group elements are all determined by one continuous parameter, the rotation angle θ . This is the commutative group, called $SO(2)$, of 2×2 orthogonal matrices with unit determinant, under matrix multiplication. We will now see that the situation gets vastly more complicated in the physically relevant case of three dimensions.

³Recall the following definitions:

Definition: The pair $(G, *)$ consisting of any set $G = \{g_1, g_2, \dots\}$ with a binary operation $*$ defined on it that obeys the four properties

- closure under $*$, *i.e.*, $\forall g_1, g_2 \in G \ g_1 * g_2 \in G$
- existence of an identity, *i.e.*, $\exists e \in G$ s.t. $\forall g \in G, \ g * e = e * g = g$
- existence of an inverse *i.e.*, $\forall g \in G \ \exists g^{-1} \in G$ s.t. $g * g^{-1} = g^{-1} * g = e$, and
- associativity of $*$, *i.e.*, $\forall g_1, g_2, g_3 \in G, \ g_1 * (g_2 * g_3) = (g_1 * g_2) * g_3$

is called a **group**.

Definition: If $\forall g_1, g_2 \in G, [g_1, g_2] = g_1 * g_2 - g_2 * g_1 = 0$ then the group $(G, *)$ is called a “commutative” or “Abelian” group. $[g_1, g_2]$ is called the commutator of the elements g_1 and g_2 .

1.5 Rotations in three dimensions

In two dimensions there is just one way to rotate the axes which, if we introduce a “ x_3 ” axis, amounts to a rotation of the $x_1 - x_2$ axes about it. In three dimensions there are three such rotations possible: the rotation of the $x_1 - x_2$ axes about the x_3 axis, the rotation of the $x_2 - x_3$ axes about the x_1 axis and the rotation of the $x_1 - x_3$ axes about the x_2 axis. In each of these rotations the axis of rotation remains fixed, and each rotation is obviously independent of the others. Thus, we now need 3×3 matrices and may write

$$\hat{R}^3(\theta) = \begin{pmatrix} \cos \theta & \sin \theta & 0 \\ -\sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad (1.5.1)$$

to represent the rotation of the $x_1 - x_2$ axes as before about the x_3 axis. Under such a rotation only the first and second component of a vector are transformed according to the rule

$$A'_i = \sum_j \hat{R}_{ij}^3(\theta) A_j \quad (1.5.2)$$

Rotations about the other two axes may be written likewise as follows:

$$\hat{R}^1(\theta) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \theta & \sin \theta \\ 0 & -\sin \theta & \cos \theta \end{pmatrix} \quad (1.5.3)$$

and⁴

$$\hat{R}^2(\theta) = \begin{pmatrix} \cos \theta & 0 & -\sin \theta \\ 0 & 1 & 0 \\ \sin \theta & 0 & \cos \theta \end{pmatrix} \quad (1.5.4)$$

The general rotation matrix in three dimensions may be constructed in many ways, one of which (originally due to Euler) is canonical:

- first rotate the (x_1, x_2) about the x_3 axis through an angle θ . This gives the new axes (ξ, η, τ) ($\tau \equiv z$),
- then rotate (ξ, η, z) about the ξ axis through an angle ϕ . This gives the new axes (ξ', η', τ') ($\xi' \equiv \xi$),
- finally rotate (ξ, η', τ') about the τ' axis through an angle ψ to get (x', y', z') .

⁴**Note** the change in sign. It is because we are using a right-handed coordinate system. Convince yourself that it should be so.

We get

$$\hat{R}(\theta, \phi, \psi) = \hat{R}^3(\psi) \cdot \hat{R}^2(\phi) \cdot \hat{R}^3(\theta) \quad (1.5.5)$$

The angles $\{\theta, \phi, \psi\}$ are called the **Euler angles** after the the originator of this particular sequence of rotations.⁵ The sequence is not unique however and there are many possible ways to make a general rotation. To count the number of ways, we need to keep in mind that three independent rotations are necessary:

- the first rotation can be performed in one of three ways, corresponding to the three independent rotations about the axes,
- the second rotation can be performed in one of two independent ways: we are not permitted to rotate about the axis around which the previous rotation was performed, and
- the third rotation can be performed in one of two independent ways: again we are not permitted to rotate about the axis around which the previous rotation was performed.

So in all there are $3 \times 2 \times 2 = 12$ possible combinations of rotations that will give the desired general rotation matrix in three dimensions. Note that any scheme you choose will involve three and only three independent angles, whereas only one angle was needed to define the general rotation matrix in two dimensions.

Three dimensional rotation matrices satisfy some interesting properties that we will now outline:

- The product of any two rotation matrices is also a rotation matrix.
- The identity matrix is just the rotation matrix $\hat{R}(0, 0, 0)$.
- All three dimensional rotation matrices, like their two dimensional counterparts, obey the condition

$$\hat{R}^T \cdot \hat{R} = \mathbf{1} \quad (1.5.6)$$

where \hat{R}^T is the transpose of \hat{R} , *i.e.*,

$$\hat{R}_{ij}^T = \hat{R}_{ji} \quad (1.5.7)$$

⁵Problem: Show that the general rotation matrix constructed with the Euler angles is

$$\hat{R}(\theta, \phi, \psi) = \begin{pmatrix} \cos \psi \cos \theta - \cos \phi \sin \theta \sin \psi & \cos \psi \sin \theta + \cos \phi \cos \theta \sin \psi & \sin \psi \sin \phi \\ -\sin \psi \cos \theta - \cos \phi \sin \theta \cos \psi & -\sin \psi \sin \theta + \cos \phi \cos \theta \cos \psi & \cos \psi \sin \phi \\ \sin \phi \sin \theta & -\sin \phi \cos \theta & \cos \phi \end{pmatrix}$$

The transpose of any rotation matrix is also a rotation matrix. It is obtained by applying the separate rotation matrices in reverse order. In the Euler parametrization,⁶

$$\hat{R}^T(\theta, \phi, \psi) = \hat{R}(-\psi, -\phi, -\theta) \quad (1.5.8)$$

Therefore, the transpose of a rotation matrix is its inverse.

- Finally, the associative property of matrix multiplication ensures that the product of rotations is associative.

The four properties listed above ensure that three dimensional rotations form a group under matrix multiplication. This the continuous, three parameter group called $SO(3)$ and is the group of all 3×3 orthogonal matrices of unit determinant, under matrix multiplication. The group is not commutative.

Rotations keep the magnitude of a vector invariant. Suppose \vec{A} has components (A_1, A_2, A_3) . Under a general rotation the components transform as

$$A'_i = \sum_j \hat{R}_{ij} A_j \quad (1.5.9)$$

Therefore,

$$\sum_i A'_i A'_i = \sum_{ijk} A_j \hat{R}_{ji}^T \hat{R}_{ik} A_k = \sum_{jk} A_j \left(\sum_i \hat{R}_{ji}^T \hat{R}_{ik} \right) A_k = \sum_{jk} A_j \delta_{jk} A_k = \sum_j A_j A_j \quad (1.5.10)$$

where δ_{jk} is the Kronecker δ ,⁷ and in the last step we use the fact that the transpose of \hat{R} is its inverse. $\sum_i A_i A_i$ is simply the length square of the vector $|\vec{A}|$, or its magnitude squared, *i.e.*,

$$|\vec{A}| = \sqrt{\sum_i A_i A_i} \quad (1.5.11)$$

is invariant under rotations.

⁶Problem: Verify this explicitly!

⁷Problem: The Kronecker δ is defined by

$$\delta_{ij} = \begin{cases} 0 & \text{if } i \neq j \\ 1 & \text{if } i = j \end{cases}$$

so it is the unit matrix. In fact, δ_{ij} is a “tensor”, *i.e.*, it transforms as two copies of a vector under rotations. Show this by showing that

$$\delta'_{ij} = \sum_{lk} \hat{R}_{il} \hat{R}_{jk} \delta_{lk} = \delta_{ij}.$$

1.6 Algebraic Operations on Vectors

We define

- Vector equality:

$$\vec{A} = \vec{B} \Leftrightarrow A_i = B_i, \quad \text{for all } i \quad (1.6.1)$$

- Scalar multiplication:

$$\vec{B} = a\vec{A} \Leftrightarrow B_i = cA_i, \quad \text{for } a \in \mathbb{R} \quad (1.6.2)$$

and

- Vector addition/subtraction:

$$\vec{C} = \vec{A} \pm \vec{B} \Leftrightarrow C_i = A_i \pm B_i \quad (1.6.3)$$

It is easy to show that the results of scalar multiplication, addition and subtraction are vectors (*i.e.*, having the correct transformation properties). Furthermore, there are two ways to define a product between two vectors.

1.6.1 The scalar product

The first is called the scalar (or inner, or dot) product and yields a scalar quantity. If \vec{A} and \vec{B} are two vectors,

$$\vec{A} \cdot \vec{B} = \sum_i A_i B_i \quad (1.6.4)$$

To show that $\vec{A} \cdot \vec{B}$ is a scalar under rotations, consider

$$\sum_i A'_i B'_i = \sum_{ijk} A_j \hat{R}_{ji}^T \hat{R}_{ik} B_k = \sum_{jk} A_j \delta_{jk} B_k = \sum_j A_j B_j. \quad (1.6.5)$$

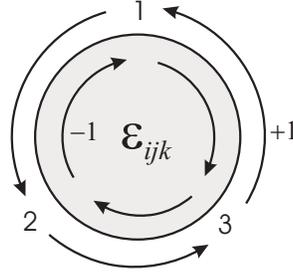
Notice that $|\vec{A}| = \sqrt{\vec{A} \cdot \vec{A}}$.

The basis vectors $\{\hat{x}_i\}$ satisfy $\hat{x}_i \cdot \hat{x}_j = \delta_{ij}$ and the component of a vector \vec{A} along any of the axes can be obtained from the scalar product of \vec{A} with the unit vector in the direction of the axis,

$$A_i = \vec{A} \cdot \hat{x}_i, \quad (1.6.6)$$

Since $A_i = |\vec{A}| \cos \alpha_i$, it can be used to define the direction cosines,

$$\cos \alpha_i = \frac{\vec{A} \cdot \hat{x}_i}{|\vec{A}|} = \frac{\vec{A} \cdot \hat{x}_i}{\sqrt{\vec{A} \cdot \vec{A}}} \quad (1.6.7)$$

Figure 1.6: Useful way to remember ϵ_{ijk}

Indeed, if \hat{u} is any unit vector, the component of \vec{A} in the direction of \hat{u} is $A_u = \vec{A} \cdot \hat{u}$. Because the scalar product is invariant under rotations, we prove this by letting α_i be the direction angles of \vec{A} and β_i be the direction angles of \hat{u} in the particular frame in which both \vec{A} and \hat{u} lie in the $x_1 - x_2$ plane (such a plane can always be found). Then $\alpha_3 = \beta_3 = \frac{\pi}{2}$ and

$$\vec{A} \cdot \hat{u} = |\vec{A}| \sum_i \cos \alpha_i \cos \beta_i = |\vec{A}| (\cos \alpha_1 \cos \beta_1 + \cos \alpha_2 \cos \beta_2) \quad (1.6.8)$$

In two dimensions, $\alpha_2 = \frac{\pi}{2} - \alpha_1$ and $\beta_2 = \frac{\pi}{2} - \beta_1$ so

$$\vec{A} \cdot \hat{u} = |\vec{A}| (\cos \alpha_1 \cos \beta_1 + \sin \alpha_1 \sin \beta_1) = |\vec{A}| \cos(\alpha_1 - \beta_1) = |\vec{A}| \cos \theta_u \quad (1.6.9)$$

where θ_u is the angle between \vec{A} and \hat{u} , because α_1 and β_1 are the angles made with the x axis. It follows, by Pythagoras' theorem, that A_u is the component of \vec{A} in the direction of \hat{u} . In a general coordinate frame, for any two vectors \vec{A} and \vec{B} ,

$$\vec{A} \cdot \vec{B} = |\vec{A}| |\vec{B}| \sum_i \cos \alpha_i \cos \beta_i = |\vec{A}| |\vec{B}| \cos \theta_{AB} \quad (1.6.10)$$

where θ_{AB} is the angle between \vec{A} and \vec{B} .

1.6.2 The vector product

The second product between two vectors yields another vector and is called the vector (or cross) product. If $\vec{C} = \vec{A} \times \vec{B}$, then

$$C_i = \sum_{jk} \epsilon_{ijk} A_j B_k \quad (1.6.11)$$

where we have introduced the three index quantity called the Levi-Civita tensor (density), defined by⁸

$$\epsilon_{ijk} = \begin{cases} +1 & \text{if } \{i, j, k\} \text{ is an even permutation of } \{1, 2, 3\} \\ -1 & \text{if } \{i, j, k\} \text{ is an odd permutation of } \{1, 2, 3\} \\ 0 & \text{if } \{i, j, k\} \text{ is not a permutation of } \{1, 2, 3\} \end{cases} \quad (1.6.12)$$

An useful mnemonic is shown in figure (1.6). One should check the following identities by direct computation

$$\begin{aligned} \sum_i \epsilon_{ijk} \epsilon_{irs} &= \delta_{jr} \delta_{ks} - \delta_{js} \delta_{kr} \\ \sum_{ij} \epsilon_{ijk} \epsilon_{ijs} &= 2\delta_{ks} \\ \sum_{ijk} \epsilon_{ijk} \epsilon_{ijk} &= 3! \end{aligned} \quad (1.6.13)$$

Note that the Levi-Civita symbol is antisymmetric under an interchange of its indices, eg., $\epsilon_{ijk} = -\epsilon_{ikj}$ etc. Using the above definition of the cross product, we could write out the components of $\vec{A} \times \vec{B}$ explicitly,

$$\vec{A} \times \vec{B} = (A_2 B_3 - A_3 B_2, A_3 B_1 - A_1 B_3, A_1 B_2 - A_2 B_1), \quad (1.6.14)$$

which is also obtained from the determinant form⁹

$$\vec{A} \times \vec{B} = \det \begin{vmatrix} \hat{x}_1 & \hat{x}_2 & \hat{x}_3 \\ A_1 & A_2 & A_3 \\ B_1 & B_2 & B_3 \end{vmatrix} \quad (1.6.16)$$

It is worth showing that the cross product is a vector. Since the Levi-Civita symbol transforms as a rank three tensor,

$$C'_i = \sum_{j,k} \epsilon'_{ijk} A'_j B'_k = \sum_{l,m,n,j,k,p,q} R_{il} R_{jm} R_{kn} R_{jp} R_{kq} \epsilon_{lmn} A_p B_q$$

⁸Prove that ϵ_{ijk} transforms as a rank three tensor, *i.e.*, according to *three* copies of a vector. Show that

$$\epsilon'_{ijk} = \sum_{lmn} \hat{R}_{il} \hat{R}_{jm} \hat{R}_{kn} \epsilon_{lmn} = \epsilon_{ijk}$$

provided that the rotation matrices are of unit determinant.

⁹The Levi-Civita symbol can be used to define the determinant of any 3×3 matrix as follows: if \hat{M} is a 3×3 matrix then

$$\det|\hat{M}| = \sum_{ijk} \epsilon_{ijk} \hat{M}_{1i} \hat{M}_{2j} \hat{M}_{3k} \quad (1.6.15)$$

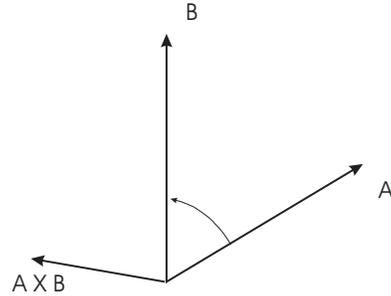


Figure 1.7: The right hand rule

$$= \sum_{l,m,n} R_{il} \epsilon_{lmn} A_m B_n = \sum_l R_{il} C_l \quad (1.6.17)$$

where we have used $\sum_k R_{kn} R_{kq} = \delta_{nq}$ and $\sum_j R_{jm} R_{jp} = \delta_{mp}$.

Notice that $\vec{A} \times \vec{A} = 0$ and that the basis vectors obey $\hat{x}_i \times \hat{x}_j = \epsilon_{ijk} \hat{x}_k$. In a coordinate frame that has been rotated so that both \vec{A} and \vec{B} lie in the $x_1 - x_2$ plane, using $\cos \alpha_2 = \sin \alpha_1$ and $\cos \beta_2 = \sin \beta_1$ together with $\cos \alpha_3 = \cos \beta_3 = 0$, we find that the only non-vanishing component of \vec{C} is C_3 given by

$$C_3 = |\vec{A}| |\vec{B}| (\cos \alpha_1 \sin \beta_1 - \sin \alpha_1 \cos \beta_1) = |\vec{A}| |\vec{B}| \sin(\beta_1 - \alpha_1) \quad (1.6.18)$$

If $\beta_1 > \alpha_1$, then C_3 is positive and \vec{C} points along the positive x_3 axis. On the contrary if $\beta_1 < \alpha_1$, then C_3 points along the negative x_3 axis. Because the magnitude of a vector independent of the frame, we conclude: in a general coordinate frame, the vector $\vec{A} \times \vec{B}$ has magnitude

$$|\vec{A} \times \vec{B}| = |\vec{A}| |\vec{B}| \left| \sum_{j,k} \epsilon_{ijk} \cos \alpha_j \cos \beta_k \right| = |\vec{A}| |\vec{B}| \sin |\theta_{AB}| \quad (1.6.19)$$

and direction given by the **right-hand rule**, which states that if the fingers of the right hand rotate \vec{A} into \vec{B} then the outstretched thumb points in the direction of \vec{C} (see figure (1.7)).

1.7 Vector Spaces

It is easy to verify that the set of all vectors in three dimensional space (\mathbb{R}^3), form an Abelian group under vector addition. The unit element is the zero vector and the inverse

of \vec{A} is $-\vec{A}$. Moreover, vector addition inherits its associativity from addition of ordinary numbers and is commutative. The space is also closed under scalar multiplication, since multiplying any vector by a real number gives another vector. Scalar multiplication is also

- associative,

$$a(b\vec{A}) = (ab)\vec{A}, \quad (1.7.1)$$

- distributive over vector addition,

$$a(\vec{A} + \vec{B}) = a\vec{A} + a\vec{B} \quad (1.7.2)$$

- as well as over scalar addition,

$$(a + b)\vec{A} = a\vec{A} + b\vec{A}, \quad (1.7.3)$$

- and admits an identity (1),

$$1(\vec{A}) = \vec{A} \quad (1.7.4)$$

In general, a vector space is any set that is a group under some binary operation (addition) over which multiplication by elements of a field, satisfying the four properties listed above, is defined.¹⁰ Although we have considered only scalar multiplication by real numbers, scalar multiplication by elements of any field (eg. the complex numbers or the rational numbers) is possible in general. The scalar and vector products we have defined are additional structures, not inherent to the definition of a vector space. The vectors we have introduced are geometric vectors in \mathbb{R}^3 .

1.8 Some Algebraic Identities

We turn to proving some simple but important identities involving the scalar and vector products. The examples given will not be exhaustive, but will serve to illustrate the general

¹⁰Definitions:

- A set of vectors, $\{\vec{A}_i\}$, is linearly independent if for scalars a_i ,

$$\sum_i a_i \vec{A}_i = 0 \Leftrightarrow a_i = 0 \forall i.$$

- A set of linearly independent vectors is **complete** if any vector in the vector space may be expressed as a linear combination of its members.
- A complete set of linearly independent vectors is said to form a **basis** for the vector space.
- The set of vectors $\hat{x}_1 = (1, 0, 0)$, $\hat{x}_2 = (0, 1, 0)$ and $\hat{x}_3 = (0, 0, 1)$, form a basis for \mathbb{R}^3 .

methods used. To simplify notation we will henceforth employ the following convention: if an index is repeated in any expression, it is automatically assumed that the index is to be summed over. Thus we will no longer write the sums explicitly (this is known as the Einstein summation convention).

1. $\vec{A} \times \vec{B} = -\vec{B} \times \vec{A}$.

We prove this for the components.

$$[\vec{A} \times \vec{B}]_i = \epsilon_{ijk} A_j B_k = \epsilon_{ikj} A_k B_j = \epsilon_{ikj} B_j A_k = -\epsilon_{ijk} B_j A_k = -[\vec{B} \times \vec{A}]_i$$

where, in the second step, we have simply renamed the indices by calling $j \leftrightarrow k$ which changes nothing as the indices j and k are summed over. In the next to last step we have used the fact that ϵ_{ijk} is antisymmetric in its indices, so that every interchange of indices in ϵ_{ijk} introduces a negative sign.

2. $\vec{A} \times (\vec{B} \times \vec{C}) = (\vec{A} \cdot \vec{C})\vec{B} - (\vec{A} \cdot \vec{B})\vec{C}$

Again take a look at the components,

$$\begin{aligned} [\vec{A} \times (\vec{B} \times \vec{C})]_i &= \epsilon_{ijk} A_j [\vec{B} \times \vec{C}]_k = \epsilon_{ijk} \epsilon_{klm} A_j B_l C_m \\ &= \epsilon_{ijk} \epsilon_{lmk} A_j B_l C_m = (\delta_{il} \delta_{jm} - \delta_{im} \delta_{jl}) A_j B_l C_m \\ &= (\vec{A} \cdot \vec{C}) B_i - (\vec{A} \cdot \vec{B}) C_i \end{aligned}$$

3. $(\vec{A} \times \vec{B}) \cdot (\vec{C} \times \vec{D}) = (\vec{A} \cdot \vec{C})(\vec{B} \cdot \vec{D}) - (\vec{A} \cdot \vec{D})(\vec{B} \cdot \vec{C})$

Write everything down in components. The left hand side is

$$\begin{aligned} (\vec{A} \times \vec{B}) \cdot (\vec{C} \times \vec{D}) &= \epsilon_{ijk} A_j B_k \epsilon_{ilm} C_l D_m = (\delta_{jl} \delta_{km} - \delta_{jm} \delta_{kl}) A_j B_k C_l D_m \\ &= (\vec{A} \cdot \vec{C})(\vec{B} \cdot \vec{D}) - (\vec{A} \cdot \vec{D})(\vec{B} \cdot \vec{C}) \end{aligned}$$

In particular, $(\vec{A} \times \vec{B})^2 = \vec{A}^2 \vec{B}^2 \sin^2 \theta$, where θ is the angle between \vec{A} and \vec{B} .

4. The triple product of three vectors \vec{A} , \vec{B} and \vec{C} is defined by

$$[\vec{A}, \vec{B}, \vec{C}] = \vec{A} \cdot (\vec{B} \times \vec{C}) = \epsilon_{ijk} A_i B_j C_k$$

This is a scalar.¹¹ It satisfies the following properties:

$$[\vec{A}, \vec{B}, \vec{C}] = [\vec{C}, \vec{A}, \vec{B}] = [\vec{B}, \vec{C}, \vec{A}] = -[\vec{B}, \vec{A}, \vec{C}] = -[\vec{C}, \vec{B}, \vec{A}] = -[\vec{A}, \vec{C}, \vec{B}] \quad (1.8.1)$$

i.e., the triple product is even under cyclic permutations and otherwise odd. Also $[\vec{A}, \vec{A}, \vec{B}] = 0$. All these properties follow directly from the properties of the Levi-Civita tensor density, ϵ_{ijk} .¹²

¹¹Problem: Verify this!

¹²Problem: Convince yourself that this is so.

$$5. (\vec{A} \times \vec{B}) \times (\vec{C} \times \vec{D}) = [\vec{A}, \vec{B}, \vec{D}]\vec{C} - [\vec{A}, \vec{B}, \vec{C}]\vec{D}$$

The left hand side is just

$$\begin{aligned} (\vec{A} \times \vec{B}) \times (\vec{C} \times \vec{D}) &= \epsilon_{ijk}\epsilon_{jrs}\epsilon_{kmn}A_rB_sC_mD_n = \epsilon_{jrs}(\delta_{im}\delta_{jn} - \delta_{in}\delta_{jm})A_rB_sC_mD_n \\ &= (\epsilon_{nrs}A_rB_sD_n)C_i - (\epsilon_{mrs}A_rB_sC_m)D_i \\ &= [\vec{A}, \vec{B}, \vec{D}]\vec{C} - [\vec{A}, \vec{B}, \vec{C}]\vec{D} \end{aligned}$$

1.9 Differentiation of Vectors

1.9.1 Time derivatives

A vector function of time is a vector whose components are functions of time. The derivative of a vector function of time is then defined in a natural way in terms of the derivatives of its components in the Cartesian basis. Let $\vec{A}(t)$ be a vector function of some parameter t , *i.e.*,

$$\vec{A}(t) = (A_1(t), A_2(t), A_3(t)) \quad (1.9.1)$$

The derivative of $\vec{A}(t)$ is another vector function, $\vec{C}(t)$, whose Cartesian components are given by

$$C_i = \frac{dA_i}{dt} \quad (1.9.2)$$

Note that the above definition is “good” only in for the Cartesian components of the vector. This is because the Cartesian basis $\{\hat{x}_i\}$ is **rigid**. In more general coordinate systems where the basis is not rigid, the derivative of a vector must be handled delicately. We will return to this later. Here, we will convince ourselves that \vec{C} is really a vector. Under a rotation

$$A_i \rightarrow A'_i \Rightarrow C'_i = \frac{dA'_i}{dt} = \frac{d}{dt}(\hat{R}_{ij}A_j) = \hat{R}_{ij}\frac{dA_j}{dt} = \hat{R}_{ij}C_j \quad (1.9.3)$$

which shows that $\vec{C}(t)$ has the correct transformation properties, inherited from $\vec{A}(t)$. This justifies the statement that the velocity, momentum, acceleration and force must all be vectors, because they are all obtained by differentiating $\vec{r}(t)$.

1.9.2 The Gradient Operator

The gradient operator is a vector **differential operator**, whose definition is motivated by a simple geometric fact. Consider some scalar function $\phi(\vec{r})$ ¹³ and the surface in \mathbb{R}^3 ,

¹³Any scalar function $\phi(\vec{r}, t)$ is called a **scalar field**.

defined by

$$\phi(\vec{r}) = \phi(x_1, x_2, x_3) = \text{const.}, \quad (1.9.4)$$

so that

$$\phi'(x'_1, x'_2, x'_3) = \phi(x_1, x_2, x_3) \quad (1.9.5)$$

By the fundamental theorem of calculus

$$d\phi = \frac{\partial\phi}{\partial x_1} dx_1 + \frac{\partial\phi}{\partial x_2} dx_2 + \frac{\partial\phi}{\partial x_3} dx_3 = \left(\frac{\partial\phi}{\partial x_1}, \frac{\partial\phi}{\partial x_2}, \frac{\partial\phi}{\partial x_3} \right) \cdot (dx_1, dx_2, dx_3) = 0 \quad (1.9.6)$$

The vector (dx_1, dx_2, dx_3) represents an infinitesimal displacement on the surface determined by the equation $\phi(x_1, x_2, x_3) = \text{const.}$ The other vector in the last scalar product is called the gradient of the function ϕ ,

$$\vec{\nabla}\phi = \left(\frac{\partial\phi}{\partial x_1}, \frac{\partial\phi}{\partial x_2}, \frac{\partial\phi}{\partial x_3} \right) \quad (1.9.7)$$

It has the form of a vector, but we need to check of course that its transformation properties under rotations are those of a vector. We will therefore look at the components of $\vec{\nabla}\phi$:

$$\frac{\partial\phi}{\partial x_i} = \partial_i\phi \rightarrow \partial'_i\phi' = \frac{\partial\phi'}{\partial x'_i} = \frac{\partial\phi}{\partial x_j} \frac{\partial x_j}{\partial x'_i} \quad (1.9.8)$$

Now

$$x'_i = \hat{R}_{ik} x_k \Rightarrow x_j = \hat{R}_{ji}^T x_i \quad (1.9.9)$$

so

$$\frac{\partial x_j}{\partial x'_i} = \hat{R}_{ji}^T = \hat{R}_{ij} \quad (1.9.10)$$

and therefore

$$\partial'_i\phi' = \frac{\partial\phi'}{\partial x'_i} = \hat{R}_{ij} \partial_j\phi \quad (1.9.11)$$

which is indeed the vector transformation law. Hence $\vec{\nabla}\phi$ is a vector if $\phi(\vec{r})$ is a scalar. Now it turns out that $\vec{\nabla}\phi$ has a nice geometric meaning. Because

$$\vec{\nabla}\phi \cdot d\vec{r} = 0 \quad (1.9.12)$$

for all infinitesimal displacements along the surface, it follows that $\vec{\nabla}\phi$, if it is not vanishing, must be *normal* to the surface given by $\phi(\vec{r}) = \text{const.}$ Thus, given any surface $\phi(\vec{r}) = \text{const.}$,

$$\hat{n} = \frac{\vec{\nabla}\phi}{|\vec{\nabla}\phi|} \quad (1.9.13)$$

is the unit normal to the surface.

Example: Take $\phi(x, y, z) = x^2 + y^2 + z^2$, then $\phi(\vec{r}) = \text{const.}$ represents a sphere centered at the origin of coordinates. The unit normal to the sphere at any point is

$$\vec{\nabla}\phi = \frac{\vec{r}}{r} \quad (1.9.14)$$

where r is the radius of the sphere and \vec{r} is the position vector of the point. The normal to the sphere is therefore in the direction of the radius.

Example: Take $\phi(x, y, z) = \frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2}$, so that $\phi(x, y, z) = 1$ represents an ellipsoid with semi-axes of lengths a , b and c respectively. We find

$$\hat{n} = \left(\frac{x}{a}, \frac{y}{b}, \frac{z}{c}\right) \quad (1.9.15)$$

which is the normal to the ellipsoid at the point (x, y, z) .

We see that $\vec{\nabla}$ is just the derivative operator in the Cartesian system, so we can think of it in component form as the collection of derivatives,

$$[\vec{\nabla}]_i = \partial_i. \quad (1.9.16)$$

Now if we introduce the concept of a **vector field** as a vector function of space and time,

$$\vec{A}(\vec{r}, t) = (A_1(\vec{r}, t), A_2(\vec{r}, t), A_3(\vec{r}, t)) \quad (1.9.17)$$

then we can define two distinct operations on $\vec{A}(\vec{r})$ using the scalar and vector products given earlier,

- the **divergence** of a vector field $\vec{A}(\vec{r}, t)$ as

$$\text{div}\vec{A} = \vec{\nabla} \cdot \vec{A} = \partial_i A_i \quad (1.9.18)$$

and

- the **curl** (or **rotation**) of a vector field $\vec{A}(\vec{r}, t)$ as

$$[\vec{\nabla} \times \vec{A}]_i = \epsilon_{ijk} \partial_j A_k \quad (1.9.19)$$

These turn out to be of fundamental importance in any field theory, electromagnetism being one of them. We will understand their physical significance in the following chapters. For now, we only prove a few identities involving the $\vec{\nabla}$ operator. Once again, the examples given are far from exhaustive, their purpose being only to illustrate the method.

1.10 Some Differential Identities

1. $\vec{\nabla} \cdot \vec{r} = 3$

This follows directly from the definition of the divergence,

$$\vec{\nabla} \cdot \vec{r} = \partial_i x_i = \delta_{ii} = 3$$

2. $\vec{\nabla} \cdot (\phi \vec{A}) = (\vec{\nabla} \phi) \cdot \vec{A} + \phi (\nabla \cdot \vec{A})$

Expand the l.h.s to get

$$\partial_i (\phi A_i) = (\partial_i \phi) A_i + \phi (\partial_i A_i) = (\vec{\nabla} \phi) \cdot \vec{A} + \phi (\nabla \cdot \vec{A}) \quad (1.10.1)$$

As a special case, take $\vec{A} = \vec{r}$, then $\vec{\nabla} \cdot (\vec{r} \phi) = \vec{r} \cdot (\vec{\nabla} \phi) + 3\phi$

3. $\vec{\nabla} \cdot (\vec{\nabla} \times \vec{A}) \equiv 0$

The proof is straightforward and relies on the antisymmetry of ϵ_{ijk} :

$$\vec{\nabla} \cdot (\vec{\nabla} \times \vec{A}) = \epsilon_{ijk} \partial_i \partial_j A_k = 0$$

which follows because $\partial_i \partial_j$ is symmetric w.r.t. $\{ij\}$ while ϵ_{ijk} is antisymmetric w.r.t. the same pair of indices.

4. $\vec{\nabla} \cdot (\vec{A} \times \vec{B}) = (\vec{\nabla} \times \vec{A}) \cdot \vec{B} - \vec{A} \cdot (\vec{\nabla} \times \vec{B})$

Expanding the l.h.s.,

$$\begin{aligned} \partial_i (\epsilon_{ijk} A_j B_k) &= \epsilon_{ijk} [(\partial_i A_j) B_k + A_j (\partial_i B_k)] \\ &= (\epsilon_{kij} \partial_i A_j) B_k - A_j (\epsilon_{jik} \partial_i B_k) \\ &= (\vec{\nabla} \times \vec{A}) \cdot \vec{B} - \vec{A} \cdot (\vec{\nabla} \times \vec{B}) \end{aligned}$$

5. $\vec{\nabla} \times \vec{r} = 0$

This also follows from the antisymmetry of the Levi-Civita tensor,

$$\vec{\nabla} \times \vec{r} = \epsilon_{ijk} \partial_j x_k = \epsilon_{ijk} \delta_{jk} = 0$$

6. $\vec{\nabla} \times \vec{\nabla} \phi \equiv 0$

This is another consequence of the same reasoning as above,

$$\vec{\nabla} \times \vec{\nabla} \phi = \epsilon_{ijk} \partial_j \partial_k \phi = 0$$

$$7. \vec{\nabla} \times (\phi \vec{A}) = (\vec{\nabla} \phi) \times \vec{A} + \phi (\vec{\nabla} \times \vec{A})$$

Consider the i^{th} component of the l.h.s.,

$$\begin{aligned} [\vec{\nabla} \times (\phi \vec{A})]_i &= \epsilon_{ijk} \partial_j (\phi A_k) = \epsilon_{ijk} (\partial_j \phi) A_k + \epsilon_{ijk} \phi (\partial_j A_k) \\ &= [\vec{\nabla} \phi \times \vec{A}]_i + \phi [\vec{\nabla} \times \vec{A}]_i \end{aligned}$$

As a special case, take $\vec{A} = \vec{r}$, then $\vec{\nabla} \times (\vec{r} \phi) = (\vec{\nabla} \phi) \times \vec{r}$.

$$8. \vec{\nabla} \times (\vec{\nabla} \times \vec{A}) = \vec{\nabla} (\vec{\nabla} \cdot \vec{A}) - \vec{\nabla}^2 \vec{A}$$

Beginning with,

$$\begin{aligned} [\vec{\nabla} \times (\vec{\nabla} \times \vec{A})]_i &= \epsilon_{ijk} \partial_j (\epsilon_{klm} \partial_l A_m) = (\delta_{il} \delta_{jm} - \delta_{im} \delta_{jl}) \partial_j \partial_l A_m \\ &= \partial_i (\partial_m A_m) - \partial_j \partial_j A_i = [\vec{\nabla} (\vec{\nabla} \cdot \vec{A})]_i - [\vec{\nabla}^2 \vec{A}]_i \end{aligned}$$

$$9. \vec{\nabla} \times (\vec{A} \times \vec{B}) = (\vec{\nabla} \cdot \vec{B}) \vec{A} - (\vec{\nabla} \cdot \vec{A}) \vec{B} + (\vec{B} \cdot \vec{\nabla}) \vec{A} - (\vec{A} \cdot \vec{\nabla}) \vec{B}$$

Again, beginning with,

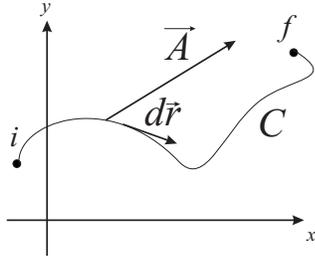
$$\begin{aligned} [\vec{\nabla} \times (\vec{A} \times \vec{B})]_i &= \epsilon_{ijk} \epsilon_{klm} \partial_j (A_l B_m) = (\delta_{il} \delta_{jm} - \delta_{im} \delta_{jl}) \partial_j (A_l B_m) \\ &= \partial_j (A_i B_j) - \partial_j (A_j B_i) \\ &= (\partial_j B_j) A_i + (B_j \partial_j) A_i - (\partial_j A_j) B_i - (A_j \partial_j) B_i \\ &= (\vec{\nabla} \cdot \vec{B}) [A]_i - (\vec{\nabla} \cdot \vec{A}) [B]_i + (\vec{B} \cdot \vec{\nabla}) [A]_i - (\vec{A} \cdot \vec{\nabla}) [B]_i \end{aligned}$$

$$10. \vec{\nabla} (\vec{A} \cdot \vec{B}) = (\vec{A} \cdot \vec{\nabla}) \vec{B} + (\vec{B} \cdot \vec{\nabla}) \vec{A} + \vec{A} \times (\vec{\nabla} \times \vec{B}) + \vec{B} \times (\vec{\nabla} \times \vec{A})$$

Consider the i^{th} component of the last two terms on the right,

$$\begin{aligned} [\vec{A} \times (\vec{\nabla} \times \vec{B}) + \vec{B} \times (\vec{\nabla} \times \vec{A})]_i &= \epsilon_{ijk} \epsilon_{klm} A_j \partial_l B_m + \epsilon_{ijk} \epsilon_{klm} B_j \partial_l A_m \\ &= (\delta_{il} \delta_{jm} - \delta_{im} \delta_{jl}) A_j \partial_l B_m \\ &= A_j \partial_i B_j - A_j \partial_j B_i + B_j \partial_i A_j - B_j \partial_j A_i \\ &= \partial_i (A_j B_j) - A_j \partial_j B_i - B_j \partial_j A_i \\ &= [\vec{\nabla} (\vec{A} \cdot \vec{B})]_i - (\vec{A} \cdot \vec{\nabla}) B_i - (\vec{B} \cdot \vec{\nabla}) A_i \end{aligned}$$

The stated result follows.



z

Figure 1.8: The Line Integral

11. A vector \vec{A} is said to be **irrotational** if $\vec{\nabla} \times \vec{A} = 0$ and it is **solenoidal** if $\vec{\nabla} \cdot \vec{A} = 0$. It turns out that $\vec{A} \times \vec{B}$ is solenoidal if both \vec{A} and \vec{B} are irrotational. Begin with

$$\begin{aligned} \vec{\nabla} \cdot (\vec{A} \times \vec{B}) &= \epsilon_{ijk} \partial_i (A_j B_k) = \epsilon_{ijk} (\partial_i A_j) B_k + \epsilon_{ijk} A_j (\partial_i B_k) \\ &= (\vec{\nabla} \times \vec{A}) \cdot \vec{B} - \vec{A} \cdot (\vec{\nabla} \times \vec{B}) = 0 \end{aligned}$$

(since both \vec{A} and \vec{B} are irrotational).

There are many more identities which we will encounter along the way and all of them can be proved using the methods above

1.11 Vector Integration

There are three types of integrations involving vector and scalar functions that lead to scalar quantities, *viz.*,

1.11.1 Line Integrals

Line integrals involve integrations along a curve, C , and generally depend upon the curve over which the integration is carried out. Three basic possibilities exist, *viz.*,

$$\int_{i,C}^f \phi(\vec{r}) d\vec{r}, \quad \int_{i,C}^f \vec{A} \times d\vec{r}, \quad \int_{i,C}^f \vec{A} \cdot d\vec{r}, \quad (1.11.1)$$

of which the first and second yield vectors and the last a scalar. C may be an open or closed curve, i and f are the beginning and endpoints of the integration on C and $d\vec{r}$ is an infinitesimal displacement along (tangent to) C . A famous example of a line integral is

the work performed by a force \vec{F} in moving a particle along some trajectory, C (see figure (1.8)). A particularly interesting case occurs when the vector \vec{A} is the gradient of a scalar function, *i.e.*, $\vec{A} = \vec{\nabla}\phi$. In this case,

$$\int_{i,C}^f \vec{A} \cdot d\vec{r} = \int_{i,C}^f \vec{\nabla}\phi \cdot d\vec{r} = \int_{i,C}^f d\phi = \phi(\vec{r}_f) - \phi(\vec{r}_i) \quad (1.11.2)$$

showing that the integral depends only on the endpoints and not on the curve C itself. A vector whose line integral is independent of the path along which the integration is carried out is called **conservative**. Every conservative vector then obeys

$$\oint_C \vec{A} \cdot d\vec{r} = 0, \quad (1.11.3)$$

for *every* closed path. Conversely, any vector that obeys (1.11.3) is expressible as the gradient of a scalar function, for .

$$\oint_C \vec{A} \cdot d\vec{r} = 0 \Rightarrow \vec{A} \cdot d\vec{r} = d\phi = \vec{\nabla}\phi \cdot d\vec{r} \quad (1.11.4)$$

and since $d\vec{r}$ is arbitrary, it follows that $\vec{A} = \vec{\nabla}\phi$.

One does not need to evaluate its line integral to determine whether or not a vector is conservative. From the fact that the curl of a gradient vanishes it follows that if \vec{A} is conservative then $\vec{\nabla} \times \vec{A} = 0$. The converse is also true, since if \vec{A} is irrotational then $\epsilon_{ijk}\partial_j A_k = 0$ for all i . These are simply integrability conditions for a function ϕ defined by $A_k = \partial_k \phi$. Therefore every irrotational vector is conservative and vice versa. The function $-\phi$ is generally called a **potential** of \vec{A} .

1.11.2 Surface integrals

Surface integrals also appear in the same three forms, the integration occurring over infinitesimal area elements, $d\vec{S}$, which are assigned the direction of the surface normal, (see figure (1.9)) Writing $d\vec{S}$ as $dS\hat{n}$, where \hat{n} is the unit normal to the surface at dS ,

$$\int_S dS(\hat{n}\phi), \quad \int_S dS(\hat{n} \times \vec{A}), \quad \int_S dS(\hat{n} \cdot \vec{A}) \quad (1.11.5)$$

where S is some arbitrary (open or closed) surface.

1.11.3 Volume Integrals

We may define volume integrals similarly but because the volume element is a scalar there are only two distinct possibilities,

$$\int_V d^3\vec{r}\phi, \quad \int_V d^3\vec{r}\vec{A}. \quad (1.11.6)$$

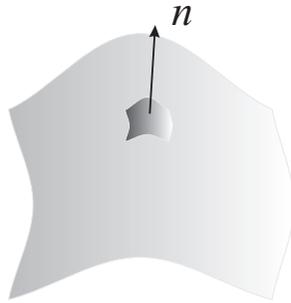


Figure 1.9: The Surface Integral

1.12 Integral Theorems

The three types of integrals that were defined in the previous section are connected by the following two theorems:¹⁴

1. **Stokes Theorem:**

$$\oint_C \vec{A} \cdot d\vec{r} = \int_S dS \hat{n} \cdot (\vec{\nabla} \times \vec{A}) \quad (1.12.1)$$

where C is a **closed** curve, S is the surface bounded by C , dS is an infinitesimal area element and \hat{n} is normal to the surface element dS .

2. **Gauss' theorem:**

$$\oint_S d\vec{S} \cdot \vec{A} = \int_V d^3\vec{r} \vec{\nabla} \cdot \vec{A} \quad (1.12.2)$$

where S is a **closed** surface and V is the volume bounded S .

While we accept these theorems without proof here, we shall now use them to prove some corollaries that will turn out to be useful in the future.

1.12.1 Corollaries of Stokes' Theorem

We will prove the following three relations:

1. $\oint_C \phi d\vec{r} = \int_S dS (\hat{n} \times \vec{\nabla} \phi)$
2. $\oint_C d\vec{r} \times \vec{A} = \int_S dS (\hat{n} \times \vec{\nabla}) \times \vec{A}$

¹⁴The proofs of these theorems can be found in any text on mathematical physics. We will leave it to the student to examine the proofs independently.

where C is a closed curve and S is the surface bounded by C in each case.

The proofs are quite simple. Define the vector $\vec{A} = \vec{a}\phi$, where \vec{a} is an arbitrary, constant vector then by Stokes' theorem,

$$\begin{aligned}\oint_C \vec{A} \cdot d\vec{r} &= \vec{a} \cdot \oint \phi d\vec{r} = \int_S dS \hat{n} \cdot (\vec{\nabla} \times \vec{a}\phi) = \int_S dS \hat{n} \cdot (\vec{\nabla}\phi \times \vec{a}) \\ &= \vec{a} \cdot \int_S dS (\hat{n} \times \vec{\nabla}\phi)\end{aligned}\quad (1.12.3)$$

Thus

$$\vec{a} \cdot \left(\oint \phi d\vec{r} - \int_S dS (\hat{n} \times \vec{\nabla}\phi) \right) = 0 \quad (1.12.4)$$

holds for arbitrary vectors implying the first identity.

The second identity may be derived similarly, by using $\vec{B} = \vec{a} \times \vec{A}$ in Stokes' theorem. Then

$$\begin{aligned}\oint_C \vec{B} \cdot d\vec{r} &= \oint_C \vec{a} \times \vec{A} \cdot d\vec{r} = - \oint_C (d\vec{r} \times \vec{A}) \cdot \vec{a} \\ &= \int_S dS \hat{n} \cdot (\vec{\nabla} \times \vec{B}) = \int_S dS [(\hat{n} \times \vec{\nabla}) \cdot \vec{B}] \\ &= \int_S dS [(\hat{n} \times \vec{\nabla}) \cdot (\vec{a} \times \vec{A})]\end{aligned}\quad (1.12.5)$$

But it is easy to show that¹⁵ $(\hat{n} \times \vec{\nabla}) \cdot (\vec{a} \times \vec{A}) = -[(\hat{n} \times \vec{\nabla}) \times \vec{A}] \cdot \vec{a}$, therefore

$$\vec{a} \cdot \left[\oint_C (d\vec{r} \times \vec{A}) - \int_S dS [\hat{n} \times (\vec{\nabla} \times \vec{A})] \right] = 0 \quad (1.12.6)$$

Again \vec{a} was arbitrary, therefore the second identity follows.

1.12.2 Corollaries of Gauss' theorem

We will prove three relations that will be quite helpful to us in the future, *viz.*,

1. $\oint_S dS (\vec{A} \times \hat{n}) = \int_V d^3\vec{r} \vec{\nabla} \times \vec{A}$
2. If ϕ and ψ are two arbitrary functions, then
 - (a) Green's **first identity**:

$$\int_V d^3\vec{r} [\vec{\nabla}\phi \cdot \vec{\nabla}\psi + \phi \vec{\nabla}^2\psi] = \int_S dS \hat{n} \cdot \phi \vec{\nabla}\psi \quad (1.12.7)$$

and

¹⁵Problem: Use the properties of the Levi-Civita tensor to show this

(b) Green's **second identity**:

$$\int_V d^3\vec{r} [\phi \vec{\nabla}^2 \psi - \psi \vec{\nabla}^2 \phi] = \int_S dS \hat{n} \cdot [\phi \vec{\nabla} \psi - \psi \vec{\nabla} \phi] \quad (1.12.8)$$

To prove the first corollary we will employ the trick we used to prove the corollaries of Stoke's theorem. For a constant vector \vec{a} , let $\vec{B} = \vec{a} \times \vec{A}$ and apply Gauss' law

$$\oint_S dS (\hat{n} \cdot \vec{B}) = \int_V d^3\vec{r} \vec{\nabla} \cdot \vec{B} \Rightarrow \oint_S dS [\hat{n} \cdot (\vec{a} \times \vec{A})] = \int_V d^3\vec{r} \vec{\nabla} \cdot (\vec{a} \times \vec{A}) \quad (1.12.9)$$

Developing the last relation, using some of the vector identities we proved earlier, we find

$$\vec{a} \cdot \left[\oint_S dS (\hat{n} \times \vec{A}) - \int_V d^3\vec{r} \vec{\nabla} \times \vec{A} \right] = 0 \quad (1.12.10)$$

But since \vec{a} is arbitrary, the identity follows.

To prove Green's two theorems is equally straightforward. Take $\vec{A} = \phi \vec{\nabla} \psi$ and apply Gauss' theorem. Since

$$\vec{\nabla} \cdot \vec{A} = \vec{\nabla} \cdot (\phi \vec{\nabla} \psi) = (\vec{\nabla} \phi) \cdot (\vec{\nabla} \psi) + \phi \vec{\nabla}^2 \psi, \quad (1.12.11)$$

it follows that the first of Green's theorems is just Gauss' theorem,

$$\int_V d^3\vec{r} [\vec{\nabla} \phi \cdot \vec{\nabla} \psi + \phi \vec{\nabla}^2 \psi] = \int_S dS \hat{n} \cdot \phi \vec{\nabla} \psi \quad (1.12.12)$$

To prove the second identity, consider $\vec{B} = \psi \vec{\nabla} \phi$ and again apply Gauss' theorem to get

$$\int_V d^3\vec{r} [\vec{\nabla} \psi \cdot \vec{\nabla} \phi + \psi \vec{\nabla}^2 \phi] = \int_S dS \hat{n} \cdot \psi \vec{\nabla} \phi \quad (1.12.13)$$

Subtracting the second from the first gives

$$\int_V d^3\vec{r} [\phi \vec{\nabla}^2 \psi - \psi \vec{\nabla}^2 \phi] = \int_S dS \hat{n} \cdot [\phi \vec{\nabla} \psi - \psi \vec{\nabla} \phi] \quad (1.12.14)$$

which is Green's second identity.

This chapter does not do justice to the vast area of vector analysis. On the contrary, most proofs have not been given and many useful identities have been neglected. What has been presented here is only an introduction to the material we will immediately need. Consequently, as we progress, expect to periodically encounter detours in which further vector analysis will be presented, often as exercises in the footnotes.

Chapter 2

Electrostatics: Introduction

2.1 Coulomb's Force Law

Electromagnetism begins with the electrostatic force law. Credit for the law is generally given to Charles Augustin de Coulomb, a french engineer who invented the torsion balance and used it to determine the distance dependence of the electric force between two charges. As its philosophical basis it had the mechanical approach of Newton. It sought to determine the force between two charge distributions so that one would then be able to predict the temporal evolution of the system of charges by a simple application of the second law

$$\frac{d\vec{p}}{dt} = \vec{F}^{\text{ext}} \quad (2.1.1)$$

After performing a series of remarkable experiments in which Coulomb measured the force between charged balls, he was able to arrive at the following conclusions

- There exist two kinds of charges in nature, which can conveniently be described as positive and negative (this conclusion should rightly be attributed to Benjamin Franklin).
- Like charges repel and unlike charges attract one another,
- The magnitude of the force exerted by one point-like charge on another point-like charge depends upon the product of the charges and the inverse square of the distance between them.
- The force between two point like charges acts along the line joining them.

These conclusions can be put together in a formula for the force exerted by one point like charge on another,

$$\vec{F}_{1 \rightarrow 2} = C \frac{q_1 q_2}{r_{12}^2} \hat{r}_{12} \quad (2.1.2)$$

where r_{12} is the distance between the two point-like charges q_1 and q_2 , \hat{r}_{12} is the unit vector that points from charge 1 to charge 2 and C is a universal constant called Coulomb's constant. In terms of the position vectors, \vec{r}_1 of charge 1 and \vec{r}_2 of charge 2, Coulomb's law can be stated as

$$\vec{F}_{1 \rightarrow 2} = C \frac{q_1 q_2 (\vec{r}_2 - \vec{r}_1)}{|\vec{r}_2 - \vec{r}_1|^3} \quad (2.1.3)$$

In the standard MKS system of units, the charge is measured in "Coulombs", roughly speaking one Coulomb is the charge contained by a collection of 0.5917×10^{19} protons. Distance is measured in meters and the constant C is often written for convenience as

$$C = \frac{1}{4\pi\epsilon_0}, \quad \epsilon_0 = 8.854 \times 10^{-12} \frac{\text{coulombs}^2}{\text{N} \cdot \text{m}^2} \quad (2.1.4)$$

where ϵ_0 is called the **permittivity** of space for reasons that will become clear only much later. In keeping with common usage, we will therefore henceforth use Coulomb's law in the form

$$\vec{F}_{1 \rightarrow 2} = \frac{1}{4\pi\epsilon_0} \frac{q_1 q_2 (\vec{r}_2 - \vec{r}_1)}{|\vec{r}_2 - \vec{r}_1|^3} \quad (2.1.5)$$

As usual, from the interaction between point particles or charges, we construct the interaction between arbitrary distributions.

Suppose Q is a point like charge and D is some distribution of point like charges. If D is a discrete distribution of N charges, q_j , then the force exerted by D on the point like charge Q is simply the vector sum of the forces exerted upon Q by the individual charges in the distribution,

$$\vec{F}_{D \rightarrow Q} = \frac{Q}{4\pi\epsilon_0} \sum_{j=1}^N \frac{q_j (\vec{r} - \vec{r}_j)}{|\vec{r} - \vec{r}_j|^3} \quad (2.1.6)$$

where \vec{r} is the position of charge Q and \vec{r}_j are the positions of the charges in the distribution. If the distribution can be considered continuous then we break it up into infinitesimal pieces, each of charge dq and add up the forces on Q due to these infinitesimal pieces as we did before. This time, however, the sum becomes an integral and we have

$$\vec{F}_{D \rightarrow Q} = \frac{Q}{4\pi\epsilon_0} \int_D dq(\vec{r}') \frac{(\vec{r} - \vec{r}')}{|\vec{r} - \vec{r}'|^3} \quad (2.1.7)$$

We can express the charge dq in terms of a volume charge density, $\rho(\vec{r}')$, of the distribution as $dq = \rho dV$ where dV is the infinitesimal volume element that contains the charge dq . Then

$$\vec{F}_{D \rightarrow Q} = \frac{Q}{4\pi\epsilon_0} \int_D d^3\vec{r}' \frac{\rho(\vec{r}') (\vec{r} - \vec{r}')}{|\vec{r} - \vec{r}'|^3} \quad (2.1.8)$$

Again, if D were a two dimensional distribution, that is distributed for example on the surface of some material then we are not interested in the volume charge density, $\rho(\vec{r}')$ but in the surface charge density $\sigma(\vec{r}')$ (defined so that $dq = \sigma dS$) and

$$\vec{F}_{D \rightarrow Q} = \frac{Q}{4\pi\epsilon_0} \int_D d^2\vec{r}' \frac{\sigma(\vec{r}')(\vec{r} - \vec{r}')}{|\vec{r} - \vec{r}'|^3} \quad (2.1.9)$$

For a lineal charge distribution, we replace the surface charge density by the lineal charge density in the integral (2.1.8) or (2.1.9) above. We have assumed a very important property of the electric force, upon which we will rely heavily throughout. We assumed that the electric force is simply additive. This implies that the interaction between a pair of charges does not affect the interaction between any of the charges of the pair and a third charge. Put in a more technical way, in electromagnetism we are able to “superpose” (just simply add) the forces exerted on any charge by a collection of charges. This is because electromagnetism is a “linear” theory. In fact, it is the only linear interaction of the four known fundamental interactions.

2.2 The Electric Field

The electric force between the charges does not require the charges to be in contact with one another. The presence of a single charge anywhere in space means that to every other point in space there must be associated a “force” that acts upon any other charge that may be located at that point. This leads us naturally to the concept of a **field** of a distribution D , which we define as

$$\vec{E} = \lim_{Q \rightarrow 0} \frac{\vec{F}_{D \rightarrow Q}}{Q} \quad (2.2.1)$$

The electric field at any point, due to a charge distribution D , is the electric force exerted by D per unit charge at that point. The limit $Q \rightarrow 0$ is taken to eliminate any effect that the “test” charge Q would have on the distribution. It follows that the electric field at \vec{r} of a point charge, q located at \vec{r}' , is

$$\vec{E}(\vec{r}) = \frac{1}{4\pi\epsilon_0} \frac{q(\vec{r} - \vec{r}')}{|\vec{r} - \vec{r}'|^3} \quad (2.2.2)$$

which can be directly generalized to a discrete distribution

$$\vec{E}(\vec{r}) = \frac{1}{4\pi\epsilon_0} \sum_{j=1}^N \frac{q_j(\vec{r} - \vec{r}_j)}{|\vec{r} - \vec{r}_j|^3} \quad (2.2.3)$$

and a continuous distribution,

$$\vec{E}(\vec{r}) = \frac{1}{4\pi\epsilon_0} \int_D d^3\vec{r}' \frac{\rho(\vec{r}')(\vec{r} - \vec{r}')}{|\vec{r} - \vec{r}'|^3} \quad (2.2.4)$$

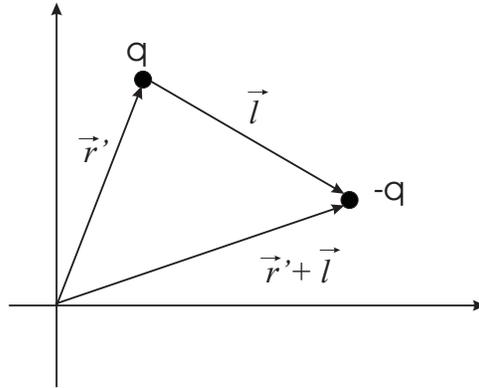


Figure 2.1: The electric dipole

or

$$\vec{E}(\vec{r}) = \frac{1}{4\pi\epsilon_0} \int_D d^2\vec{r}' \frac{\sigma(\vec{r}')(\vec{r} - \vec{r}')}{|\vec{r} - \vec{r}'|^3} \quad (2.2.5)$$

by superposition. Notice that none of these expressions depend on time if the distribution D is static and that any time dependence of \vec{E} must come about only via a time dependence of the source charge distribution. If the charge distributions under consideration are time independent then we have the “electrostatic” field.

The electric dipole serves a very simple, but important, example of the application of the above superpositions. In its purest form, the electric dipole consists of two opposite charges separated by a fixed distance. In this section we will find the electric field due to an electric dipole.

Consider the charges as located in the figure, so that the position vector of the positive charge is \vec{r}' and let \vec{l} be the position vector of the negative charge relative to the positive (see figure (2.1)). The electric field at any point $P(\vec{r})$ is then just the sum of the electric fields due to the individual charges

$$\vec{E}(\vec{r}) = \frac{q}{4\pi\epsilon_0} \left[\frac{\vec{r} - \vec{r}'}{|\vec{r} - \vec{r}'|^3} - \frac{\vec{r} - \vec{r}' - \vec{l}}{|\vec{r} - \vec{r}' - \vec{l}|^3} \right] \quad (2.2.6)$$

Expand the right hand side for small $|\vec{l}|/|\vec{r} - \vec{r}'|$, *i.e.*, for distances that are very much larger than the separation between the charges,

$$\vec{E}(\vec{r}) = \frac{q}{4\pi\epsilon_0} \left[\frac{\vec{r} - \vec{r}'}{|\vec{r} - \vec{r}'|^3} - \frac{\vec{r} - \vec{r}' - \vec{l}}{|\vec{r} - \vec{r}' - \vec{l}|^3} \left[1 + \frac{\vec{l}^2}{|\vec{r} - \vec{r}'|^2} - \frac{2\vec{l} \cdot (\vec{r} - \vec{r}')}{|\vec{r} - \vec{r}'|^2} \right]^{-3/2} \right]$$

$$\approx \frac{q}{4\pi\epsilon_0} \left[\frac{\vec{l}}{|\vec{r} - \vec{r}'|^3} - \frac{3\vec{l} \cdot (\vec{r} - \vec{r}')(\vec{r} - \vec{r}')}{|\vec{r} - \vec{r}'|^5} \right] \quad (2.2.7)$$

up to terms of $\mathcal{O}(l)$ (assuming that $|\vec{r} - \vec{r}'| \gg |\vec{l}|$, we can neglect the higher order terms in the expansion). The electric dipole moment is defined as

$$\vec{p} = q\vec{l} \quad (2.2.8)$$

and the approximation for the electric field in (2.2.7) should be re-written as

$$\vec{E}(\vec{r}) = \frac{1}{4\pi\epsilon_0} \left[\frac{\vec{p}}{|\vec{r} - \vec{r}'|^3} - \frac{3\vec{p} \cdot (\vec{r} - \vec{r}')(\vec{r} - \vec{r}')}{|\vec{r} - \vec{r}'|^5} \right] \quad (2.2.9)$$

where its dependence on the dipole moment becomes explicit. One can define the electric dipole moment of any distribution analogously: if the distribution is discrete,

$$\vec{p} = \sum_{j=1}^n q_j \vec{r}_j \quad (2.2.10)$$

suffices, and if the distribution is continuous,

$$\vec{p} = \int_V dq(\vec{r}') \vec{r}' = \int_V d^3\vec{r}' \rho(\vec{r}') \vec{r}' \quad (2.2.11)$$

where the integral is over the volume occupied by the charge distribution and $\rho(\vec{r}')$ is its density function.

2.3 Two Properties of the Electrostatic Field

Two properties of the electrostatic field are of great importance. The first is that the field is irrotational,

$$\vec{\nabla} \times \vec{E} = 0 \quad (2.3.1)$$

This is easy to prove, and it is sufficient to prove it for the electric field due to a single charge, say q , situated at the origin. Since the electric field due to any charge distribution is constructed by superposing the electric fields due to the individual parts, the result will follow for arbitrary charge distributions. This is just one example of the simplification allowed by linearity (superposition).

So let us consider

$$\vec{\nabla} \times \vec{E} = \frac{q}{4\pi\epsilon_0} \vec{\nabla} \times \frac{\vec{r} - \vec{r}'}{|\vec{r} - \vec{r}'|^3} = \frac{q}{4\pi\epsilon_0} \left[\frac{\vec{\nabla} \times (\vec{r} - \vec{r}')}{|\vec{r} - \vec{r}'|^3} + (\vec{r} - \vec{r}') \times \vec{\nabla} \frac{1}{|\vec{r} - \vec{r}'|^3} \right] \quad (2.3.2)$$

Now

$$[\vec{\nabla} \times (\vec{r} - \vec{r}')]_i = \epsilon_{ijk} \partial_j (x_k - x'_k) = \epsilon_{ijk} \delta_{jk} \equiv 0 \quad (2.3.3)$$

by the antisymmetry of the Levi-Civita symbol, and

$$\begin{aligned} \left[\vec{\nabla} \frac{1}{|\vec{r} - \vec{r}'|^3} \right]_i &= \partial_i \left[\sum_j (x_j - x'_j)^2 \right]^{-3/2} = -3 \left[\sum_j (x_j - x'_j)^2 \right]^{-5/2} (x_j - x'_j) \delta_{ij} \\ &= - \left[\frac{3(\vec{r} - \vec{r}')}{|\vec{r} - \vec{r}'|^5} \right]_i \end{aligned} \quad (2.3.4)$$

so

$$\vec{\nabla} \times \vec{E} = -\frac{3q}{4\pi\epsilon_0} (\vec{r} - \vec{r}') \times \frac{(\vec{r} - \vec{r}')}{|\vec{r} - \vec{r}'|^5} = 0 \quad (2.3.5)$$

Because \vec{E} is irrotational it can be expressed as the gradient of a scalar function and we can write

$$\vec{E} = -\vec{\nabla}\phi \quad (2.3.6)$$

(the minus sign is inserted for later convenience). The scalar function ϕ is called the **electrostatic potential**. Therefore the electrostatic field has really not three but one continuous degree of freedom and all of its three components are obtained from the gradient of a scalar function. Knowing \vec{E} we can easily determine ϕ . For a single charge, q , situated at \vec{r}' ,¹

$$\phi(\vec{r}) = \frac{q}{4\pi\epsilon_0 |\vec{r} - \vec{r}'|}. \quad (2.3.7)$$

Because we can superpose the electrostatic force and the electrostatic field, we can also superpose the electrostatic potential, ϕ . Therefore, for a discrete charge distribution, made up of charges q_j each located at \vec{r}'_j ,

$$\phi_D(\vec{r}) = \sum_{j \in D} \frac{q_j}{4\pi\epsilon_0 |\vec{r} - \vec{r}'_j|} \quad (2.3.8)$$

and for a continuous charge distribution,

$$\phi_D(\vec{r}) = \int_D d^3\vec{r}' \frac{\rho(\vec{r}')}{4\pi\epsilon_0 |\vec{r} - \vec{r}'|}. \quad (2.3.9)$$

Again,

$$\vec{F}_{D \rightarrow Q} = Q\vec{E}_D = -Q\vec{\nabla}\phi_D \quad (2.3.10)$$

¹Problem: Check this by taking its gradient and recovering the electrostatic field of a point charge

where we have explicitly used the suffix D to indicate that the field and potential are of the distribution D . The electrostatic force is conservative. The potential energy of the charge Q in the neighborhood of the distribution is

$$U_Q(\vec{r}) = - \int_*^{\vec{r}} \vec{F}_{D \rightarrow Q} \cdot d\vec{r} = Q \int_*^{\vec{r}} \vec{\nabla} \phi_D \cdot d\vec{r} = Q[\phi_D(\vec{r}) - \phi_D(*)] \quad (2.3.11)$$

where $*$ is used to represent the standard fixed point to which the potential energy is referred. As an example, the electrostatic potential due to a dipole is just ²

$$\begin{aligned} \phi(\vec{r}) &= \frac{q}{4\pi\epsilon_0} \left[\frac{1}{|\vec{r} - \vec{r}'|} - \frac{1}{|\vec{r} - \vec{r}' - \vec{l}'|} \right] \\ &\approx \frac{q}{4\pi\epsilon_0} \left[\frac{\vec{l}' \cdot (\vec{r} - \vec{r}')}{|\vec{r} - \vec{r}'|^3} \right] = \frac{\vec{p} \cdot (\vec{r} - \vec{r}')}{4\pi\epsilon_0 |\vec{r} - \vec{r}'|^3} \end{aligned} \quad (2.3.12)$$

The potential energy of the charge in the electric field of a dipole is,

$$U_Q(\vec{r}) = Q [\phi(\vec{r}) - \phi(*)] = \frac{Q}{4\pi\epsilon_0} \frac{\vec{p} \cdot (\vec{r} - \vec{r}')}{|\vec{r} - \vec{r}'|^3}, \quad (2.3.13)$$

taking the reference point at infinity, so that $\phi(*) = 0$.

Alternatively, if we place the dipole in an external electric field, $\vec{E}^{\text{ext}} = -\vec{\nabla}\Phi^{\text{ext}}$, then what would the potential energy of the dipole be? Evidently,

$$U_p(\vec{r}) = q\Phi^{\text{ext}}(\vec{r}) - q\Phi^{\text{ext}}(\vec{r} + \vec{l}) \quad (2.3.14)$$

Now

$$\Phi^{\text{ext}}(\vec{r} + \vec{l}) = \Phi^{\text{ext}}(\vec{r}) + \vec{l} \cdot \vec{\nabla}\Phi^{\text{ext}}(\vec{r}) \quad (2.3.15)$$

so

$$U_p(\vec{r}) = -\vec{p} \cdot \vec{E}^{\text{ext}} \quad (2.3.16)$$

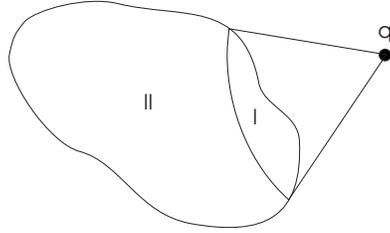
gives the potential energy in the external electric field up to $\mathcal{O}(\vec{l})$.

A consequence, following directly from Stokes theorem, of the fact that the electrostatic field is irrotational, is that

$$\mathcal{E} = \oint_C \vec{E} \cdot d\vec{r} = 0 \quad (2.3.17)$$

about any closed curve, C . For future reference, the quantity \mathcal{E} is called the **emf**. The vanishing of \mathcal{E} is equivalent to the statement that the electrostatic force is conservative, but this is true only for *electrostatic* fields.

²Problem: Show that $\vec{E} = -\vec{\nabla}\phi$ gives the electric field as calculated from first principles.

Figure 2.2: Gauss' Law: Charge not enclosed by S

The second property of the electrostatic field, known as Gauss' Law, follows from Gauss' theorem:

$$\vec{\nabla} \cdot \vec{E} = \frac{\rho}{\epsilon_0}. \quad (2.3.18)$$

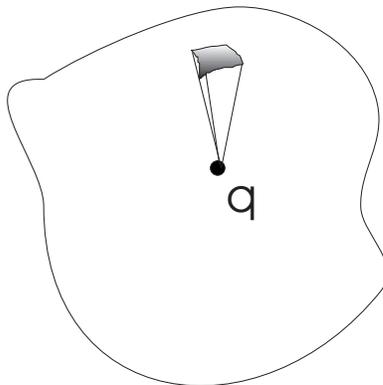
Again, if we can prove this for a point charge, then it will hold true for arbitrary charge distributions, simply because of superposition. Let us begin by evaluating

$$\oint_S dS \hat{n} \cdot \vec{E} = \frac{q}{4\pi\epsilon_0} \oint_S dS \frac{\hat{n} \cdot (\vec{r} - \vec{r}')}{|\vec{r} - \vec{r}'|^3} = \frac{q}{4\pi\epsilon_0} \oint_S dS \frac{\cos \theta}{|\vec{r} - \vec{r}'|^2} \quad (2.3.19)$$

where θ is the angle between the normal to the surface and the outward radial vector from the charge. Now $dS \hat{n} \cdot (\vec{r} - \vec{r}')/|\vec{r} - \vec{r}'|^3$ is just the projection of the surface area dS on the sphere of radius $|\vec{r} - \vec{r}'|$, therefore the integrand is simply the solid angle subtended by the projection of dS on the sphere of that radius at the charge q . If we sum up these infinitesimal solid angles, then two cases may arise: (a) the original surface S does not enclose the charge q (see figure(2.2)), or (b) the charge is enclosed by the surface S (see figure (2.3)). In case (a), the net result of summing up the solid angles subtended at q will be identically zero due to the two equal and opposite contributions from region I, where $\cos \theta$ is negative, and from region II, where $\cos \theta$ is positive. The contribution from region I is negative, whereas the contribution from region II is positive and equal in magnitude to the contribution from region I. In case (b), the infinitesimal solid angles add up to give precisely the total solid angle of 4π . Thus we find that

$$\oint_S dS \hat{n} \cdot \vec{E} = \frac{q_{\text{in}}}{\epsilon_0} \quad (2.3.20)$$

where the suffix in q_{in} is to indicate that only the charge *contained within* the surface S contributes. Equation (2.3.20) is the integral form of Gauss' law. Exploiting the fact that the electric field due to a distribution of charges is a simple superposition of the electric fields due to the individual charges, we may directly write down the integral form of Gauss'

Figure 2.3: Gauss' Law: Charge enclosed by S

law for a discrete distribution

$$\oint_S dS \hat{n} \cdot \vec{E} = \sum_{j=1}^N \frac{q_{j,\text{in}}}{\epsilon_0} \quad (2.3.21)$$

and a continuous volume distribution

$$\oint_S dS \hat{n} \cdot \vec{E} = \frac{1}{\epsilon_0} \int_V d^3\vec{r}' \rho(\vec{r}') \quad (2.3.22)$$

The differential form (2.3.18) of Gauss' law can now be obtained by exploiting Gauss' theorem,

$$\oint_S dS \hat{n} \cdot \vec{E} = \int_V d^3\vec{r}' \vec{\nabla} \cdot \vec{E} = \frac{q_{\text{in}}}{\epsilon_0} = \frac{1}{\epsilon_0} \int_V d^3\vec{r}' \rho(\vec{r}') \quad (2.3.23)$$

where we have replaced the charge contained by an integral over the volume of the density of charge. Eq. (2.3.18) then follows directly.³

2.4 Simple Applications of Gauss' Law

In its integral form, Gauss' law is useful to determine the electric fields of highly symmetric distributions. Otherwise, the using the integral form of Gauss' law to determine the electric

³**Note:** Outside of charges, Gauss' reads simply reads $\vec{\nabla} \cdot \vec{E} = 0$. Thus, this is also the equation for a single charge. Assuming spherical symmetry, prove that

$$\vec{E} = \frac{q\hat{r}}{4\pi\epsilon_0 r^2}$$

is a solution of $\vec{\nabla} \cdot \vec{E} = 0$. Where does the constant q come from?

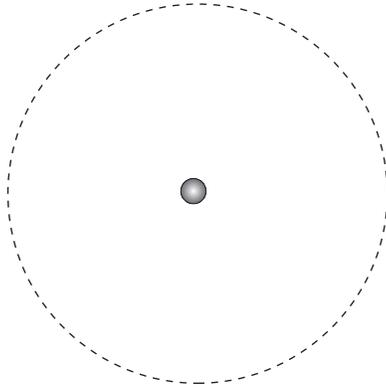


Figure 2.4: Gaussian surface for a point charge.

field is not recommended. The idea is that if the distribution is sufficiently symmetric and a closed surface that mimics the symmetry of the distribution is chosen, the integral becomes trivial to evaluate.

2.4.1 Point charge

The symmetry of a point charge is spherical, so, for a Gaussian surface, choose a sphere with the charge as its center (see figure (2.4)). By the symmetry, we expect that the electric field will point along the unit radial vector, so $\vec{E} = E\hat{r}$ and, moreover, that its magnitude will be constant *on the surface of the sphere*. The unit radial vector, \hat{r} , is also the normal to the sphere, so

$$\oint_S dS \hat{n} \cdot \vec{E} = E \oint_S dS = 4\pi r^2 E = \frac{q}{\epsilon_0} \Rightarrow \vec{E} = \frac{q\hat{r}}{4\pi\epsilon_0 r^2} \quad (2.4.1)$$

This argument may, of course, be extended to any spherically symmetric charge distribution. At points outside it, a spherical charge distribution will therefore behave as if it were a point charge situated at the center of the distribution.

2.4.2 Infinite line of charge of constant linear charge density:

The symmetry of an infinite line of charge is that of a right circular cylinder. For a Gaussian surface, choose an infinite cylinder with the line of charge along its axis (see figure (2.5)). Let ρ represent the radius of the cylinder and $\hat{\rho}$ the unit radial vector. By the symmetry, we expect that the electric field will point along the unit radial vector, so $\vec{E} = E\hat{\rho}$ and, moreover, that its magnitude will be constant *on the surface of the cylinder*,

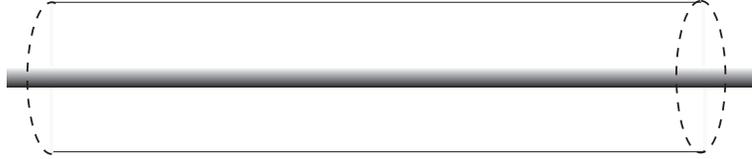


Figure 2.5: Gaussian surface for an infinite line of charge.

i.e., $E = E(\rho)$. We take the length of the line of charge to be L (with the understanding that the limit as $L \rightarrow \infty$ is to be taken in the end), so that if λ represents the linear charge density on the line then $q = \lambda L$. The unit radial vector $\hat{\rho}$ is also normal to the cylinder, so

$$\oint_S dS \hat{n} \cdot \vec{E} = E \oint_S dS = 2\pi\rho LE = \frac{q}{\epsilon_0} = \frac{\lambda L}{\epsilon_0} \Rightarrow \vec{E} = \frac{\lambda\hat{\rho}}{2\pi\epsilon_0\rho} \quad (2.4.2)$$

where q is the total charge on the line. What do we mean by an “infinite” line of charge? If the line of charge is finite, of length L , then its symmetry is not really cylindrical and in fact the problem becomes more complicated by the presence of edges. An infinite line of charge is the approximation in which these edge effects can be neglected, *i.e.*, when the point at which the field is measured is close to the line so that $\rho/L \ll 1$.

2.4.3 Infinite sheet of charge of constant areal charge density:

Choose a “pill-box”, *i.e.*, a cylinder closed at its two ends as shown in figure (2.6) for a Gaussian surface. By the planar symmetry of the charge distribution, we expect the electric field to be normal to the sheet at all points, therefore only the integrations over the upper and lower ends of the pill-box will yield non-vanishing contributions. Furthermore, as the magnitude of the electric field may depend at most on the perpendicular distance from the sheet and we can arrange the pill-box so that its two faces are equidistant from the sheet, the contributions from these two faces will be identical. If ΔS is the area of each face, integrating over the pill-box shaped Gaussian surface then gives

$$\oint_S dS \hat{n} \cdot \vec{E} = 2E\Delta S = \frac{q_{\text{in}}}{\epsilon_0} = \frac{\sigma\Delta S}{\epsilon_0} \Rightarrow \vec{E} = \frac{\sigma\hat{n}}{2\epsilon_0} \quad (2.4.3)$$

where \hat{n} is normal to the sheet.

2.4.4 Electric field of a conductor of arbitrary shape:

Let us now consider a conductor of arbitrary shape. A conductor is a material which has the property that charges may flow freely from any point within it to any other point.

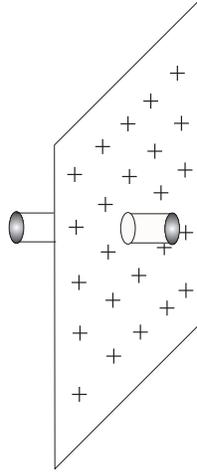


Figure 2.6: A “pill-box” Gaussian surface appropriate for an infinite sheet of charge.

Suppose we charge a conductor. Because the charges are free to move and because they will exert a repulsive force on one another, every charge will attempt to maximize the distance between itself and its nearest neighbor. After a characteristic time (typical of the conductor: for a good conductor this is approximately 10^{-16} s), they will therefore achieve an equilibrium state by arranging themselves on the surface of the conductor. A conductor in equilibrium has the following properties:

- If the conductor is isolated, its charge resides on its surface.
- The electric field inside the conductor vanishes everywhere.
- The electric field outside the conductor and “close” to its surface is perpendicular to it and depends only on the surface charge density at that point
- The surface charge density is not necessarily uniform. On an irregularly shaped conductor, it is largest at points on the surface of highest curvature.

Because the interior of the conductor will be free of charge it follows, from Gauss' law, that the electric flux across any closed surface *inside* the conductor will vanish. Since this is true for *any* Gaussian surface, it means that the electric field inside the conductor will also vanish.

To find the electric field in the exterior and close to the surface as before, choose a pill-box shaped Gaussian surface. As before, close enough to the surface we essentially have an “infinite” sheet of charge so the electric field is perpendicular to the surface at all

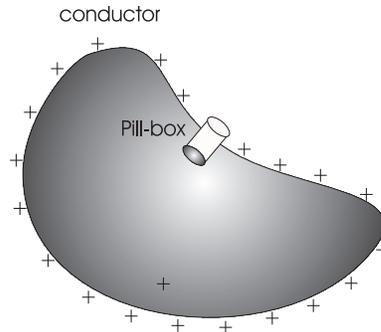


Figure 2.7: Solid Conductor and pill box.

points (see figure (2.7)). However, because the electric field vanishes inside the conductor, only the outer end of the pill-box contributes to the surface integral

$$\oint_S dS \hat{n} \cdot \vec{E} = E \Delta S = \frac{q_{\text{in}}}{\epsilon_0} = \frac{\sigma \Delta S}{\epsilon_0} \Rightarrow \vec{E} = \frac{\sigma \hat{n}}{\epsilon_0} \quad (2.4.4)$$

where \hat{n} is normal to the surface of the conductor.

2.5 Poisson's Equation and Laplace's Equation

If Gauss' Law is combined with the fact that \vec{E} is irrotational, we find that

$$\vec{\nabla} \cdot \vec{E} = -\vec{\nabla} \cdot \vec{\nabla} \phi = -\vec{\nabla}^2 \phi = \frac{\rho}{\epsilon_0} \quad (2.5.1)$$

or

$$\vec{\nabla}^2 \phi = -\frac{\rho}{\epsilon_0} \quad (2.5.2)$$

This equation is called Poisson's equation. It is the starting point for finding configurations of the electric field, *i.e.*, given any charge distribution in space, the problem of finding \vec{E} is reduced to the boundary value problem of finding the appropriate solution to (2.5.2). Most often we are interested in finding the electric field outside of the sources; in this case, the scalar potential satisfies Laplace's equation

$$\vec{\nabla}^2 \phi = 0 \quad (2.5.3)$$

If there are no non-trivial boundaries and the electric potential is required to fall off to zero at infinity, the solution can be no different from the solution we have already obtained,

$$\phi(\vec{r}) = \int_D d^3 \vec{r}' \frac{\rho(\vec{r}')}{4\pi\epsilon_0 |\vec{r} - \vec{r}'|}, \quad (2.5.4)$$

where the integration is performed over the charge distribution D . We will now prove that this is indeed the solution and, in doing so, encounter a new type of object called a **distribution**, in the form of the δ -**function**. Taking the Laplacian of the solution gives

$$\vec{\nabla}^2 \phi = \frac{1}{4\pi\epsilon_0} \int_D d^3\vec{r}' \rho(\vec{r}') \vec{\nabla}^2 \frac{1}{|\vec{r} - \vec{r}'|} \quad (2.5.5)$$

Now

$$\vec{\nabla}^2 \frac{1}{|\vec{r} - \vec{r}'|} = \partial_i \frac{x_i - x'_i}{|\vec{r} - \vec{r}'|^3} \equiv 0, \quad (2.5.6)$$

provided that $\vec{r} \neq \vec{r}'$. It is certainly not zero and, in fact, is badly defined at $\vec{r} = \vec{r}'$, which is precisely the limit that supports the integral above to give a non-zero right hand side. It is called a “ δ -function”, although it is not a function at all but a “distribution”. If we define, for future use,

$$\vec{\nabla}^2 \frac{1}{|\vec{r} - \vec{r}'|} = -4\pi\delta^3(\vec{r} - \vec{r}'), \quad (2.5.7)$$

then, because

$$\vec{\nabla}^2 \phi(\vec{r}) = -\frac{\rho(\vec{r})}{\epsilon_0}, \quad (2.5.8)$$

the distribution $\delta^3(\vec{r} - \vec{r}')$ must have the following property

$$\int d^3\vec{r}' f(\vec{r}') \delta^3(\vec{r} - \vec{r}') = f(\vec{r}) \quad (2.5.9)$$

for any function, $f(\vec{r}')$. In particular, taking $f(\vec{r}') = 1$, we should have

$$\int d^3\vec{r}' \delta(\vec{r} - \vec{r}') = 1 \quad (2.5.10)$$

As an exercise, let us convince ourselves that

$$-\frac{1}{4\pi} \vec{\nabla}^2 \left(\frac{1}{r} \right) = \delta^3(\vec{r}) \quad (2.5.11)$$

has this property. Applying Gauss' theorem,

$$-\frac{1}{4\pi} \int d^3\vec{r}' \vec{\nabla}^2 \left(\frac{1}{r} \right) = -\frac{1}{4\pi} \oint_S d\vec{S} \cdot \vec{\nabla} \left(\frac{1}{r} \right) \quad (2.5.12)$$

where S is a closed surface bounding the region of integration. Clearly, only the radial component of the surface normal will contribute to the integral. In other words, given any arbitrary bounding surface, only the projection of this surface on a sphere of radius r is

relevant. This means that, without loss of generality, we can take S to be a sphere and write the integral as

$$-\frac{1}{4\pi} \oint_S d\vec{S} \cdot \vec{\nabla} \left(\frac{1}{r} \right) = \frac{1}{4\pi} \oint_S d\theta d\phi r^2 \left(\frac{1}{r^2} \right) \equiv 1 \quad (2.5.13)$$

The δ -function provides a convenient way to represent the charge density of a discrete distribution. Notice how the distribution is essentially zero except at one point, its point of support, where it is infinite. In fact, by (2.5.9), it is “sufficiently infinite” for its integral to be finite and non-zero. Therefore, we may define the charge density of a discrete distribution as

$$\rho(\vec{r}) = \sum_{j=1}^N q_j \delta^3(\vec{r} - \vec{r}_j) \quad (2.5.14)$$

Inserting this into the solution for $\phi(\vec{r})$ in (2.3.9) gives

$$\phi(\vec{r}) = \int_D d^3\vec{r}' \frac{\rho(\vec{r}')}{4\pi\epsilon_0|\vec{r} - \vec{r}'|} = \sum_{j=1}^N \int_D d^3\vec{r}' \frac{q_j \delta^3(\vec{r}' - \vec{r}_j)}{4\pi\epsilon_0|\vec{r} - \vec{r}'|} = \sum_{j=1}^N \frac{q_j}{4\pi\epsilon_0|\vec{r} - \vec{r}_j|} \quad (2.5.15)$$

exactly as given in (2.3.8). The use of δ -functions to represent the charge density of a discrete distribution serves to unify our description of charge distributions, which will be convenient in what follows.

A word of caution: only solutions obeying the same boundary conditions can be superposed. For example, the solution (2.3.9) is a result of superposing the contributions to ϕ from various an infinite number of charges, $dq(\vec{r}')$. All these contributions were subject to the same trivial boundary conditions. When the boundary conditions are not trivial, the solutions will not be given simply by (2.3.9). We will show that then there is an additional contribution coming from the bounding surfaces on which the boundary conditions are given.

Chapter 3

Boundary Value Problems: Laplace's Equation

3.1 Preliminaries

We have seen that, the electrostatic field is completely specified by solutions to Poisson's equation for the electrostatic potential,

$$\vec{\nabla}^2 \phi = -\frac{\rho}{\epsilon_0} \quad (3.1.1)$$

from which the electric field is obtained as

$$\vec{E} = -\vec{\nabla} \phi \quad (3.1.2)$$

Not every solution of Poisson's equation is meaningful, but only those that satisfy certain boundary conditions, which are specific to any given problem, are of physical interest. We will undertake a study of methods to solve Poisson's equation in the following chapter. Here, we examine a simpler case of this equation that arises in the absence of charges, *i.e.*, in regions of space in which there are no charges present. In this case, $\rho(\vec{r}) = 0$ and the electrostatic potential is completely given by a solutions to *Laplace's* equation,

$$\vec{\nabla}^2 \phi = 0 \quad (3.1.3)$$

subject to given boundary conditions. The first question that arises therefore is what boundary conditions must be specified so that a unique solution is obtained.

3.2 Boundary Conditions

In physics a potential field of any sort is generally required to be well behaved. It should be finite everywhere except possibly at sources, and C^1 *i.e.*, continuous and differentiable

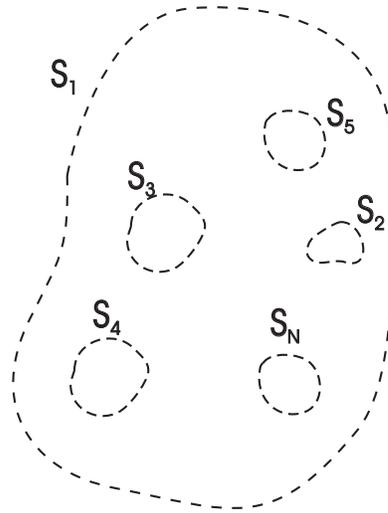


Figure 3.1: Electric field in a volume V bounded by the surface $S = S_1 + S_2 + S_3 + \dots$

everywhere. In the case of electrostatics this is equivalent to requiring that the electric field exists and is at least piecewise continuous everywhere. Imagine that we want to find the electrostatic field in a region V and let S be its bounding surface (which may be at infinity, or may be made up of many surfaces, $S = S_1 + S_2 + S_3 + \dots$, as shown in the figure (3.1). Before we begin to address this problem, we must keep in mind that one does not actually measure a potential. One measures the first derivatives of a potential (the electric field in the case of electrostatics), so it is impossible to distinguish between two potentials that differ only by a constant. It follows, as we now will see, that whether we are dealing with Laplace's equation or Poisson's equation, a sufficiently unique solution is obtained if *either* the value of the field *or* the value of the normal derivative of the field at S is supplied. To show this, we have to show that any two solutions of Poisson's equations in V , obtained by specifying *either* the value of the field *or* the value of the normal derivative at that surface, may differ by at most a constant.

We exploit Green's first identity. Let $\phi_{1,2}(\vec{r})$ be two solutions of Poisson's equation, valid inside a volume V and obtained either by specifying the value of the potential on its bounding surface S , or by specifying the value of the normal derivative of the potential on S . Let $\psi(\vec{r}) = \phi_1(\vec{r}) - \phi_2(\vec{r})$, then by superposition, $\psi(\vec{r})$ satisfies Laplace's equation

$$\vec{\nabla}^2 \psi = 0 \tag{3.2.1}$$

inside V . Moreover, since the two solutions $\phi_{1,2}$ were obtained by specifying the value of the potential on its bounding surface S , or by specifying the value of their normal

derivatives on S , it follows that either ψ or $\hat{n} \cdot \vec{\nabla}\psi = 0$ on S . By Green's first identity (1.12.7)

$$\int_V d^3\vec{r}[(\vec{\nabla}\psi)^2 + \psi\vec{\nabla}^2\psi] = \int_S dS \hat{n} \cdot \psi\vec{\nabla}\psi \quad (3.2.2)$$

But clearly the second term in brackets on the left vanishes together with the right hand side (no matter which boundary conditions were chosen!), so

$$\int_V d^3\vec{r}(\vec{\nabla}\psi)^2 = 0 \quad (3.2.3)$$

which is possible if and only if $\vec{\nabla}\psi = 0$ or $\psi = \text{const.}$ Therefore, in either case, the two functions ϕ_1 and ϕ_2 differ by an (irrelevant) constant.

- When the value of $\phi(\vec{r})$ on S is supplied, this is called a **Dirichlet** problem and the boundary condition is said to be of the “Dirichlet” type.
- When the normal derivative of $\phi(\vec{r})$ on S is supplied, this is a **Neumann** problem and the boundary condition is of the “Neumann” type.

A specification of *both* the field *and* its normal derivative on S (Cauchy problem, Cauchy boundary conditions) is an over specification of the problem and a solution agreeing with both conditions may not in fact exist. On the other hand, situations may arise in which Dirichlet conditions are appropriate to parts of the bounding surface and Neumann boundary conditions to other parts. Solutions will in general exist for this “mixed boundary condition” problem.

If the only problems encountered were of localized distributions of charge with no boundary surfaces except at infinity where the potential and its normal derivative vanish, then (2.3.9) would be a complete formal solution to the electrostatic problem and we would only need to perform one integration to obtain the electric field everywhere. Unfortunately, this is not the case and most interesting problems involve finite regions of space bounded by surfaces at which either Dirichlet or Neumann conditions hold. We can exploit the fact that if $\tilde{\phi}_\rho(\vec{r})$ is a solution to Poisson's equation, then

$$\phi(\vec{r}) = \tilde{\phi}_\rho(\vec{r}) + \Phi(\vec{r}) \quad (3.2.4)$$

will also solve Poisson's equation provided that $\Phi(\vec{r})$ satisfies Laplace's equation. In particular we could take

$$\tilde{\phi}_\rho(\vec{r}) = \int d^3\vec{r}' \frac{\rho(\vec{r}')}{4\pi\epsilon_0|\vec{r}-\vec{r}'|} \quad (3.2.5)$$

and add to it an appropriate solution of Laplace's equation so that $\phi(\vec{r})$ satisfies the boundary conditions. Are we assured that *all* boundary conditions can be accounted for

in this way? We can now use Green's second identity to show that the answer is "yes". In the second of Green's identities (1.12.8), take ϕ to be the electrostatic potential and $\psi = 1/|\vec{r} - \vec{r}'|$, then

$$\begin{aligned} \int_V d^3\vec{r}' [\phi(\vec{r}') \cdot \vec{\nabla}'^2(1/|\vec{r} - \vec{r}'|) - (\vec{\nabla}'^2\phi)(1/|\vec{r} - \vec{r}'|)] \\ = \int_S dS \hat{n} \cdot [\phi(\vec{r}') \vec{\nabla}'(1/|\vec{r} - \vec{r}'|) - (1/|\vec{r} - \vec{r}'|) \vec{\nabla}'\phi(\vec{r}')] \end{aligned} \quad (3.2.6)$$

Using the facts that $\vec{\nabla}'^2(1/|\vec{r} - \vec{r}'|) = -4\pi\delta(\vec{r} - \vec{r}')^1$ and that $\vec{\nabla}'^2\phi = -\rho(\vec{r}')/\epsilon_0$ we see that

$$\phi(\vec{r}) = \int d^3\vec{r}' \frac{\rho(\vec{r}')}{4\pi\epsilon_0|\vec{r} - \vec{r}'|} - \frac{1}{4\pi} \int_S dS [\phi(\vec{r}')(\hat{n} \cdot \vec{\nabla}'(1/|\vec{r} - \vec{r}'|)) - (1/|\vec{r} - \vec{r}'|)(\hat{n} \cdot \vec{\nabla}'\phi(\vec{r}'))] \quad (3.2.7)$$

If the only surface is the one at infinity, where the field and its normal derivatives vanish, only the first term survives and we have the result in (2.3.9). When non-trivial boundaries are present, (2.3.9) is modified by the second term on the right, which satisfies Laplace's equation and is therefore to be identified with $\Phi(\vec{r})$. The modification may be thought of as arising from charges induced or residing on the surfaces themselves and, in fact, the boundary conditions themselves may be thought of as specifying an effective charge distribution on the boundary surfaces.

Thus situations may arise in which the electric potential (and the electric field) is non-vanishing in a charge free volume. In this case, the electric potential in that region depends only on the field and its normal derivative on the boundary. This is surprising because, as we have just seen, the specification of both the field and its normal derivative is an over specification of the problem. Thus one cannot, for example, obtain any desired field in a cavity, by simply adjusting ϕ and $\hat{n} \cdot \vec{\nabla}'\phi$ on the boundary.

The solution (3.2.7) is only formal in the sense that both terms in the surface integral on the right cannot generally be given. It should be thought of only as an integral relation for $\phi(\vec{r})$. One can write this relation in an more general fashion as

$$\phi(\vec{r}) = \frac{1}{4\pi\epsilon_0} \int d^3\vec{r}' \mathcal{G}(\vec{r}, \vec{r}') \rho(\vec{r}') - \frac{1}{4\pi} \int_S dS [\phi(\vec{r}')(\hat{n} \cdot \vec{\nabla}'\mathcal{G}(\vec{r}, \vec{r}')) - \mathcal{G}(\vec{r}, \vec{r}')(\hat{n} \cdot \vec{\nabla}'\phi(\vec{r}'))] \quad (3.2.8)$$

provided that

$$\vec{\nabla}'^2\mathcal{G}(\vec{r}, \vec{r}') = -4\pi\delta^3(\vec{r} - \vec{r}') \quad (3.2.9)$$

Any function $\mathcal{G}(\vec{r}, \vec{r}')$ satisfying (3.2.9) is called a Green function. The function $1/|\vec{r} - \vec{r}'|$ is only one of a family of Green functions. Quite generally, we could take

$$\mathcal{G}(\vec{r}, \vec{r}') = \frac{1}{|\vec{r} - \vec{r}'|} + \mathcal{Q}(\vec{r}, \vec{r}') \quad (3.2.10)$$

¹Clearly, $\vec{\nabla}'^2(1/|\vec{r} - \vec{r}'|) = \vec{\nabla}^2(1/|\vec{r} - \vec{r}'|)$. Convince yourself that this is true.

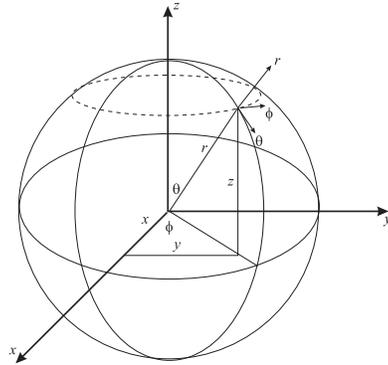


Figure 3.2: Spherical Coordinates

where $Q(\vec{r}, \vec{r}')$ satisfies Laplace's equation, *i.e.*,

$$\vec{\nabla}^2 Q(\vec{r}, \vec{r}') = 0 \quad (3.2.11)$$

With this generalized Green function and its additional freedom, we could choose $Q(\vec{r}, \vec{r}')$ in such a way as to eliminate one or other of the surface integrals, obtaining a result that involves either Dirichlet or Neumann boundary conditions. We turn to this problem in the next chapter. Here we will concentrate on general solutions of Laplace's equation with symmetries.

3.3 Symmetries

When symmetries are present it is always convenient to select a coordinate system that is adapted to them. We will be interested principally in problems with rectangular, spherical or axial symmetry. In these cases it is convenient to turn respectively to Cartesian coordinates, Spherical coordinates or Cylindrical coordinates respectively. Later we will learn to work with general general coordinate systems, but then we will introduce more powerful techniques.

3.3.1 Spherical Coordinates

Thus, for example, if we know that the source charge distribution and boundary conditions are spherically symmetric it is convenient to apply the spherical coordinates, defined in terms of the cartesian coordinate system by the transformations

$$r = \sqrt{x^2 + y^2 + z^2}$$

$$\begin{aligned}\varphi &= \tan^{-1}\left(\frac{y}{x}\right) \\ \theta &= \cos^{-1}\left(\frac{z}{\sqrt{x^2 + y^2 + z^2}}\right)\end{aligned}\tag{3.3.1}$$

and the inverse transformations

$$\begin{aligned}x &= r \sin \theta \cos \varphi \\ y &= r \sin \theta \sin \varphi \\ z &= r \cos \theta\end{aligned}\tag{3.3.2}$$

These transformations are of course valid only away from the origin. Recall that unit vectors representing the directions of increasing r , θ and φ are given by

$$\begin{aligned}\hat{r} &= (\sin \theta \cos \varphi, \sin \theta \sin \varphi, \cos \theta) \\ \hat{\theta} &= (\cos \theta \cos \varphi, \cos \theta \sin \varphi, -\sin \theta) \\ \hat{\varphi} &= (-\sin \varphi, \cos \varphi, 0)\end{aligned}\tag{3.3.3}$$

and that

$$\begin{aligned}\frac{\partial \hat{r}}{\partial \theta} &= \hat{\theta}, & \frac{\partial \hat{r}}{\partial \varphi} &= \hat{\varphi} \sin \theta \\ \frac{\partial \hat{\theta}}{\partial \theta} &= -\hat{r}, & \frac{\partial \hat{\theta}}{\partial \varphi} &= \hat{\varphi} \cos \theta \\ \frac{\partial \hat{\varphi}}{\partial \theta} &= 0, & \frac{\partial \hat{\varphi}}{\partial \varphi} &= -\hat{r} \sin \theta - \hat{\theta} \cos \theta\end{aligned}\tag{3.3.4}$$

The unit vectors in the original Cartesian system may be given in terms of these unit vectors as follows

$$\begin{aligned}\hat{x} &= (\hat{x} \cdot \hat{r})\hat{r} + (\hat{x} \cdot \hat{\theta})\hat{\theta} + (\hat{x} \cdot \hat{\varphi})\hat{\varphi} = \hat{r} \sin \theta \cos \varphi + \hat{\theta} \cos \theta \cos \varphi - \hat{\varphi} \sin \varphi \\ \hat{y} &= (\hat{y} \cdot \hat{r})\hat{r} + (\hat{y} \cdot \hat{\theta})\hat{\theta} + (\hat{y} \cdot \hat{\varphi})\hat{\varphi} = \hat{r} \sin \theta \sin \varphi + \hat{\theta} \cos \theta \sin \varphi + \hat{\varphi} \cos \varphi \\ \hat{z} &= (\hat{z} \cdot \hat{r})\hat{r} + (\hat{z} \cdot \hat{\theta})\hat{\theta} + (\hat{z} \cdot \hat{\varphi})\hat{\varphi} = \hat{r} \cos \theta - \hat{\theta} \sin \theta\end{aligned}\tag{3.3.5}$$

Of interest to us is the Laplacian in these coordinates, so let us begin by evaluating the gradient operator. Transforming its definition in Cartesian coordinates,

$$\vec{\nabla} = \hat{x} \frac{\partial}{\partial x} + \hat{y} \frac{\partial}{\partial y} + \hat{z} \frac{\partial}{\partial z}\tag{3.3.6}$$

to spherical coordinates by using the transformations just given, we find²

$$\vec{\nabla} = \hat{r} \frac{\partial}{\partial r} + \frac{\hat{\theta}}{r} \frac{\partial}{\partial \theta} + \frac{\hat{\varphi}}{r \sin \theta} \frac{\partial}{\partial \varphi}\tag{3.3.7}$$

²Problem: Show this.

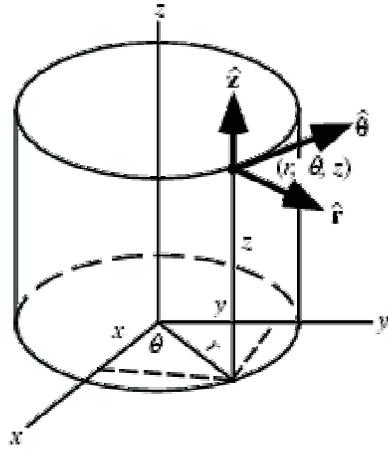


Figure 3.3: Cylindrical Coordinates

and taking the inner product³ we obtain the Laplacian in spherical coordinates

$$\vec{\nabla}^2 = \frac{1}{r^2} \partial_r (r^2 \partial_r) + \frac{1}{r^2 \sin \theta} \partial_\theta (\sin \theta \partial_\theta) + \frac{1}{r^2 \sin^2 \theta} \partial_\varphi^2 \quad (3.3.8)$$

The Euclidean distance between two points may be given in terms of r , θ and φ by transforming its expression in Cartesian coordinates,

$$d\vec{r}^2 = dx^2 + dy^2 + dz^2 = dr^2 + r^2(d\theta^2 + \sin^2 \theta d\varphi^2) \quad (3.3.9)$$

The volume element must account for the Jacobian of the transformation from the Cartesian system to the spherical system.

$$\int d^3\vec{r} = \int dx dy dz = \int dr d\theta d\varphi \det(\mathbb{J}) = \int (r^2 \sin \theta) dr d\theta d\varphi \quad (3.3.10)$$

gives the volume of spheres.

3.3.2 Cylindrical coordinates

Cylindrical coordinates are defined by the following transformations from a Cartesian system:

$$\rho = \sqrt{x^2 + y^2}$$

³Problem: Show this. Note that the gradient is a vector *operator*, so be careful to take appropriate derivatives when necessary while taking the inner product.

$$\begin{aligned}\varphi &= \tan^{-1} \left(\frac{y}{x} \right) \\ z &= z\end{aligned}\tag{3.3.11}$$

where, we have assumed that the axial symmetry is about the “ z ” axis. The inverse transformations are simple to obtain

$$\begin{aligned}x &= \rho \cos \varphi \\ y &= \rho \sin \varphi \\ z &= z\end{aligned}\tag{3.3.12}$$

and this time the system is badly defined along the entire z -axis. Nevertheless, for points away from the z -axis, we may define the unit vectors

$$\begin{aligned}\hat{\rho} &= (\cos \varphi, \sin \varphi, 0) \\ \hat{\varphi} &= (-\sin \varphi, \cos \varphi, 0) \\ \hat{z} &= (0, 0, 1)\end{aligned}\tag{3.3.13}$$

(ρ and φ are now just polar coordinates in the $x - y$ plane) which satisfy

$$\frac{\partial \hat{\rho}}{\partial \varphi} = \hat{\varphi}, \quad \frac{\partial \hat{\varphi}}{\partial \varphi} = -\hat{\rho}\tag{3.3.14}$$

(all other derivatives vanish). The Cartesian unit vectors can be expressed in terms of the cylindrical ones as we did in the case of spherical symmetry,

$$\begin{aligned}\hat{x} &= \hat{\rho} \cos \varphi - \hat{\varphi} \sin \varphi \\ \hat{y} &= \hat{\rho} \sin \varphi + \hat{\varphi} \cos \varphi \\ \hat{z} &= \hat{z}\end{aligned}\tag{3.3.15}$$

and the gradient operator can be transformed to the cylindrical system as before to get

$$\vec{\nabla} = \hat{\rho} \partial_{\rho} + \frac{\hat{\varphi}}{\rho} \partial_{\varphi} + \hat{z} \partial_z\tag{3.3.16}$$

giving

$$\vec{\nabla}^2 = \frac{1}{\rho} \partial_{\rho} (\rho \partial_{\rho}) + \frac{1}{\rho^2} \partial_{\varphi}^2 + \partial_z^2\tag{3.3.17}$$

The Euclidean distance between two points in these coordinates is

$$d\vec{r}^2 = d\rho^2 + \rho^2 d\varphi^2 + dz^2\tag{3.3.18}$$

Again, the volume element must account for the Jacobian of the transformation from the Cartesian system to the spherical system,

$$\int d^3\vec{r} = \int dx dy dz = \int \rho d\rho d\varphi dz \det(\mathbb{J}) = \int \rho d\rho d\varphi dz\tag{3.3.19}$$

gives the volume of cylinders.

3.4 General Solutions of Laplace's Equation with symmetries

Before going on to examine the general solutions (no boundary conditions as yet imposed) when high amounts of symmetry (spherical and cylindrical, for example) are involved, let us remember that, because Laplace's equation is linear, superposition of solutions holds. Therefore if we have N solutions of Laplace's equation, the linear combination,

$$\phi(\vec{r}) = \sum_{j=1}^N c_j \phi_j(\vec{r}), \quad (3.4.1)$$

also solves Laplace's equation.

3.4.1 One dimensional solutions

If the symmetry is rectangular and if ϕ is independent of two of the three dimensions, the problem is effectively one dimensional and Cartesian coordinates are appropriate. Laplace's equation in Cartesian coordinates is just

$$\frac{d^2\phi}{dx^2} = 0 \Rightarrow \phi(x) = ax + b \quad (3.4.2)$$

where a and b are arbitrary constants (b is physically irrelevant) that must be determined from the boundary conditions. The electric field in this case is constant and points in the x -direction. It may be thought of as arising from an infinite sheet of charge, whose surface charge density is related to a . If the symmetry is spherical and there is no angular dependence, Laplace's equation in spherical coordinates is to be used and one has

$$\frac{1}{r^2} \partial_r (r^2 \partial_r \phi) = 0 \Rightarrow \phi(r) = \frac{a}{r} + b \quad (3.4.3)$$

Again, the constant b is physically irrelevant and the electric field is that of a point charge (a related to this charge) located at the origin. With cylindrical symmetry we solve Laplace's equation in cylindrical coordinates

$$\frac{1}{\rho} \partial_\rho (\rho \partial_\rho \phi) = 0 \Rightarrow \phi(\rho) = a \ln |\rho| + b \quad (3.4.4)$$

The electric field due to this potential is evidently due to an infinite line of charge, located on the z -axis and a is related to the line density of charge. All of these solutions have been obtained before using Gauss' law. So, indeed the high degree of symmetry we required to effectively use the integral form of Gauss' law turned out to yield only the most trivial of solutions.

3.4.2 Two dimensional solutions

If the electrostatic potential does not depend on one dimension, but only on two, we write Laplace's in Cartesian coordinates as

$$(\partial_x^2 + \partial_y^2)\phi = 0 \quad (3.4.5)$$

and transform the complex coordinates, $z = x + iy$, $z^* = x - iy$. Then we have

$$\begin{aligned} \partial_x &= \frac{\partial z}{\partial x} \partial_z + \frac{\partial z^*}{\partial x} \partial_{z^*} = \partial_z + \partial_{z^*} \\ \partial_y &= i(\partial_z - \partial_{z^*}) \end{aligned} \quad (3.4.6)$$

so that $\vec{\nabla}^2 \phi = 0$ is equivalent to

$$\partial_z \partial_{z^*} \phi = 0 \quad (3.4.7)$$

whose general solution,

$$\phi(z, z^*) = F(z) + G(z^*), \quad (3.4.8)$$

must then be subjected to suitable boundary conditions. Clearly, even though this form of the solution was obtained using Cartesian coordinates, it can equally well be written in polar coordinates, using $z = re^{i\varphi}$. This is a “conformal field theory”. Usually there are so many solutions that it is not an easy matter to disentangle the meaningful ones among them, even given the boundary conditions. Techniques based on conformal mapping are often used.

3.4.3 Three dimensional solutions

If it is possible to choose coordinates adapted to the symmetries of the problem, one may look for solutions that are separable in these coordinates. Separability means that, if w_1 , w_2 and w_3 are the adapted coordinates, the solution chosen is such that

$$\phi(\vec{r}) = \phi(w_1, w_2, w_3) = \prod_i \phi^i(w_i) = \phi^1(w_1)\phi^2(w_2)\phi^3(w_3) \quad (3.4.9)$$

Laplace's equation will in general break up into three ordinary differential equations, one for each $\phi^i(w_i)$.

The solutions of these ordinary differential equations will be complete sets of “orthogonal” functions within the bounding surface. Consider therefore what a complete set of orthogonal functions means in one dimension. Let n label the solutions of the ordinary differential equation obtained and let the solutions be denoted by $\phi_n(w)$, valid in some interval, say (a, b) . Orthogonality means that

$$\int_a^b dw \phi_n^*(w)\phi_m(w) = C_n \delta_{n,m} \quad \forall n, m \quad (3.4.10)$$

3.4. GENERAL SOLUTIONS OF LAPLACE'S EQUATION WITH SYMMETRIES 55

where C_n is some constant that depends on the length of the interval and possibly on n . The relation is analogous to the usual scalar product between the basis vectors in a finite dimensional space and should be thought of as such. The functions $\phi_n(w)$ can be thought of as a set of basis “vectors” in the space of square integrable functions on the interval (a, b) . The integral is, in fact, often referred to as a scalar product in this space. Completeness of the functions means that *any* square integrable solution of Laplace's equation, $f(w)$, can be represented as some linear combination of the set $\phi_n(w)$ in that interval, *i.e.*,

$$f(w) = \sum_n a_n \phi_n(w) \quad (3.4.11)$$

One can always “normalize” the functions $\phi_n(w)$, that is to say multiply them by a suitable constant (possibly dependent on n) so that (3.4.10) becomes

$$\int_a^b dw \phi_n^*(w) \phi_m(w) = \delta_{n,m} \quad \forall n, m \quad (3.4.12)$$

The functions $\phi_n(w)$ so normalized are then said to be “orthonormal”. Let us work with orthonormal functions from now on. Multiplying (3.4.11) by $\phi_m^*(w)$ and integrating over the interval (a, b) , we find

$$\int_a^b dw \phi_m^*(w) f(w) = \sum_n a_n \int_a^b dw \phi_m^*(w) \phi_n(w) = \sum_n a_n \delta_{m,n} = a_m \quad (3.4.13)$$

where we have used the orthonormality of the ϕ_n in the second to last step. This can be thought of as using the scalar product to obtain the projection of the function $f(w)$ along (or “in the direction of”) the basis “vector” ϕ_n . The coefficient a_n is then seen to be just this projection. We can go further and insert this expression for a_n into the expansion (3.4.11) to find

$$f(w) = \int_a^b dw' \left[\sum_n \phi_n^*(w') \phi_n(w) \right] f(w') \Rightarrow \sum_n \phi_n^*(w') \phi_n(w) = \delta(w - w') \quad (3.4.14)$$

The last is called the “completeness” relation. The completeness relation must be satisfied if $f(w)$ is to be faithfully represented by the expansion in (3.4.11). It is similar to the orthonormality relation, except that the roles of the set label, n , and the coordinate w are interchanged. If the completeness relation is not satisfied, the set $\{\phi_n(w)\}$ is unable to faithfully reproduce any square solution of Laplace's equation in the interval of definition. The set is then said to be incomplete. Let us work out some simple examples to illustrate the method.

3.5 Examples in three dimensions

3.5.1 Problems with Rectangular Symmetry

If the symmetry is rectangular, Cartesian coordinates are most convenient. In Cartesian coordinates, assume a solution of the form $\phi(x, y, z) = X(x)Y(y)Z(z)$, where X , Y and Z are functions only of x , y and z respectively. Then

$$\vec{\nabla}^2 \phi = YZ\partial_x^2 X + ZX\partial_y^2 Y + XY\partial_z^2 Z = 0 \quad (3.5.1)$$

Dividing by XYZ gives

$$\frac{1}{Y}\partial_y^2 Y + \frac{1}{Z}\partial_z^2 Z = -\frac{1}{X}\partial_x^2 X \quad (3.5.2)$$

Now the left hand side depends only on y, z and the right hand side depends only on x so they can be equal only if each is constant. Call this constant λ , then

$$\partial_x^2 X + \lambda X = 0 \Rightarrow X(x) = A_{\pm} e^{\pm i\sqrt{\lambda}x} \quad (3.5.3)$$

where A_{\pm} are constants of the integration. Using the same argument for the equation

$$\frac{1}{Y}\partial_y^2 Y + \frac{1}{Z}\partial_z^2 Z = \lambda \Rightarrow \frac{1}{Y}\partial_y^2 Y = \lambda - \frac{1}{Z}\partial_z^2 Z = -\sigma \quad (3.5.4)$$

Then

$$Y(y) = B_{\pm} e^{\pm i\sqrt{\sigma}y} \quad (3.5.5)$$

and

$$Z(z) = C_{\pm} e^{\pm\sqrt{\lambda+\sigma}z} \quad (3.5.6)$$

are also linear combinations of exponentials. The constants $A_{\pm}, B_{\pm}, C_{\pm}$ and λ, σ are all to be determined from the boundary conditions. We may write the solutions as linear combinations of

$$e^{\pm i\alpha x} e^{\pm i\beta y} e^{\pm\sqrt{\alpha^2+\beta^2}z} \quad (3.5.7)$$

where $\alpha = \sqrt{\lambda}$ and $\beta = \sqrt{\sigma}$.

In a given problem, it is better to think of the separate directions independently. Thus consider the problem of finding the electrostatic potential inside an empty, rectangular box made of conducting sides of lengths (a, b, c) , situated so that one vertex is at the origin and the opposite vertex is at the point (a, b, c) . Suppose that we require all faces of the box to be at constant potential, say ϕ_0 , except the face $z = c$, which is at some given potential $\tilde{\phi}(x, y)$. We first find the solution that gives $\phi(\vec{r}) = 0$ at all faces except the face $z = c$. For the solution of the form described above to vanish at $x = 0$ and $y = 0$ we should have

$$X(x) = X_0 \sin(\alpha x), \quad Y(y) = Y_0 \sin(\beta y) \quad (3.5.8)$$

and for these to vanish at $x = a$ and $y = b$, the coefficients should take values $\alpha = n\pi/a$, $\beta = m\pi/b$, where m, n are integers. The two sets of functions

$$\left\{ \sin\left(\frac{n\pi x}{a}\right) \right\}, \quad \left\{ \sin\left(\frac{m\pi y}{b}\right) \right\}, \quad n, m \in \mathbb{N} \quad (3.5.9)$$

are each complete (as we mentioned they would be) in the intervals $(0, a)$ and $(0, b)$ respectively and the functions themselves are all orthogonal.⁴ The solution for $Z(z)$ is of the form

$$Z(z) = Z_{\pm} e^{\pm \sqrt{\frac{n^2}{a^2} + \frac{m^2}{b^2}} z} \quad (3.5.10)$$

which can vanish at $z = 0$ only if we take the special combination

$$Z(z) = Z_0 \sinh\left(\sqrt{\frac{n^2}{a^2} + \frac{m^2}{b^2}} z\right) \quad (3.5.11)$$

The general solution satisfying $\phi = \phi_0$ at all faces except $z = c$ can be written as

$$\phi(\vec{r}) = \phi_0 + \sum_{n,m} a_{nm} \sin\left(\frac{n\pi x}{a}\right) \sin\left(\frac{m\pi y}{b}\right) \sinh\left(\sqrt{\frac{n^2}{a^2} + \frac{m^2}{b^2}} z\right) \quad (3.5.12)$$

where the a_{mn} are coefficients determined by the value $\tilde{\phi}(x, y)$, of $\phi(\vec{r})$ in the plane $z = c$. $\tilde{\phi}$ is given by

$$\tilde{\phi}(x, y) = \phi_0 + \sum_{n,m} A_{nm} \sin\left(\frac{n\pi x}{a}\right) \sin\left(\frac{m\pi y}{b}\right) \quad (3.5.13)$$

and

$$A_{nm} = a_{nm} \sinh\left(\sqrt{\frac{n^2}{a^2} + \frac{m^2}{b^2}} c\right) \quad (3.5.14)$$

We recognize the coefficients A_{mn} as the Fourier coefficients of the function $\tilde{\phi}(x, y) - \phi_0$. Thus⁵

$$A_{mn} = \frac{4}{ab} \int_0^a dx \int_0^b dy (\tilde{\phi}(x, y) - \phi_0) \sin\left(\frac{n\pi x}{a}\right) \sin\left(\frac{m\pi y}{b}\right) \quad (3.5.15)$$

Given any $\tilde{\phi}(x, y)$ we can find the coefficients a_{mn} in the solution. As a consequence, if the face $z = c$ is maintained at the same potential as the others, ϕ_0 , then all the coefficients vanish and the only solution inside the box is the trivial one, $\phi(\vec{r}) = \phi_0$.

⁴Problem: Prove this, *i.e.*, show that

$$\int_0^a dx \sin\left(\frac{n\pi x}{a}\right) \sin\left(\frac{m\pi x}{a}\right) = \frac{a}{2} \delta_{mn}$$

⁵Question: Why has the factor $4/ab$ been introduced?

3.5.2 Problems with Cylindrical Symmetry

In cylindrical coordinates Laplace's equation has the form

$$\vec{\nabla}^2 \phi = \frac{1}{\rho} \partial_\rho(\rho \partial_\rho \phi) + \frac{1}{\rho^2} \partial_\varphi^2 \phi + \partial_z^2 \phi = 0. \quad (3.5.16)$$

If the boundary conditions allow for the separability of the electrostatic potential, say $\phi(\rho, \varphi, z) = R(\rho)\Phi(\varphi)Z(z)$, then we can go through the same steps as we did for the Cartesian system to get

$$\frac{1}{\rho R} \partial_\rho(\rho \partial_\rho R) + \frac{1}{\rho^2 \Phi} \partial_\varphi^2 \Phi = -\frac{1}{Z} \partial_z^2 Z = \lambda \text{ const.} \quad (3.5.17)$$

The solution for $Z(z)$ is evidently

$$Z(z) = e^{\pm i\sqrt{\lambda}z} \quad (3.5.18)$$

The other equation

$$\frac{1}{\rho R} \partial_\rho(\rho \partial_\rho R) - \lambda = -\frac{1}{\rho^2 \Phi} \partial_\varphi^2 \Phi \quad (3.5.19)$$

can be re-written in the form

$$\frac{\rho}{R} \partial_\rho(\rho \partial_\rho R) - \lambda \rho^2 = -\frac{1}{\Phi} \partial_\varphi^2 \Phi = \sigma \quad (3.5.20)$$

This immediately gives the solution for $\Phi(\varphi)$ in the form

$$\Phi(\varphi) = e^{\pm i\sqrt{\sigma}\varphi} \quad (3.5.21)$$

Now φ is a compact coordinate of period 2π . For the function $\phi(\vec{r})$ to be single-valued, it must be periodic in φ . This is only possible if $\sqrt{\sigma} = n$, where n is an integer,

$$\Phi(\varphi) = e^{\pm in\varphi}. \quad (3.5.22)$$

The radial equation is then

$$\frac{1}{\rho} \partial_\rho(\rho \partial_\rho R) - \left(\lambda + \frac{n^2}{\rho^2} \right) R = 0 \quad (3.5.23)$$

The solutions to this equation are given in terms of the *modified* Bessel functions, which are just Bessel functions of purely imaginary argument. We have

$$R(\rho) = J_{\pm n}(i\sqrt{\lambda}\rho) \quad (3.5.24)$$

provided that $\lambda \neq 0$. Unfortunately, the Bessel functions $J_n(z)$ and $J_{-n}(z)$ are not linearly independent. This is evident from the series representation

$$J_{\pm n}(z) = \left(\frac{z}{2}\right)^{\pm n} \sum_{j=0}^{\infty} \frac{(-)^j}{j! \Gamma(j \pm n + 1)} \left(\frac{z}{2}\right)^{2j} \quad (3.5.25)$$

from which it follows that

$$J_{-n}(z) = (-)^n J_n(z). \quad (3.5.26)$$

We must therefore find another linearly independent set of solutions, which are customarily taken to be the Hankel function of the first kind. The usual choices of independent solutions are denoted by

$$\begin{aligned} I_n(z) &= i^{-n} J_n(iz) \\ K_n(z) &= \frac{\pi}{2} i^{n+1} H_n^{(1)}(iz) \end{aligned} \quad (3.5.27)$$

where $H^{(1)}(z)$ is the Hankel function of the first kind. These are real functions of their arguments, whose limiting behaviors are

$$\begin{aligned} x \ll 1 \quad I_n(x) &\rightarrow \frac{1}{\Gamma(n+1)} \left(\frac{x}{2}\right)^n \quad \forall n \\ K_n(x) &\rightarrow -0.5772 - \ln \frac{x}{2} \dots \quad n = 0 \\ &\rightarrow \frac{\Gamma(n)}{2} \left(\frac{2}{x}\right)^n \quad n \neq 0 \\ x \gg 1 \quad I_n(x) &\rightarrow \frac{1}{\sqrt{2\pi x}} e^x \left[1 + \mathcal{O}\left(\frac{1}{x}\right)\right] \\ K_n(x) &\rightarrow -0.5772 - \ln \frac{x}{2} \dots \end{aligned}$$

The solutions may be written as linear combinations of

$$I_n(\sqrt{\lambda}\rho) e^{\pm i\sqrt{\lambda}z} e^{\pm in\varphi}, \quad K_n(\sqrt{\lambda}\rho) e^{\pm i\sqrt{\lambda}z} e^{\pm in\varphi}, \quad (3.5.28)$$

analogous to the case of rectangular symmetry.

Many problems of interest in electrostatics are at least effectively cylindrically symmetric, which means that there is no significant dependence on z , or $Z(z) = 1$. Let us examine this case first because it is easier. No z dependence is only possible if $\lambda = 0$. Then,

$$\rho^2 \frac{d^2 R}{d\rho^2} + \rho \frac{dR}{d\rho} - n^2 R = 0 \quad (3.5.29)$$

But, we can use the fact that if $u = \ln \rho$ then

$$\frac{d^2 R}{du^2} = \rho^2 \frac{d^2 R}{d\rho^2} + \rho \frac{dR}{d\rho} \quad (3.5.30)$$

so that the equation, written in terms of u is

$$\frac{d^2 R}{du^2} - n^2 R = 0 \Rightarrow R(\rho) = \begin{cases} \rho^{\pm n} & n \neq 0 \\ \ln \rho & n = 0 \end{cases} \quad (3.5.31)$$

The full solution explicitly as a superposition of all the solutions we have obtained can be written as

$$\phi(\rho, \varphi) = A + B \ln \rho + \sum_{n=-\infty, n \neq 0}^{\infty} \rho^n (C_n \cos n\varphi + D_n \sin n\varphi) \quad (3.5.32)$$

and all the constants must be evaluated through the boundary conditions.

Consider the following example: suppose that we want to obtain the electric field both inside and outside a hollow cylindrical shell of radius R and of infinite extent along the z axis, whose surface potential is fixed to vary as some known function $\phi(R, \phi) = \tilde{\phi}(\varphi)$. This can always be arranged by building the shell of mutually insulated conductors, each of which is held at the appropriate fixed potential. We can set $B = 0$ simply because there is no line charge along the z axis and the contribution of $\ln \rho$ would correspond to precisely such a line charge. We will consider the interior separately from the exterior. If the potential is to be finite along the z axis ($\rho = 0$), we should eliminate all terms with negative powers of ρ . Then

$$\tilde{\phi}(\varphi) = A + \sum_{n=1}^{\infty} R^n (C_n \cos n\varphi + D_n \sin n\varphi) \quad (3.5.33)$$

giving

$$A = \frac{1}{2\pi} \int_0^{2\pi} d\varphi \tilde{\phi}(\varphi)$$

$$\begin{aligned}
C_n &= \frac{1}{\pi R^n} \int_0^{2\pi} d\varphi \tilde{\phi}(\varphi) \cos n\varphi \\
D_n &= \frac{1}{\pi R^n} \int_0^{2\pi} d\varphi \tilde{\phi}(\varphi) \sin n\varphi
\end{aligned} \tag{3.5.34}$$

On the other hand, in the exterior we eliminate all terms that have *positive* powers of ρ to ensure the finiteness of ϕ at infinity. Then

$$\tilde{\phi}(\varphi) = A + \sum_{n=1}^{\infty} R^{-n} (C_{-n} \cos n\varphi - D_{-n} \sin n\varphi) \tag{3.5.35}$$

and, as before,

$$\begin{aligned}
A &= \frac{1}{2\pi} \int_0^{2\pi} d\varphi \tilde{\phi}(\varphi) \\
C_{-n} &= \frac{R^n}{\pi} \int_0^{2\pi} d\varphi \tilde{\phi}(\varphi) \cos n\varphi \\
D_{-n} &= -\frac{R^n}{\pi} \int_0^{2\pi} d\varphi \tilde{\phi}(\varphi) \sin n\varphi
\end{aligned} \tag{3.5.36}$$

As a special example, suppose $\tilde{\phi}(\varphi) = \phi_0 \sin \varphi$. Then in the interior, clearly $A = 0$ and, by the orthogonality of the functions $\sin n\varphi$ and $\cos n\varphi$, we find also that $C_n = 0 \forall n$, $D_n = 0 \forall n > 1$ and $n < 0$, $D_1 = \phi_0/R$. The solution in the interior is then

$$\phi_{\text{in}}(\rho, \varphi) = \phi_0 \frac{\rho}{R} \sin \varphi \tag{3.5.37}$$

and the electric field is

$$\vec{E}_{\text{in}} = -\frac{\phi_0}{R} (\sin \varphi \hat{\rho} + \cos \varphi \hat{\varphi}) = -\frac{\phi_0}{R} \hat{y} \tag{3.5.38}$$

and points in the (negative) y direction. This is not surprising, given the symmetry of the potential on the cylindrical shell.

In the exterior of the shell, the same arguments give $A = 0$, $C_{-n} = 0 \forall n$, $D_{-n} = 0 \forall n \neq 1$, $D_{-1} = -\phi_0 R$. We obtain the exterior solution

$$\phi_{\text{out}}(\rho, \varphi) = \phi_0 \frac{R}{\rho} \sin \varphi \tag{3.5.39}$$

The electric field is

$$\vec{E}_{\text{out}} = \phi_0 \frac{R}{\rho^2} (\sin \varphi \hat{\rho} - \cos \varphi \hat{\varphi}) \tag{3.5.40}$$

so the azimuthal component of \vec{E} is continuous across the shell, but the radial component is discontinuous. The electrostatic potential is continuous across the shell.

We return to the case $\lambda \neq 0$. Let us consider the following concrete problem, which is similar in spirit to the problem we considered while examining solutions in Cartesian coordinates. Imagine that we want the electrostatic potential inside a closed right circular cylinder (instead of inside a rectangular box) along the z axis, of radius R and length L situated so that one flat face of the cylinder lies in the $x - y$ plane. Suppose we arrange for all faces of the cylinder to be at some constant potential ϕ_0 , except for the face at $z = L$, which we assume is maintained at some potential $\tilde{\phi}(\rho, \varphi)$. As before, let us first look for the potential that is identically zero on all the faces except the one at $z = L$. (We can always add a constant to the potential later.) Now the functions $K_n(z)$ are badly behaved in the limit as z approaches zero, so we must consider a linear combination only of the $I_n(z)$, or equivalently of the modified Bessel functions, $J_{\pm n}(i\sqrt{\lambda}\rho)$. We require the electrostatic potential to vanish at $\rho = R$, the cylindrical surface, which means that the Bessel function should admit at least one root. This is possible only if the argument of the Bessel function is real, so $\lambda < 0$. The Bessel function, J_n , then has a countably infinite number of roots, which we label by q_{nm} , and the vanishing of $\phi(\vec{r})$ at $\rho = R$ implies that

$$\sqrt{|\lambda_{nm}|} = \frac{q_{nm}}{R} \quad (3.5.41)$$

Now, because $\lambda < 0$ the function $Z(z)$ is a linear combination of

$$Z(z) = e^{\pm\sqrt{|\lambda_{nm}|}z} \quad (3.5.42)$$

If it is to vanish on the $x - y$ plane, only the combination

$$Z(z) = Z_0 \sinh(\sqrt{|\lambda_{nm}|}z) \quad (3.5.43)$$

is possible. Therefore we have the solution

$$\phi(\rho, \varphi, z) = \phi_0 + \sum_{n=-\infty, m=1}^{\infty} c_{nm} J_n(\sqrt{|\lambda_{nm}|}\rho) \sinh(\sqrt{|\lambda_{nm}|}z) e^{in\varphi} \quad (3.5.44)$$

As we have seen earlier, the $J_{-n}(w)$ are not independent of the $J_n(w)$. In fact, because

$$J_{-n}(w) = (-)^n J_n(w), \quad w \in \mathbb{C}, \quad (3.5.45)$$

the solution can also be written as

$$\phi(\rho, \varphi, z) = \phi_0 + \sum_{n=0, m=1}^{\infty} J_n(\sqrt{|\lambda_{nm}|}\rho) \sinh(\sqrt{|\lambda_{nm}|}z) (a_{nm} \cos n\varphi + b_{nm} \sin n\varphi) \quad (3.5.46)$$

On the face at $z = L$, the solution must be of the form

$$\tilde{\phi}(\rho, \varphi) = \phi_0 + \sum_{n=0, m=1}^{\infty} J_n(\sqrt{|\lambda_{nm}|}\rho)(A_{nm} \cos n\varphi + B_{nm} \sin n\varphi) \quad (3.5.47)$$

where

$$\begin{aligned} A_{nm} &= a_{nm} \sinh(\sqrt{|\lambda_{nm}|}L) \\ B_{nm} &= b_{nm} \sinh(\sqrt{|\lambda_{nm}|}L) \end{aligned} \quad (3.5.48)$$

The coefficients are determined using the orthogonality properties of the Bessel functions

$$\int_0^a dw w J_n\left(q_{nm} \frac{w}{a}\right) J_n\left(q_{nl} \frac{w}{a}\right) = \delta_{ml} \frac{a^2}{2} J_{n+1}^2(q_{nm}). \quad (3.5.49)$$

Applying this directly to our case,

$$\begin{aligned} A_{nm} &= \frac{2}{\pi R^2 J_{n+1}^2(\sqrt{|\lambda_{nm}|}R)} \int_0^{2\pi} d\varphi \int_0^R d\rho \rho (\tilde{\phi}(\rho, \varphi) - \phi_0) J_n(\sqrt{|\lambda_{nm}|}\rho) \cos n\varphi \\ B_{nm} &= \frac{2}{\pi R^2 J_{n+1}^2(\sqrt{|\lambda_{nm}|}R)} \int_0^{2\pi} d\varphi \int_0^R d\rho \rho (\tilde{\phi}(\rho, \varphi) - \phi_0) J_n(\sqrt{|\lambda_{nm}|}\rho) \sin n\varphi \end{aligned} \quad (3.5.50)$$

provided that for $n = 0$ we use half the value of A_{0m} calculated from the above. We see again that if the potential at the face $z = L$ is held at the same potential as the others, the only possible solution is the trivial one, $\phi(\vec{r}) = \phi_0$.

3.5.3 Problems with Spherical Symmetry

Laplace's equation in spherical coordinates may be separated by requiring

$$\phi(\vec{r}) = R(r)\Theta(\theta)\Phi(\varphi) \quad (3.5.51)$$

where $R(r)$ is the "radial" function, $\Theta(\theta)$ is the polar function and $\Phi(\varphi)$ is the azimuthal function. We begin by writing the equation for $\phi(\vec{r})$ as

$$\frac{\sin^2 \theta}{R(r)} \partial_r (r^2 \partial_r) R + \frac{\sin \theta}{\Theta(\theta)} \partial_\theta (\sin \theta \partial_\theta) \Theta = -\frac{1}{\Phi(\varphi)} \partial_\varphi^2 \Phi = m^2 \quad (3.5.52)$$

where we have already used the condition that the arbitrary constant that would appear on the extreme right hand side must be the square of an integer in order that $\Phi(\varphi)$ be one to one. This gives

$$\Phi(\varphi) = e^{\pm im\varphi} \quad (3.5.53)$$

and we may write the remaining equation in the form

$$\frac{1}{R(r)}\partial_r(r^2\partial_r)R = -\frac{1}{\Theta\sin\theta}\partial_\theta(\sin\theta\partial_\theta)\Theta + \frac{m^2}{\sin^2\theta} = \sigma \quad (3.5.54)$$

where σ is (so far) an arbitrary constant. This gives us two equations, *viz.*, the polar and the radial respectively,

$$\begin{aligned} -\frac{1}{\sin\theta}\partial_\theta(\sin\theta\partial_\theta)\Theta + \frac{m^2\Theta}{\sin^2\theta} - \sigma\Theta &= 0 \\ \partial_r(r^2\partial_r)R - \sigma R &= 0 \end{aligned} \quad (3.5.55)$$

The first of these equations is solved by a method called the Frobenius expansion. It turns out that solutions exist if and only if $\sigma = l(l+1)$, where l is a whole number and $|m| \leq l$. In this case,

$$\frac{m^2\Theta}{\sin^2\theta} - \frac{1}{\sin\theta}\partial_\theta(\sin\theta\partial_\theta)\Theta - l(l+1)\Theta = 0 \quad (3.5.56)$$

and the solutions are the associated Legendre polynomials in $\cos\theta$, $\mathcal{P}_{l,m}(\cos\theta)$. Some associated Legendre polynomials are given below:

$$\mathcal{P}_{0,0}(\cos\theta) = 1$$

$$\mathcal{P}_{1,0}(\cos\theta) = \cos\theta$$

$$\mathcal{P}_{1,1}(\cos\theta) = \sin\theta$$

$$\mathcal{P}_{2,0}(\cos\theta) = \frac{1}{2}(3\cos^2\theta - 1)$$

$$\mathcal{P}_{2,1}(\cos\theta) = 3\cos\theta\sin\theta$$

$$\mathcal{P}_{2,2}(\cos\theta) = 3\sin^2\theta$$

$$\mathcal{P}_{3,0}(\cos\theta) = \frac{1}{2}(5\cos^3\theta - 3\cos\theta)$$

$$\mathcal{P}_{3,1}(\cos\theta) = \frac{3}{2}(5\cos^2\theta - 1)\sin\theta$$

$$\mathcal{P}_{3,2}(\cos\theta) = 15\cos\theta\sin^2\theta$$

$$\mathcal{P}_{3,3}(\cos\theta) = 15\sin^3\theta$$

$$\dots \quad (3.5.57)$$

and

$$\mathcal{P}_{l,-m}(x) = (-)^m \frac{(l-m)!}{(l+m)!} \mathcal{P}_{l,m}(x) \quad (3.5.58)$$

is used to find the associated Legendre polynomials when $m < 0$.

The radial equation becomes

$$\frac{d^2 R}{dr^2} + \frac{2}{r} \frac{dR}{dr} - \frac{l(l+1)}{r^2} R = 0 \quad (3.5.59)$$

This can be reduced to a simple linear, second order differential equation with constant coefficients in the variable $z = \ln r$. The two independent solutions are found to be

$$R_l(r) = r^l, \quad r^{-(l+1)} \quad (3.5.60)$$

Therefore the general solution to Laplace's equation in spherical coordinates is

$$\phi(r, \theta, \phi) = \sum_{l=0}^{\infty} \sum_{m=-l}^l [a_{lm} r^l + b_{lm} r^{-(l+1)}] \mathcal{P}_{l,m}(\cos \theta) e^{im\phi}. \quad (3.5.61)$$

Many problems of interest involve azimuthal symmetry, where there is no ϕ dependence. The absence of any dependence on the azimuthal angle implies that $m = 0$. In this case, the associated Legendre polynomials reduce to the Legendre polynomials, $P_l(\cos \theta) = \mathcal{P}_{l,0}(\cos \theta)$, and the general solution takes the form

$$\phi(r, \theta) = \sum_{l=0}^{\infty} [a_l r^l + b_l r^{-(l+1)}] P_l(\cos \theta) \quad (3.5.62)$$

If we are interested in the electrostatic potential inside a spherical surface of radius R , then $b_l = 0 \forall l$, assuming that there are no charges at the origin so that the potential is finite there. We have,

$$\tilde{\phi}(\theta) = \sum_{l=0}^{\infty} A_l P_l(\cos \theta), \quad (3.5.63)$$

where $A_l = a_l R^l$. Using the orthogonality properties of the Legendre polynomials, we find the coefficients a_l from⁶

$$A_l = \left(l + \frac{1}{2}\right) \int_0^\pi d\theta \sin \theta V(\theta) P_l(\cos \theta). \quad (3.5.65)$$

⁶Problem: Work out the following example: imagine a spherical surface made up of two hemispheres of radius R that are maintained at opposite potentials, *i.e.*,

$$\phi(R, \theta) = \begin{cases} +\phi_0 & 0 \leq \theta < \frac{\pi}{2} \\ -\phi_0 & \frac{\pi}{2} < \theta \leq \pi \end{cases} \quad (3.5.64)$$

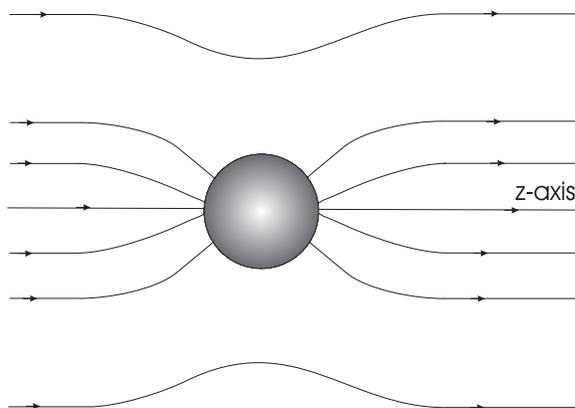


Figure 3.4: Spherical conductor in a uniform electric field.

As another application, let us examine the problem of a conducting sphere of radius R placed in a uniform electric field, which we take for convenience to be in the direction of the positive z axis. The presence of the conducting sphere will distort the electric field in its neighborhood as the field lines must be normal to the sphere at its surface. There is azimuthal symmetry in this problem, so we can use the result above. First we note that, far from the sphere, the effect of the conducting sphere should vanish, *i.e.*,

$$\lim_{r \rightarrow \infty} E(\vec{r}) = E_0 \hat{z} \quad (3.5.66)$$

This means that

$$\lim_{r \rightarrow \infty} \phi(\vec{r}) = -E_0 z = -E_0 r \cos \theta + \text{const.} \quad (3.5.67)$$

Observe therefore, that $a_l = 0 \forall l > 1$ and

$$\phi(\vec{r}) = a_0 + a_1 r \cos \theta + \frac{b_0}{r} + \frac{b_1}{r^2} \cos \theta + \frac{b_2}{r^3} (3 \cos^2 \theta - 1) + \dots \quad (3.5.68)$$

with $a_1 = -E_0$. Again, the potential can only be constant on the surface of the conducting sphere, so $\phi(r = R, \theta) = \text{const.}$ As there is no θ dependence on the surface, this gives

$$\begin{aligned} -E_0 R + \frac{b_1}{R^2} &= 0 \Rightarrow b_1 = E_0 R^3 \\ b_j &= 0 \quad \forall j \geq 2 \end{aligned} \quad (3.5.69)$$

The appearance of term b_0/r would indicate that the conductor bears a net charge, which is not the case, so $b_0 = 0$. Then, we have

$$\phi(\vec{r}) = \phi_0 - E_0 r \cos \theta \left(1 - \frac{R^3}{r^3} \right) \quad (3.5.70)$$

where $a_0 = \phi_0$ is the potential on the conductor. We can now calculate the electric field everywhere,

$$\vec{E}(\vec{r}) = -\vec{\nabla}\phi(\vec{r}) = E_0 \left(1 + 2\frac{R^3}{r^3}\right) \cos\theta \hat{r} - E_0 \left(1 - \frac{R^3}{r^3}\right) \sin\theta \hat{\theta} \quad (3.5.71)$$

When $r \rightarrow \infty$ we find that indeed $\vec{E} = E_0 \hat{z}$. Having the electric field everywhere, we obtain also the charge density on the conductor by noting that

$$\lim_{r \rightarrow R} \vec{E}(\vec{r}) = \frac{\sigma}{\epsilon_0} \hat{r} \Rightarrow \sigma = 3\epsilon_0 E_0 \cos\theta \quad (3.5.72)$$

Since we have assumed that the sphere was uncharged, the total charge on the sphere should vanish. We can integrate over the surface of the conductor to find that the total charge,

$$Q = 3\epsilon_0 E_0 R^2 \int_0^{2\pi} d\varphi \int_0^\pi d\theta \sin\theta \cos\theta \equiv 0 \quad (3.5.73)$$

vanishes as it should. Although the total charge on the sphere is non-vanishing the sphere acquires an electric dipole moment on account of the way in which the charge is distributed on its surface. Being confined to the surface of the conductor, we may describe the charge distribution as

$$\rho(\vec{r}') = 3\epsilon_0 E_0 \cos\theta \delta(r' - R) \quad (3.5.74)$$

where the δ -function takes into account the fact that the charge resides on the surface of the conductor. Integrating this we find

$$\begin{aligned} \vec{p} &= 3\epsilon_0 E_0 \int d\Omega' \int_0^\infty dr' r' \hat{r}' \cos\theta' \delta(r' - R) \\ &= 3\epsilon_0 E_0 R \int_0^{2\pi} d\varphi' \int_0^\pi d\theta' \sin\theta' \cos\theta' (\sin\theta' \cos\varphi', \sin\theta' \sin\varphi', \cos\theta') \end{aligned} \quad (3.5.75)$$

Naturally, only the z component is non-vanishing. We easily find it to be $p_z = 4\pi\epsilon_0 E_0 R$.

Chapter 4

Boundary Value Problems: Poisson's Equation

So far we have concentrated on obtaining general solutions to Laplace's equation, *i.e.*, on electrostatics in a vacuum. When sources are present we must solve Poisson's equation, which is generally a more difficult problem. In this case it is often useful to exploit the boundary conditions from the beginning. This is the subject of this chapter. Methods developed in this chapter will naturally be applicable to vacuum electrostatics too, since this condition is obtained from the general situation when the charge density, $\rho(\vec{r}')$, vanishes everywhere. We begin with the "method of images" and then develop more formal and systematic methods.

4.1 The Method of Images

Solutions to Poisson's equation that obey the boundary conditions, either Dirichlet or Neumann, are unique regardless of the methods used to arrive at them. So far we have considered a systematic procedure to find solutions of the electrostatic vacuum and we have seen that it is limited by the high degree of symmetry necessary to make Laplace's equation tractable. Poisson's equation is even more difficult in general and different approaches to the problem of finding solutions to it are always welcome. In this section we introduce one such method, which is more of an art form than an algorithmic procedure but nonetheless very important. It is called the "method of images" and usually applies to problems involving one or more point charges in the presence of conducting (or equipotential) surfaces. Often these problems do not exhibit any definite symmetry, but sometimes from the geometry of the constant potential surfaces it is possible that a small number of additional charges of appropriate magnitudes placed at judicious points in space can yield a solution. These charges are called "image" charges and the electrostatic potential is

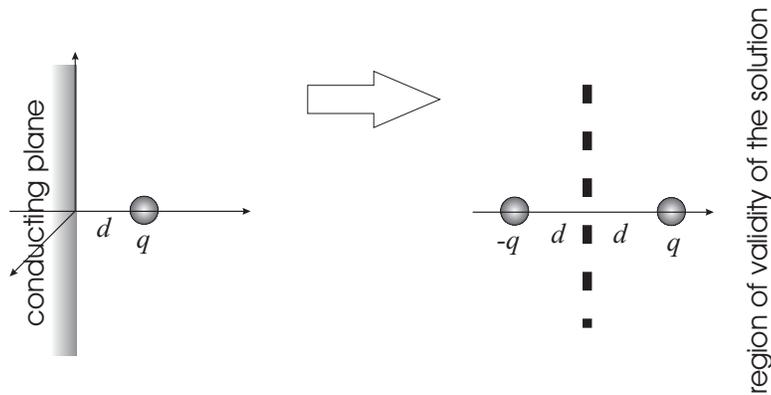


Figure 4.1: Problem of a point charge and a conducting plane.

obtained only in regions *excluding* those occupied by the image charges. *Without* actually solving Poisson's equation, this method can yield solutions to problems that would be difficult if not impossible to otherwise obtain.

4.1.1 Point charge near infinite conducting planes.

To motivate the “method of images”, imagine that we are interested in the electric field due to a charge that is placed near an infinite conducting plane (see figure (4.1)). The boundary conditions are that the electric potential must vanish on the plane and at infinity. We desire the field on the side of the plane containing the charge, q . The symmetry is neither rectangular nor spherical, as it would be either for the plane or the point charge alone.

For definiteness, let the $y - z$ plane be the desired surface of constant potential and let the charge be situated a distance d from it on the positive x axis. Consider now an alternate problem, which we shortly show to equivalent. Ignore the fact that our problem is defined only to the right of the plane and place an equal and opposite charge to the left of it at the same distance, d . The electric potential at any point \vec{r} , due to the two charges, is easy to write down

$$\phi(\vec{r}) = \phi_q(\vec{r}) + \phi_{-q}(\vec{r}) = \frac{q}{4\pi\epsilon_0} \left[\frac{1}{r_+} - \frac{1}{r_-} \right] \quad (4.1.1)$$

and

$$r_+ = \sqrt{(x-d)^2 + y^2 + z^2}, \quad r_- = \sqrt{(x+d)^2 + y^2 + z^2} \quad (4.1.2)$$

At $x = 0$, which is supposed to be the plane at zero potential, we have $r_+ = r_-$ and therefore $\phi(\vec{r}) = 0$ as required. Furthermore, the potential approaches zero at infinity

and, because we have used only the electric potentials due to the point charges, $\phi(\vec{r})$ is a solution of Laplace's equation. It is therefore the unique solution to the problem, but only on the right hand side of the plane because the charge $-q$ that was introduced on the left does not really exist – it is an “image charge”. The electric field is

$$\vec{E}(\vec{r}) = -\vec{\nabla}\phi(\vec{r}) = \frac{q}{4\pi\epsilon_0} \left[\frac{\vec{r}_+}{r_+^3} - \frac{\vec{r}_-}{r_-^3} \right] \quad (4.1.3)$$

Very near the plane, in the limit as $x \rightarrow 0$, $r_+^3 \approx r_-^3$ and $(\vec{r}_+ - \vec{r}_-) = -2d\hat{x}$, therefore

$$\lim_{x \rightarrow 0} \vec{E}(\vec{r}) = \frac{\sigma}{\epsilon_0} \hat{x} = -\frac{qd}{2\pi\epsilon_0(d^2 + y^2 + z^2)^{3/2}} \hat{x} \Rightarrow \sigma = -\frac{qd}{2\pi(d^2 + y^2 + z^2)^{3/2}} \quad (4.1.4)$$

gives the surface charge density induced by the presence of the charge q on the conducting plane. The total charge induced on the conducting plane is obtained by integrating the surface charge density

$$Q = -\frac{qd}{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{dxdy}{(d^2 + x^2 + y^2)^{3/2}} \quad (4.1.5)$$

This integral may be resolved using polar coordinates in the $x - y$ plane,

$$Q_{\text{plane}} = -\frac{qd}{2\pi} \int_0^{2\pi} d\varphi \int_0^{\infty} \frac{rdr}{(d^2 + r^2)^{3/2}} = \frac{qd}{\sqrt{d^2 + r^2}} \Big|_0^{\infty} = -q \quad (4.1.6)$$

as we expect. The charge q exerts a force on the conducting plane, and vice versa by Newton's third law. It is conceptually easier to see what the force exerted by the plane on q would be, because this must be precisely the force exerted by the image charge, $-q$ on q ,

$$\vec{F}_q = -\frac{q^2}{16\pi\epsilon_0 d^2} \hat{x} \quad (4.1.7)$$

This implies that

$$\vec{F}_{\text{plane}} = +\frac{q^2}{16\pi\epsilon_0 d^2} \hat{x}. \quad (4.1.8)$$

is the force exerted by the point charge on the plane.

To generalize this problem, consider two conducting planes, making an angle θ relative to each other as shown in figure (4.2), where θ (measured in radians) divides 2π . Suppose we are interested in the electrostatic field within the opening between the planes and suppose that opening contains a physical charge q . Suppose also that the electrostatic potential vanishes on the planes. We may solve this problem by introducing image charges as follows: think of the planes as mirrors and introduce image charges, $(-)^n q$, at each point at which the “mirrors” would form images of the physical charge where n is the order of

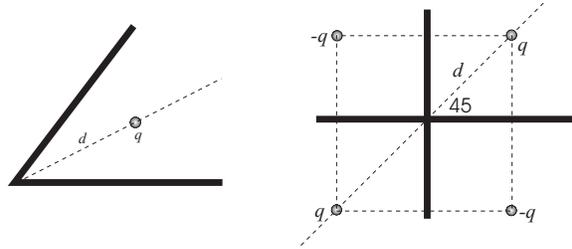


Figure 4.2: Problem of a point charge and two conducting planes.

the reflection. If θ divides 2π (the only case we consider), there will be $N = 2\pi/\theta - 1$ such images, and we may write down the electrostatic potential due to all $2\pi/\theta$ charges (including the physical charge). It should be clear that the electrostatic potential so obtained will vanish on the planes and, moreover, will approach zero at infinity. The solution is unique because the boundary conditions are satisfied, but it is valid only in the opening between the planes that contains the physical charge. The special case of perpendicular conducting planes as shown on the right of figure (4.2) will illustrate the method.

For mutually perpendicular planes, we need three image charges. The electrostatic potential due to the four charges is

$$\phi(\vec{r}) = \frac{q}{4\pi\epsilon_0} \left[\frac{1}{|\vec{r} - \vec{r}_0|} - \frac{1}{|\vec{r} - \vec{r}_2|} + \frac{1}{|\vec{r} + \vec{r}_0|} - \frac{1}{|\vec{r} - \vec{r}_4|} \right] \quad (4.1.9)$$

where \vec{r}_0 is the position of the physical charge $+q$, and $\vec{r}_{2,4}$ locate the image charge in the second and fourth quadrant respectively. The image charge in the third quadrant has position vector $-\vec{r}_0$ and, because it corresponds to a second order reflection, must have charge $+q$.¹

4.1.2 Point charge outside a grounded, conducting sphere.

A simple example involving curved conducting surfaces would be that of a point charge located outside a grounded conducting sphere of radius R and at a distance d from its center (see figure (4.3)). With no loss of generality we select the z axis to run from the center of the sphere through the point charge. There is azimuthal symmetry in this problem, but the methods above are not very easy to implement. Instead, let us place a point like image charge, q' , on the inside of the sphere and at a distance a from the center on the z axis.

¹Problem: Determine the charge on the conducting planes and the forces exerted by the charge q on each of them. Calculate the total force on the charge q .

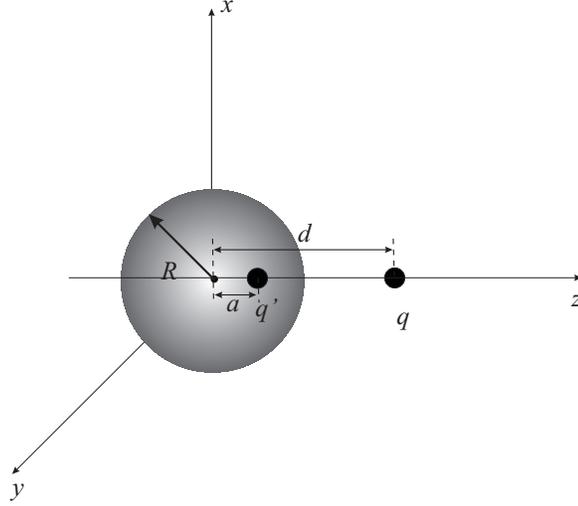


Figure 4.3: Problem of a point charge and a conducting sphere.

We shall calculate q' and a so that the electrostatic potential vanishes on the conducting surface and, because our sources are point like, it will automatically vanish at infinity. The electrostatic potential due to our point charges is given by

$$\phi(\vec{r}) = \frac{1}{4\pi\epsilon_0} \left[\frac{q}{|\vec{r} - \vec{r}_0|} + \frac{q'}{|\vec{r} - \vec{r}'_0|} \right] \quad (4.1.10)$$

where $\vec{r}_0 = (0, 0, d)$ and $\vec{r}'_0 = (0, 0, a)$ represent the positions of the charges q and q' respectively. Using

$$|\vec{r} - \vec{r}_0| = \sqrt{r^2 + r_0^2 - 2\vec{r} \cdot \vec{r}_0} = \sqrt{r^2 + d^2 - 2dr \cos \theta} \quad (4.1.11)$$

where θ is the polar angle. For points on the sphere, we therefore have

$$\phi(R, \theta) = \frac{1}{4\pi\epsilon_0} \left[\frac{q}{\sqrt{R^2 + d^2 - 2Rd \cos \theta}} + \frac{q'}{\sqrt{R^2 + a^2 - 2Ra \cos \theta}} \right], \quad (4.1.12)$$

which should vanish. This is possible to accomplish because we have two adjustable parameters, *viz.*, q' and a . Setting $q' = -\alpha q$, a vanishing potential on the sphere implies that

$$\alpha^2(R^2 + d^2 - 2Rd \cos \theta) = R^2 + a^2 - 2Ra \cos \theta \quad (4.1.13)$$

Comparing the $\cos \theta$ terms gives

$$\alpha^2 = \frac{a}{d} \quad (4.1.14)$$

and therefore that

$$\frac{a}{d}(R^2 + d^2) = R^2 + a^2 \quad (4.1.15)$$

which is a quadratic equation for a with solutions $a = d$ and $a = R^2/d$. We must reject the first of these solutions as it places the image charge outside the sphere, indeed it places an image charge of equal magnitude and opposite sign precisely at the location of the physical charge, q , thereby annihilating it. The second solution is reasonable and gives $q' = -\alpha q = -Rq/d$.

Together with the previous example, this example illustrates why this way of finding solutions is called the method of images. Imagine that the sphere is a convex mirror, and the charge is a point like object situated at a distance o from it on the z axis. From elementary geometric optics we expect an image of the “object” to form a distance i from the surface, where

$$\frac{1}{i} + \frac{1}{o} = -\frac{1}{R} \quad (4.1.16)$$

where R is the radius of curvature. By placing the charge q a distance d from the center of the sphere we have, in this analogy, an object situated at a distance $o = d - R$ from the “mirror”. We expect its image to be formed at

$$\frac{1}{i} = -\frac{1}{R} - \frac{1}{o} \Rightarrow i = -\frac{R(d - R)}{d} \quad (4.1.17)$$

This is measured from the “mirror” and the negative sign implies that it is to the left. Thus the image would form a distance $R - |i| = R^2/d$ from the center, which is precisely where the image charge should be located. The magnification would be

$$M = \frac{i}{o} = -\frac{R}{d}, \quad (4.1.18)$$

which we recognize as the proportionality factor between the image charge and the physical one.

The electrostatic potential is given everywhere by

$$\phi(\vec{r}) = \frac{q}{4\pi\epsilon_0} \left[\frac{1}{\sqrt{r^2 + d^2 - 2rd\cos\theta}} - \frac{R}{\sqrt{r^2d^2 + R^4 - 2rdR^2\cos\theta}} \right]. \quad (4.1.19)$$

It is interesting to consider the electric charge induced on the conductor. We evaluate the electric field on the surface of the conductor as usual, the only surviving component being the radial one,²

$$E_r(R, \theta) = -\partial_r\phi(\vec{r})|_{r=R} = -\frac{q}{4\pi\epsilon_0 R} \frac{d^2 - R^2}{(R^2 + d^2 - 2Rd\cos\theta)^{3/2}} \quad (4.1.20)$$

²Exercise: Verify this statement.

This is related to the surface charge density induced on the spherical surface,

$$\sigma = \epsilon_0 E_r(R, \theta) = -\frac{q}{4\pi R} \frac{d^2 - R^2}{(R^2 + d^2 - 2Rd \cos \theta)^{3/2}} \quad (4.1.21)$$

which should yield a charge on the sphere precisely equal to the image charge introduced, as we find by integrating,

$$Q = -\frac{qR(d^2 - R^2)}{4\pi} \int_0^{2\pi} d\varphi \int_0^\pi \frac{d\theta \sin \theta}{(R^2 + d^2 - 2Rd \cos \theta)^{3/2}} = -q \frac{R}{d} = q' \quad (4.1.22)$$

As before, we may also calculate the force exerted on the sphere by the charge q . The simplest way to do this is to first calculate the force exerted by the image charge q' on q , which would give the force exerted on the charge q by the sphere. The force exerted by the charge q on the sphere follows by Newton's third law.³

4.1.3 Point charge outside an insulated, conducting sphere.

In the previous problem, we have seen that the external charge induces a charge on the sphere, equal to $-qR/d$, which distributes itself on the surface of the sphere in such a way as to achieve equilibrium. If the sphere itself is charged with total charge $Q \neq q'$, we can build the solution on the outside of the sphere by superposition. Imagine that the sphere is initially grounded so that the solution we have obtained in the previous section holds. Then disconnect the sphere from ground and charge the sphere by introducing an amount of charge equal to $Q - q'$. The charge we have just introduced will distribute itself uniformly over the sphere because the forces due to the external charge q are precisely balanced by the induced charge q' . Appealing to Gauss' law, the potential due to the remaining charge is just that of a point charge located at the center of the sphere, and we can write the solution as

$$\phi(\vec{r}) = \frac{1}{4\pi\epsilon_0} \left[\frac{q}{|\vec{r} - \vec{r}'_0|} - \frac{qR/d}{|\vec{r} - \vec{r}'_0|} + \frac{Q + qR/d}{|\vec{r}|} \right] \quad (4.1.23)$$

where, as before, \vec{r}'_0 locates the image charge. The force on q is just the force due to the image charge and the additional charge Q (located at the center).

4.1.4 Conducting sphere in a uniform external electric field.

For our last example of the method of images, we will reconsider the problem of a grounded, conducting sphere in an uniform, external electric field from the point of view of image

³Problem: Alternatively, calculate the force exerted by the sphere on the charge q by integrating the force exerted on q by elements of the sphere, using the charge density in (4.1.21).

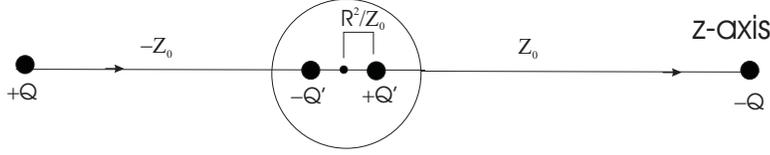


Figure 4.4: Problem of a conducting sphere in an uniform, external field.

charges. The idea here is that the uniform electric field can be simulated by two point charges of opposite sign and magnitude Q located at $\pm\infty$ on the z -axis. Actually, we will begin by considering the point charges to be located at $\pm Z_0$ on the z -axis and take the limit as $Z_0 \rightarrow \infty$ in the end in order to recover the conditions of the problem.

We must introduce two image charges, $Q' = \pm QR/Z_0$, inside the conducting sphere as shown in figure, located at R^2/Z_0 from the center. The electric potential outside the sphere, at point $P(\vec{r})$, is

$$\begin{aligned} \phi(\vec{r}) = & \frac{Q}{\sqrt{r^2 + Z_0^2 + 2rZ_0 \cos \theta}} - \frac{Q}{\sqrt{r^2 + Z_0^2 - 2rZ_0 \cos \theta}} \\ & - \frac{aQ/Z_0}{\sqrt{r^2 + \frac{R^4}{Z_0^2} + \frac{2R^2r}{Z_0} \cos \theta}} + \frac{aQ/Z_0}{\sqrt{r^2 + \frac{R^4}{Z_0^2} - \frac{2R^2r}{Z_0} \cos \theta}} \end{aligned} \quad (4.1.24)$$

Retaining only the leading terms as $Z_0 \rightarrow \infty$, we find

$$\phi(\vec{r}) \rightarrow -\frac{2Qr}{Z_0^2} \cos \theta \left(1 - \frac{R^3}{r^3}\right) = -E_0 r \cos \theta \left(1 - \frac{R^3}{r^3}\right) \quad (4.1.25)$$

which is precisely the result we obtained in (3.5.70). The first term, $-E_0 r \cos \theta = -E_0 z$ is just the potential of a uniform field. The second term is the potential due to the induced surface charge density on the sphere or, as we have now come to think of it, the potential due to the image charges.

4.2 The Green Function

The Green function provides a systematic approach to finding solutions of Poisson's equation, while taking into account the nature of the boundary conditions from the start. Let us return to the general solution of this problem as presented in (3.2.7),

$$\phi(\vec{r}) = \frac{1}{4\pi\epsilon_0} \int d^3\vec{r}' G(\vec{r}, \vec{r}') \rho(\vec{r}') - \frac{1}{4\pi} \int_S dS [\phi(\vec{r}')(\hat{n} \cdot \vec{\nabla}' G(\vec{r}, \vec{r}')) - G(\vec{r}, \vec{r}')(\hat{n} \cdot \vec{\nabla}' \phi(\vec{r}'))] \quad (4.2.1)$$

where we have set

$$G(\vec{r}, \vec{r}') = \frac{1}{|\vec{r} - \vec{r}'|}. \quad (4.2.2)$$

The function $G(\vec{r}, \vec{r}')$ satisfies

$$\vec{\nabla}^2 G(\vec{r}, \vec{r}') = \vec{\nabla}'^2 G(\vec{r}, \vec{r}') = -4\pi\delta(\vec{r} - \vec{r}') \quad (4.2.3)$$

and is called the “Green function”. The Green function is not uniquely defined; if $\mathcal{Q}(\vec{r}, \vec{r}')$ is any solution of Laplace's equation, then evidently $\mathcal{G}(\vec{r}, \vec{r}') = G(\vec{r}, \vec{r}') + \mathcal{Q}(\vec{r}, \vec{r}')$ also satisfies (4.2.3) and is therefore also a Green function. We will take the Green function to be quite generally of the form,

$$\mathcal{G}(\vec{r}, \vec{r}') = \frac{1}{|\vec{r} - \vec{r}'|} + \mathcal{Q}(\vec{r}, \vec{r}'), \quad \vec{\nabla}'^2 \mathcal{Q}(\vec{r}, \vec{r}') = 0. \quad (4.2.4)$$

In this way, the problem is reduced to solving Laplace's equation, subject to appropriate conditions. To determine what boundary conditions are applicable when determining a solution to Laplace's equation for $\mathcal{Q}(\vec{r}, \vec{r}')$, consider the following argument. From Green's second identity we have as before [see (3.2.7)]

$$\phi(\vec{r}) = \frac{1}{4\pi\epsilon_0} \int_V d^3\vec{r}' \mathcal{G}(\vec{r}, \vec{r}') \rho(\vec{r}') - \frac{1}{4\pi} \int_S dS \left[\phi \frac{\partial \mathcal{G}}{\partial n} - \mathcal{G} \frac{\partial \phi}{\partial n} \right], \quad (4.2.5)$$

where we have used the short-form

$$\frac{\partial \mathcal{G}}{\partial n} = \hat{n} \cdot \vec{\nabla} \mathcal{G}, \quad \frac{\partial \phi}{\partial n} = \hat{n} \cdot \vec{\nabla} \phi \quad (4.2.6)$$

for the directional derivatives normal to the surface, S . Now although (4.2.5) may appear to be a general solution to the electrostatic problem, it is in fact not usable in the given form because either the Dirichlet or the Neumann boundary conditions on S determines a unique solution to the problem and both cannot be generally specified. We can, however, exploit the ambiguity in the definition of \mathcal{G} to simplify (4.2.5). Suppose that we are given Dirichlet boundary conditions on S . We can use the freedom we have in choosing the function \mathcal{Q} so that \mathcal{G} itself vanishes on S ,

$$\mathcal{G}_D(\vec{r}, \vec{r}')|_S = \left[\frac{1}{|\vec{r} - \vec{r}'|} + \mathcal{Q}_D(\vec{r}, \vec{r}') \right]_S = 0. \quad (4.2.7)$$

Then from (4.2.5) it follows that,

$$\phi(\vec{r}) = \int_V d^3\vec{r}' \rho(\vec{r}') \mathcal{G}_D(\vec{r}, \vec{r}') - \frac{1}{4\pi} \int_S dS \phi(\vec{r}') \hat{n} \cdot \vec{\nabla}' \mathcal{G}_D(\vec{r}, \vec{r}') \quad (4.2.8)$$

and we are left with just the value of $\phi(\vec{r})$ on S , which is given by the boundary condition. This presumes that it is possible to find $\mathcal{Q}(\vec{r}, \vec{r}')$ such that (4.2.7) holds, which is always the case. The situation changes when Neumann boundary conditions are given. One would think that the best way to pick $\mathcal{Q}(\vec{r}, \vec{r}')$ would be so that $\partial\mathcal{G}/\partial n$ vanishes, but this is incompatible with the properties of \mathcal{G} , for

$$\int_S dS \hat{n} \cdot \vec{\nabla} \mathcal{G}(\vec{r}, \vec{r}') = \int_V d^3\vec{r} \vec{\nabla}^2 \mathcal{G}(\vec{r}, \vec{r}') = -4\pi \quad (4.2.9)$$

by Gauss' theorem, so the best we can really do is ask for

$$\hat{n} \cdot \vec{\nabla}' \mathcal{G}_N(\vec{r}, \vec{r}') \Big|_S = -\frac{4\pi}{S}, \quad (4.2.10)$$

where S is the area of the entire surface. Substitution into (4.2.5) then gives

$$\phi(\vec{r}) = \langle \phi \rangle_S + \int_V d^3\vec{r}' \rho(\vec{r}') \mathcal{G}_N(\vec{r}, \vec{r}') + \frac{1}{4\pi} \int_S dS \mathcal{G}_N(\vec{r}, \vec{r}') \hat{n} \cdot \vec{\nabla}' \phi(\vec{r}') \quad (4.2.11)$$

where $\langle \phi \rangle_S$ is the average value of the potential on the bounding surface.

Let us take a concrete example. Suppose we wanted to solve the problem that we solved earlier using the method of images, *i.e.*, the problem of a conducting sphere with (or without) a charge located somewhere outside it, using the Green function. Let us relax the condition that the sphere of radius R is a grounded conductor and simply specify the potential on it as an arbitrary function of the angular variables. The boundary condition is then Dirichlet, so we need to compute the Dirichlet Green function, $\mathcal{G}_D(\vec{r}, \vec{r}')$. But, from our work with the method of images we could guess that

$$\mathcal{Q}_D(\vec{r}, \vec{r}') = -\frac{R/|\vec{r}'|}{|\vec{r} - \frac{R^2}{|\vec{r}'|}\hat{r}'|}, \quad \mathcal{G}_D(\vec{r}, \vec{r}') = \frac{1}{|\vec{r} - \vec{r}'|} - \frac{R/|\vec{r}'|}{|\vec{r} - \frac{R^2}{|\vec{r}'|}\hat{r}'|} \quad (4.2.12)$$

for $\mathcal{Q}_D(\vec{r}, \vec{r}')$ obviously satisfies Laplace's equation and $\mathcal{G}_D(\vec{r}, \vec{r}')$ vanishes on S . Furthermore, when we use this in (4.2.8), with $\rho(\vec{r}') = q\delta(\vec{r}' - \vec{r}_0)$ corresponding to the presence of a physical charge and $\tilde{\phi}(R, \theta, \phi) = 0$, we find that the first integral gives exactly the electrostatic potential we obtained using the image charge. The surface integral in (4.2.8) vanishes because the potential vanishes on the sphere. However, the Green function is useful even when $\tilde{\phi}(R, \theta, \phi) \neq 0$.

Suppose γ is the angle between \vec{r} and \vec{r}' then \mathcal{G}_D can be written as

$$\mathcal{G}_D(\vec{r}, \vec{r}') = \frac{1}{\sqrt{r^2 + r'^2 - 2rr' \cos \gamma}} - \frac{R}{\sqrt{r^2 r'^2 + R^4 - 2rr'R^2 \cos \gamma}} \quad (4.2.13)$$

and, outside the sphere, gives

$$\frac{\partial \mathcal{G}_D}{\partial n} = \hat{n} \cdot \vec{\nabla}' \mathcal{G}(\vec{r}, \vec{r}') \Big|_S = -\frac{r^2 - R^2}{(r^2 + R^2 - 2rR \cos \gamma)^{3/2}} \quad (4.2.14)$$

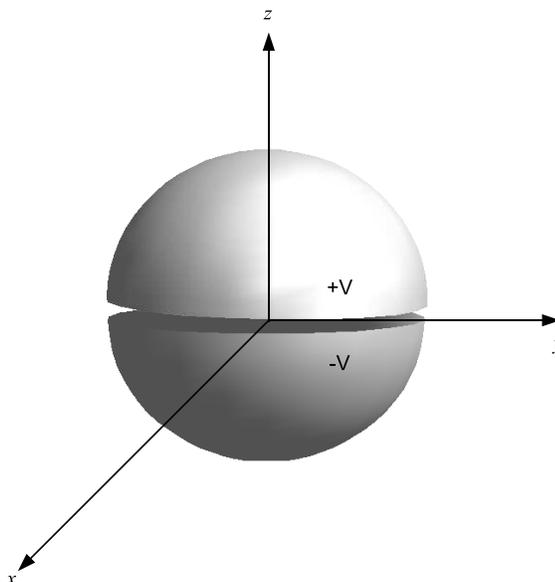


Figure 4.5: Two mutually insulated conducting hemispheres.

(the negative sign arises because the normal to the surface is inwards). Therefore, the potential *outside* a sphere with the potential specified on its surface but no other charge would be

$$\phi(\vec{r}) = \frac{1}{4\pi} \int d\Omega' \phi(R, \theta', \varphi') \frac{R^2(r^2 - R^2)}{(r^2 + R^2 - 2rR \cos \gamma)^{3/2}} \quad (4.2.15)$$

where $d\Omega'$ is the solid angle at the point (R, θ', φ') and $\cos \gamma = \hat{r} \cdot \hat{r}'$. For the interior problem, the normal derivative is radially outwards, so the sign of $\partial \mathcal{G}_D / \partial n$ is opposite to that given in (4.2.14). When a charge distribution is present, we must include the charge density term in (4.2.8) and we should get

$$\phi(\vec{r}) = \int d^3\vec{r}' \frac{\rho(\vec{r}')}{|\vec{r} - \vec{r}'|} + \frac{1}{4\pi} \int d\Omega' \phi(R, \theta', \varphi') \frac{R^2(r^2 - R^2)}{(r^2 + R^2 - 2rR \cos \gamma)^{3/2}} \quad (4.2.16)$$

The integrals above are not generally easy to do! Consider the simple problem of two conducting hemispheres making up a sphere but isolated from each other by an insulating ring as shown in figure (4.5). The solution to Laplace's equation everywhere outside the

sphere is given by

$$\phi(x) = -\frac{V}{4\pi} \int_0^{2\pi} d\varphi' \left[\int_0^{\pi/2} d\theta' \sin \theta' - \int_{-\pi/2}^0 d\theta' \sin \theta' \right] \frac{R^2(r^2 - R^2)}{(r^2 + R^2 - 2rR \cos \gamma)^{3/2}} \quad (4.2.17)$$

where $\cos \gamma = \cos \theta \cos \theta' + \sin \theta \sin \theta' \cos(\varphi - \varphi')$. As a special case, consider the integral on the z -axis. Then $\theta = 0$ and $\cos \gamma = \cos \theta'$. We easily find

$$\phi(z) = V \left[1 - \frac{z^2 - a^2}{z\sqrt{z^2 + a^2}} \right] \quad (4.2.18)$$

In general the integral cannot be performed in closed form. In this case one expands the denominator in power series and integrates term by term, which we will not trouble ourselves with doing here.⁴

4.3 Expansions of the Green function

In the previous chapter, we found it very useful to expand the electrostatic potential in complete, orthonormal sets of functions when the symmetry of a given problem is high. Expanding the Green function in a complete orthonormal set (see section 3.4) of functions can likewise be useful when the symmetry is such that one or more compact coordinate is present as in the case of spherical or cylindrical symmetry. We begin with the former.

4.3.1 Spherical Symmetry

If the symmetry of a given problem is spherical it is convenient to express the (spherical) Green function as an expansion in the complete, orthogonal set of spherical harmonics

$$Y_{lm}(\theta, \varphi) = P_{l,m}(\cos \theta) e^{im\varphi} \quad (4.3.1)$$

where the $P_{l,m}$ are the Legendre polynomials ($m \leq |l|$). Thus,

$$\mathcal{G}(\vec{r}, \vec{r}') = \sum_{l,m} A_{lm}(r|r', \theta', \varphi') Y_{lm}(\theta, \varphi) \quad (4.3.2)$$

(the coefficients A_{lm} depend only on the primed coordinates and r , as the dependence of $\mathcal{G}(\vec{r}, \vec{r}')$ on the unprimed angular coordinates occurs via the spherical harmonics) and satisfies

$$\vec{\nabla}^2 \mathcal{G}(\vec{r}, \vec{r}') = -4\pi \delta(\vec{r} - \vec{r}') \quad (4.3.3)$$

⁴Problem: Obtain the solution to (4.2.17) as a power series in $\cos \theta$ up to the third order and show that this is really nothing but the expansion in Legendre polynomials as we had before.

The δ -function may also be expanded in terms of the Legendre polynomials as

$$\delta(\vec{r} - \vec{r}') = \frac{1}{r^2} \delta(r - r') \sum_{l=0}^{\infty} \sum_{m=-l}^l Y_{lm}^*(\theta', \varphi') Y_{lm}(\theta, \varphi) \quad (4.3.4)$$

where we have used the completeness property,

$$\sum_{l=0}^{\infty} \sum_{m=-l}^l Y_{lm}^*(\theta', \varphi') Y_{lm}(\theta, \varphi) = \frac{1}{\sin \theta} \delta(\theta - \theta') \delta(\varphi - \varphi') \quad (4.3.5)$$

of the spherical harmonics. Substituting (4.3.2) into the Green function equation gives

$$A_{lm}(r|r', \theta', \varphi') = g_l(r|r') Y_{lm}^*(\theta', \varphi') \quad (4.3.6)$$

where $g_l(r|r')$ satisfies

$$\frac{1}{r} \partial_r^2 (r g_l) - \frac{l(l+1)}{r^2} g_l = -\frac{4\pi}{r^2} \delta(r - r') \quad (4.3.7)$$

The left hand side is the same as that of the radial equation in (3.5.59), but here we have a δ -function source term. Nevertheless, the δ -function is supported only on the spherical shells, when $r = r'$, so we may borrow the solutions we had before for $r > r'$ and $r < r'$. We will have

$$g_l(r|r') = \begin{cases} A_l r^l + B_l r^{-(l+1)}, & r < r' \\ A'_l r^l + B'_l r^{-(l+1)}, & r > r' \end{cases} \quad (4.3.8)$$

where A, B and A', B' depend on r' . We are interested in the spherical Green function only when the symmetry of the problem is spherical, so imagine that we have boundary conditions on spheres, say of radii a and b , and let $a < b$. Then, with Dirichlet boundary conditions such that $\mathcal{G}(\vec{r}, \vec{r}')$ vanishes at $r = a$ and $r = b$, we see that

$$\begin{aligned} A_l a^l + B_l a^{-(l+1)} = 0 & \Rightarrow B_l = -A_l a^{2l+1} \\ A'_l b^l + B'_l b^{-(l+1)} = 0 & \Rightarrow A'_l = -B'_l b^{-(2l+1)} \end{aligned} \quad (4.3.9)$$

giving

$$g_l(r|r') = \begin{cases} A_l \left(r^l + \frac{a^{2l+1}}{r^{l+1}} \right), & r < r' \\ B'_l \left(\frac{1}{r^{l+1}} - \frac{r'^l}{b^{2l+1}} \right), & r > r' \end{cases} \quad (4.3.10)$$

The symmetry of $g(r|r')$ implies that we should be able to interchange r and r' with no effect, *i.e.* for example,

$$A_l(r') \left(r^l + \frac{a^{2l+1}}{r^{l+1}} \right) = B_l(r) \left(\frac{1}{r^{l+1}} - \frac{r'^l}{b^{2l+1}} \right) \quad (4.3.11)$$

which is possible only if

$$\begin{aligned} A_l(r') &= C \left(\frac{1}{r'^{l+1}} - \frac{r'^l}{b^{2l+1}} \right) \\ B_l(r') &= C \left(r'^l - \frac{a^{2l+1}}{r'^{l+1}} \right), \end{aligned} \quad (4.3.12)$$

where C is constant. Therefore

$$g_l(r|r') = \begin{cases} C \left(\frac{1}{r'^{l+1}} - \frac{r'^l}{b^{2l+1}} \right) \left(r^l - \frac{a^{2l+1}}{r^{l+1}} \right), & r < r' \\ C \left(r'^l - \frac{a^{2l+1}}{r'^{l+1}} \right) \left(\frac{1}{r^{l+1}} - \frac{r^l}{b^{2l+1}} \right), & r > r' \end{cases} \quad (4.3.13)$$

and we can make this look nicer if we define

$$\begin{aligned} r_{>} &\equiv \max\{r, r'\} \\ r_{<} &\equiv \min\{r, r'\} \end{aligned} \quad (4.3.14)$$

in terms of which

$$g(r|r') = C \left(\frac{1}{r_{>}^{l+1}} - \frac{r_{>}^l}{b^{2l+1}} \right) \left(r_{<}^l - \frac{a^{2l+1}}{r_{<}^{l+1}} \right). \quad (4.3.15)$$

It's easy to see that the function $g(r|r')$ is continuous as $r \rightarrow r'$, but it is not differentiable. This is as it should be, of course, because there is a δ -function source on the right. We will now use this fact to obtain the value of C . Returning to the radial equation for $g(r|r')$ in (4.3.7), multiplying it by r and integrating over r from $r = r' - \epsilon$ to $r = r' + \epsilon$ taking the limit as ϵ vanishes, we find

$$[\partial_r(rg_l)]_{r+\epsilon} - [\partial_r(rg_l)]_{r-\epsilon} = -\frac{4\pi}{r'} \quad (4.3.16)$$

But, using the expression (4.3.15) for $g_l(r|r')$, the left hand side is easily calculated and compared to the right. We find that they are equal only if⁵

$$C = \frac{4\pi}{(2l+1) \left[1 - \left(\frac{a}{b} \right)^{2l+1} \right]} \quad (4.3.17)$$

We finally have the general spherical Green function with two spherical boundaries of radii $a < b$, on which it vanishes,

$$\mathcal{G}_D(\vec{r}, \vec{r}') = 4\pi \sum_{l=0}^{\infty} \sum_{m=-l}^l \left[r_{<}^l - \frac{a^{2l+1}}{r_{<}^{l+1}} \right] \left[\frac{1}{r_{>}^{l+1}} - \frac{r_{>}^l}{b^{2l+1}} \right] \frac{Y_{lm}^*(\theta', \varphi') Y_{lm}(\theta, \varphi)}{(2l+1) \left[1 - \left(\frac{a}{b} \right)^{2l+1} \right]} \quad (4.3.18)$$

⁵Problem: Show this result by explicit calculation.

We now consider the special case in which the outer surface approaches infinity ($b \rightarrow \infty$). In this case there are just two regions, the exterior of a and the interior. Consider first the exterior. In this case, we easily find

$$\mathcal{G}_D(\vec{r}, \vec{r}') = 4\pi \sum_{l=0}^{\infty} \sum_{m=-l}^l \left[\frac{r_{<}^l}{r_{>}^{l+1}} - \frac{1}{a} \left(\frac{a^2}{rr'} \right)^{l+1} \right] \frac{Y_{lm}^*(\theta'\varphi')Y_{lm}(\theta, \varphi)}{(2l+1)} \quad (4.3.19)$$

which clearly vanishes when either r or r' coincides with a as well as when either of these approaches infinity. If $a = 0$ then

$$\mathcal{G}_D(\vec{r}, \vec{r}') = 4\pi \sum_{l=0}^{\infty} \sum_{m=-l}^l \frac{r_{<}^l}{r_{>}^{l+1}} \frac{Y_{lm}^*(\theta'\varphi')Y_{lm}(\theta, \varphi)}{(2l+1)} = \frac{1}{|\vec{r} - \vec{r}'|} \quad (4.3.20)$$

which is proportional to the potential at \vec{r} due to a point charge at \vec{r}' . The left hand side is just the expansion of the right hand side in spherical harmonics.

4.3.2 Cylindrical Symmetry

From our experience with Laplace's equation, let's begin with the following ansatz for the Green function, if the symmetry of a given problem is cylindrical:

$$\mathcal{G}(\vec{r}, \vec{r}') = \frac{1}{4\pi^2} \sum_{m=-\infty}^{\infty} \int_{-\infty}^{\infty} dk g_m(k; \rho|\rho', z', \varphi') e^{ikz} e^{im\varphi}. \quad (4.3.21)$$

The delta function in cylindrical coordinates is

$$\begin{aligned} \delta(\vec{r} - \vec{r}') &= \frac{1}{\rho} \delta(\rho - \rho') \delta(z - z') \delta(\varphi - \varphi') \\ &= \frac{1}{\rho} \delta(\rho - \rho') \frac{1}{4\pi^2} \sum_{m=-\infty}^{\infty} \int_{-\infty}^{\infty} dk e^{ik(z-z')} e^{im(\varphi-\varphi')}, \end{aligned} \quad (4.3.22)$$

the last expression being just a statement of the completeness relations

$$\begin{aligned} \delta(z - z') &= \frac{1}{2\pi} \int_{-\infty}^{\infty} dk e^{ik(z-z')} \\ \delta(\varphi - \varphi') &= \frac{1}{2\pi} \sum_{m=-\infty}^{\infty} e^{im(\varphi-\varphi')} \end{aligned} \quad (4.3.23)$$

Inserting (4.3.21) into the Green function equation then gives

$$g_m(k; \rho|\rho', z', \varphi') = g_m(k; \rho|\rho') e^{-ikz} e^{-im\varphi} \quad (4.3.24)$$

and

$$\frac{1}{\rho} \partial_\rho (\rho \partial_\rho g_m) - \left(k^2 + \frac{m^2}{\rho^2} \right) g_m = -\frac{4\pi}{\rho} \delta(\rho - \rho') \quad (4.3.25)$$

As before, when $\rho > \rho'$ or $\rho < \rho'$ we may borrow the solutions we had earlier in (3.5.27). So write the solutions for $g_m(k; \rho|\rho')$ as

$$g_m(k; \rho|\rho') = \begin{cases} A_m I_m(k\rho) + B_m K_m(k\rho) & \rho < \rho' \\ A'_m I_m(k\rho) + B'_m K_m(k\rho) & \rho > \rho' \end{cases} \quad (4.3.26)$$

where A, B and A', B' are functions of ρ' . Suppose that the Green function vanishes on cylindrical surfaces of radii a and b (for definiteness, assume that $a < b$) then

$$\begin{aligned} B_m &= -A_m \frac{I_m(ka)}{K_m(ka)} \\ A'_m &= -B'_m \frac{K_m(kb)}{I_m(kb)} \end{aligned} \quad (4.3.27)$$

giving

$$g_m(k; \rho|\rho') = \begin{cases} A_m \left(I_m(k\rho) - \frac{I_m(ka)K_m(k\rho)}{K_m(ka)} \right) & \rho < \rho' \\ B'_m \left(K_m(k\rho) - \frac{K_m(kb)I_m(k\rho)}{I_m(kb)} \right) & \rho > \rho' \end{cases} \quad (4.3.28)$$

But, by the symmetry of the Green function under an interchange of ρ and ρ' ,

$$A_m(\rho') = C \left(K_m(k\rho') - \frac{K_m(kb)I_m(k\rho')}{I_m(kb)} \right) \quad (4.3.29)$$

and

$$B'_m(\rho') = C \left(I_m(k\rho') - \frac{I_m(ka)K_m(k\rho')}{K_m(ka)} \right) \quad (4.3.30)$$

This gives

$$g_m(k; \rho|\rho') = \begin{cases} C \left(K_m(k\rho') - \frac{K_m(kb)I_m(k\rho')}{I_m(kb)} \right) \left(I_m(k\rho) - \frac{I_m(ka)K_m(k\rho)}{K_m(ka)} \right) & \rho < \rho' \\ C \left(I_m(k\rho') - \frac{I_m(ka)K_m(k\rho')}{K_m(ka)} \right) \left(K_m(k\rho) - \frac{K_m(kb)I_m(k\rho)}{I_m(kb)} \right) & \rho > \rho' \end{cases} \quad (4.3.31)$$

where C is some constant that must be fixed using by the discontinuity in the slope implied by the δ -function in the Green function equation.

Consider the greatly simplified, special case when the inner cylinder is shrunk to zero radius ($a = 0$) and the outer cylinder is taken out to infinity ($b = \infty$). Using the asymptotic

behaviors of the modified Bessel functions, it is easy to see that the solution above reduces in this case to

$$g_m(k; \rho|\rho') = \begin{cases} CK_m(k\rho')I_m(k\rho) & \rho < \rho' \\ CI_m(k\rho')K_m(k\rho) & \rho > \rho' \end{cases} \quad (4.3.32)$$

or, simply,

$$g_m(k; \rho|\rho') = CI_m(k\rho_<)K_m(k\rho_>) \quad (4.3.33)$$

where we have defined

$$\begin{aligned} \rho_> &\equiv \max\{\rho, \rho'\} \\ \rho_< &\equiv \min\{\rho, \rho'\} \end{aligned} \quad (4.3.34)$$

A simple exercise, similar to the one we performed when determining the constant in the case of spherical symmetry, then shows that $C = 4\pi$, thus

$$\mathcal{G}(\vec{r}, \vec{r}') = \frac{2}{\pi} \sum_{m=-\infty}^{\infty} \int_0^{\infty} dk I_m(k\rho_<)K_m(k\rho_>) \cos k(z - z') e^{im(\varphi - \varphi')} \quad (4.3.35)$$

which may also be written in terms of real functions.⁶

4.4 Eigenfunction expansion of the Green function

An approach to obtaining expansions of the Green function that is intimately connected with the methods used in the previous section is the so called eigenfunction expansion.

If a function, $\psi(\vec{r})$ satisfies some elliptic differential equation of the form

$$\hat{O}\psi(\vec{r}) = \vec{\nabla}^2\psi(\vec{r}) + V(\vec{r})\psi(\vec{r}) = \lambda\psi(\vec{r}) \quad (4.4.1)$$

where λ is a constant and if $\psi(\vec{r})$ satisfies homogeneous boundary conditions on some bounding surface of a volume, V , of interest then $\psi(\vec{r})$ is said to be an “eigenfunction” of

⁶Problem: Use the properties of the modified Bessel functions to write the Green function for the special case $a = 0$, $b = \infty$ as

$$\mathcal{G}(\vec{r}, \vec{r}') = \frac{4}{\pi} \int_0^{\infty} dk \left[\frac{1}{2} I_0(k\rho_<)K_0(k\rho_>) + \sum_{m=1}^{\infty} I_m(k\rho_<)K_m(k\rho_>) \cos m(\varphi - \varphi') \right] \cos k(z - z')$$

This is therefore the expansion of $1/|\vec{r} - \vec{r}'|$ in cylindrical coordinates. Using this identity and taking $\rho' = 0$, show that

$$\frac{1}{\sqrt{\rho^2 + z^2}} = \frac{2}{\pi} \int_0^{\infty} dk K_0(k\rho) \cos kz$$

the (elliptic) differential operator $\hat{O} = \vec{\nabla}^2 + V(\vec{r})$ and λ is an “eigenvalue” of the operator. Generally, the equation (4.4.1) will admit well behaved solutions for a set of values of λ , giving a set of eigenvalues and eigenfunctions of the differential operator in question. If we label the members of this set by α (α may be a continuous variable, corresponding to a continuous eigenvalue spectrum but, depending on the boundary conditions, it may also be discrete, corresponding to a discrete spectrum) then we may write

$$\hat{O}\psi_\alpha(\vec{r}) = \vec{\nabla}^2\psi_\alpha(\vec{r}) + V(\vec{r})\psi_\alpha(\vec{r}) = \lambda_\alpha\psi_\alpha(\vec{r}) \quad (4.4.2)$$

The eigenfunctions are orthogonal with respect to the inner product

$$\langle\phi, \psi\rangle = \int_V d^3\vec{r}\phi^*(\vec{r})\psi(\vec{r}). \quad (4.4.3)$$

To show this, note that $\hat{O} = \vec{\nabla}^2 + V(\vec{r})$ is a hermitean operator (as long as $V(r)$ is real) and consider two eigenfunctions, ϕ and ψ of \hat{O} with eigenvalues λ_1 and λ_2 respectively. By definition, Hermiticity of \hat{O} means that

$$\langle\hat{O}\phi, \psi\rangle = \langle\phi, \hat{O}\psi\rangle \quad (4.4.4)$$

from which it follows that λ_1 and λ_2 are real. Combined, these imply that

$$0 = \langle\hat{O}\phi, \psi\rangle - \langle\phi, \hat{O}\psi\rangle = (\lambda_1 - \lambda_2)\langle\phi, \psi\rangle, \quad (4.4.5)$$

i.e., if $\lambda_1 \neq \lambda_2$ then $\langle\phi, \psi\rangle = 0$. If $\lambda_1 = \lambda_2$ then the functions ϕ and ψ may differ by at most a constant. We can rid ourselves of this superfluous constant by normalizing the eigenfunctions to unity, *i.e.*, by setting

$$\langle\psi_\alpha, \psi_\beta\rangle = \delta_{\alpha\beta} \quad (4.4.6)$$

If, now, we want to find the Green function for the operator $\hat{O} - \lambda$, *i.e.*, satisfying the equation

$$[\hat{O} - \lambda]\mathcal{G}(\vec{r}, \vec{r}') = -4\pi\delta(\vec{r} - \vec{r}') \quad (4.4.7)$$

where λ is *not* an eigenvalue of \hat{O} , and obeying the same boundary conditions as do the eigenfunctions of \hat{O} , then consider the following expansion of $\mathcal{G}(\vec{r}, \vec{r}')$ in terms of the eigenfunctions of \hat{O}

$$\mathcal{G}(\vec{r}, \vec{r}') = \sum_\alpha a_\alpha(\vec{r}')\psi_\alpha(\vec{r}) \quad (4.4.8)$$

where the sum turns into an integral if the eigenvalue spectrum is continuous. Substituting this expansion into the Green function equation we find

$$\sum_\alpha a_\alpha(\vec{r}')(\lambda_\alpha - \lambda)\psi_\alpha(\vec{r}) = -4\pi\delta(\vec{r} - \vec{r}') \quad (4.4.9)$$

Multiplying both sides by $\psi_\beta^*(\vec{r}')$ and integrating over the volume we find

$$\sum_{\alpha} a_{\alpha}(\vec{r}')(\lambda_{\alpha} - \lambda)\delta_{\alpha\beta} = -4\pi\psi_{\beta}^*(\vec{r}') \quad (4.4.10)$$

or

$$a_{\alpha}(\vec{r}') = -4\pi\frac{\psi_{\alpha}^*(\vec{r}')}{\lambda_{\alpha} - \lambda} \quad (4.4.11)$$

Thus,

$$\mathcal{G}(\vec{r}, \vec{r}') = -4\pi\sum_{\alpha}\frac{\psi_{\alpha}^*(\vec{r}')\psi(\vec{r})}{\lambda_{\alpha} - \lambda} \quad (4.4.12)$$

As a simple example, we take the Poisson equation

$$[\vec{\nabla}^2 + k^2]\psi = 0 \quad (4.4.13)$$

($k \neq 0$) whose solutions are plane waves

$$\psi_{\vec{k}}(\vec{r}) = \frac{1}{(2\pi)^{3/2}}e^{i\vec{k}\cdot\vec{r}} \quad (4.4.14)$$

where \vec{k} is the wave vector satisfying $\vec{k} \cdot \vec{k} = k^2$. These eigenfunctions form a complete, orthonormal set, *i.e.*,

$$\begin{aligned} \int d^3\vec{k}\psi_{\vec{k}}^*(\vec{r})\psi_{\vec{k}'}(\vec{r}') &= \delta(\vec{r} - \vec{r}') \quad (\text{completeness}) \\ \int d^3\vec{r}\psi_{\vec{k}}^*(\vec{r})\psi_{\vec{k}'}(\vec{r}) &= \delta(\vec{k} - \vec{k}') \quad (\text{orthonormality}) \end{aligned} \quad (4.4.15)$$

If we want to represent the Green function satisfying

$$\vec{\nabla}^2\mathcal{G}(\vec{r}, \vec{r}') = -4\pi\delta(\vec{r} - \vec{r}') \quad (4.4.16)$$

with $\lambda = 0$ in (4.4.9) we see that it has the form

$$\mathcal{G}(\vec{r}, \vec{r}') = -\frac{1}{2\pi^2}\int d^3\vec{k}\frac{e^{i\vec{k}\cdot(\vec{r}-\vec{r}')}}{k^2} \quad (4.4.17)$$

which will be shown later to be nothing but the three dimensional Fourier representation of $1/|\vec{r} - \vec{r}'|$. We have assumed no bounding surfaces, in selecting the complete orthonormal set in (4.4.14). In a Dirichlet problem care must be taken to choose eigenfunctions that obey the given conditions on the bounding surfaces.

Chapter 5

Dielectric Media

So far we have learned two laws that govern the electric field, *viz.*, Gauss' law

$$\vec{\nabla} \cdot \vec{E} = \frac{\rho}{\epsilon_0} \quad (5.0.1)$$

which states that charge is the source of electric flux, and

$$\vec{\nabla} \times \vec{E} = 0 \quad (5.0.2)$$

which expresses the fact that the electrostatic field is conservative. The last law is valid only when the physical system is *static* and must be modified in the presence of moving charges or boundaries, or when any other form of time dependence is present. This involves the magnetic field, which we will presently introduce and instead of a conservation law we will have “Faraday’s law” of induction. In this chapter, however, we will continue to be interested in static solutions, so the electric field will continue to be conservative.

The first law is valid whether or not there is time dependence in the system. Although the law itself does not discriminate between them, it is convenient to distinguish between two kinds of charges or charge densities that may be present in any given system. The first kind we will call the “fixed” or “bound” charge. This will include all the charges that are held together by some external means (the external means include binding to atoms or to lattice sites). The second kind we will call the “free” charge. As its name predicts, this charge is free to move about within system, its motion being determined by all the fields present. Such a distinction is particularly useful when we study electric fields in materials.

5.1 Polarization

We have learned that the electric field inside a conductor is identically zero in equilibrium because all the charges are free to rearrange themselves and will do so in such a way that

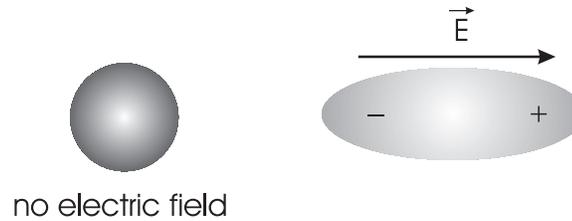


Figure 5.1: Distortion of molecules in an externally applied electric field.

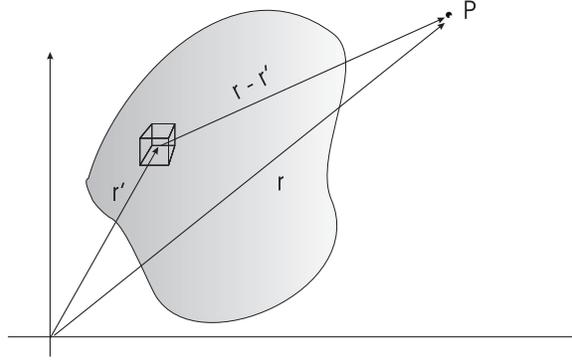
the net force on any charge is zero. When this occurs the conductor is in equilibrium, and it is only possible when the net electric field vanishes. Dielectric media do not have free charges. All charges are bound rather strongly to the atoms (molecules) of the medium. When an electric field is applied to such a medium the net effect is that the molecules get distorted, the positive charges in the nucleus being pulled in one way by the external electric field and the negative charges being pulled the other way as shown in figure (5.6). Each molecule then acts as a little electric dipole, and all the distorted molecules together will generate an electric field that will be macroscopically appreciable. The strength of each dipole, *i.e.*, its dipole moment, will depend on the separation between the charges, which will in turn depend on the strength of the externally applied field. Therefore the electric field generated by the dipoles will depend on the applied field. However, the dipole electric field and the externally applied field combine and it is really the resultant field that creates the dipoles themselves. This feedback means that the system is quite non-linear. The process by which the dielectric medium acquires an electric dipole moment is called **polarization**. The application of an external electric field to a dielectric medium electrically polarizes it.

Consider a volume element Δv of the medium. We will take this volume element to be macroscopically small (even taking the limit as $\Delta v \rightarrow 0$ in the end) but always assume that it is *microscopically* large in the sense that it contains a very large number of molecules. This is the so-called “continuum” approximation, in which the molecules are themselves taken to be point-like so that *any* volume contains a very large (essentially infinite) number of them. Let \vec{p}_m represent the electric dipole moment of the m^{th} dipole (molecule) in the volume Δv , then the total dipole moment of the volume is

$$\Delta\vec{p} = \sum_{m \in \Delta v} \vec{p}_m \quad (5.1.1)$$

We define the electric polarization, \vec{P} , of the medium by

$$\vec{P} = \lim_{\Delta v \rightarrow 0} \frac{\Delta\vec{p}}{\Delta v} \quad (5.1.2)$$

Figure 5.2: Dielectric field at a point P , external to the dielectric

The vector \vec{P} is in general a function of location within the medium, $\vec{P} = \vec{P}(\vec{r})$, and is a density. We will refer to it simply as the “polarization” of the medium.

5.2 Electric field outside a Dielectric

First let’s consider a polarized medium, characterized by a polarization vector $\vec{P}(\vec{r}')$, and the electric field due to its polarization at a point P external to the medium as shown in figure (5.2). We already know from (2.3.12) that the electric potential at P due to an infinitesimal volume of the dielectric located at \vec{r}' within it would be

$$d\phi(\vec{r}) = \frac{\Delta \vec{p}(\vec{r}') \cdot (\vec{r} - \vec{r}')}{4\pi\epsilon_0 |\vec{r} - \vec{r}'|^3} = d^3\vec{r}' \frac{\vec{P}(\vec{r}') \cdot (\vec{r} - \vec{r}')}{4\pi\epsilon_0 |\vec{r} - \vec{r}'|^3} \quad (5.2.1)$$

This can be integrated over the entire dielectric to give the electric field at P ,

$$\phi(\vec{r}) = \int_{V_D} d^3\vec{r}' \frac{\vec{P}(\vec{r}') \cdot (\vec{r} - \vec{r}')}{4\pi\epsilon_0 |\vec{r} - \vec{r}'|^3} \quad (5.2.2)$$

where V_D is the dielectric volume. Thus, if we knew $\vec{P}(\vec{r})$ we could calculate $\phi(\vec{r})$ in this simple example with no non-trivial boundaries, but what does the solution really mean? We get a better physical understanding of the solution as follows: use the fact that

$$\vec{\nabla}' \frac{1}{|\vec{r} - \vec{r}'|} = \frac{\vec{r} - \vec{r}'}{|\vec{r} - \vec{r}'|^3}$$

($\vec{\nabla}'$ acts only on \vec{r}') to express

$$\frac{\vec{P}(\vec{r}') \cdot (\vec{r} - \vec{r}')}{|\vec{r} - \vec{r}'|^3} = \vec{P}(\vec{r}') \cdot \vec{\nabla}' \frac{1}{|\vec{r} - \vec{r}'|} = \vec{\nabla}' \cdot \left(\frac{\vec{P}(\vec{r}')}{|\vec{r} - \vec{r}'|} \right) - \frac{\vec{\nabla}' \cdot \vec{P}(\vec{r}')}{|\vec{r} - \vec{r}'|} \quad (5.2.3)$$

Inserting the right hand side above into the expression for the electrostatic potential gives

$$\phi(\vec{r}) = \frac{1}{4\pi\epsilon_0} \int_{V_D} d^3\vec{r}' \left(\vec{\nabla}' \cdot \left(\frac{\vec{P}(\vec{r}')}{|\vec{r}' - \vec{r}|} \right) - \frac{\vec{\nabla}' \cdot \vec{P}(\vec{r}')}{|\vec{r}' - \vec{r}|} \right) \quad (5.2.4)$$

Now using Gauss' theorem, express the first integral above as an integral over the surface of the medium,

$$\phi(\vec{r}) = \frac{1}{4\pi\epsilon_0} \int_{S_D} dS \frac{\hat{n} \cdot \vec{P}(\vec{r}')}{|\vec{r}' - \vec{r}|} - \frac{1}{4\pi\epsilon_0} \int_{V_D} d^3\vec{r}' \frac{\vec{\nabla}' \cdot \vec{P}(\vec{r}')}{|\vec{r}' - \vec{r}|} \quad (5.2.5)$$

Now it is easy to see that the first integral corresponds to a contribution from a surface charge density

$$\sigma_P(\vec{r}') = \hat{n} \cdot \vec{P}(\vec{r}') \quad (5.2.6)$$

and the second from a *volume* charge density

$$\rho_P(\vec{r}') = -\vec{\nabla}' \cdot \vec{P}(\vec{r}') \quad (5.2.7)$$

These are, respectively, the polarization surface charge density and polarization volume charge density. As far as the electric field in the exterior of the dielectric is concerned, we may think of the dielectric medium as made up of two kinds of distributions: (a) a volume charge distribution with density $\rho_P(\vec{r}')$ and (b) a surface distribution with density $\sigma_P(\vec{r}')$.

5.3 Electric field inside a Dielectric

Let us use the result above to determine the electric field inside the dielectric. The electric field at a point, say \vec{r} inside the dielectric will be due to all the dipoles surrounding it. Consider then an infinitesimally thin needle shaped cavity inside the dielectric as shown in figure (5.3) and apply the conservation equation for the electric field (which is valid as long as the situation is static) in its integral form,

$$\oint_C \vec{E} \cdot d\vec{r} = 0, \quad (5.3.1)$$

to the rectangular curve shown in the figure. Evaluating the integral and assuming that $\Delta\vec{l}$ is small, so that \vec{E} is more or less constant along it, we find

$$\vec{E}_d \cdot \Delta\vec{l}_d + \vec{E}_{\text{vac}} \cdot \Delta\vec{l}_{\text{vac}} = 0 \quad (5.3.2)$$

where we have neglected the contribution from the edges because the needle shaped cavity can be made as thin as we wish, and the contribution from the edges will be negligible

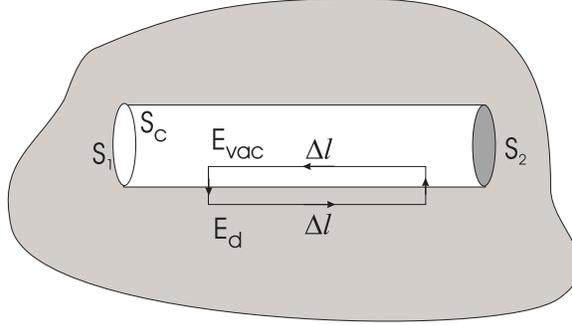


Figure 5.3: Needle shaped cavity inside a Dielectric

in the limit as the width goes to zero. Now $\Delta \vec{l}_d$ and $\Delta \vec{l}_{vac}$ are equal in magnitude and oppositely directed, therefore we have

$$E_{d,t} = E_{vac,t} \quad (5.3.3)$$

where the suffix t represents the component of \vec{E} that is tangent to the needle shaped cavity. This result is valid no matter how the cavity is oriented w.r.t. the electric field inside the dielectric. In particular, if it is oriented parallel to the electric field in the dielectric we would have, simply,

$$E_d = E_{vac} \quad (5.3.4)$$

We conclude that the electric field in a dielectric is identical to the electric field in a cylindrical (needle shaped) cavity whose axis is in the direction of the field. If V_1 is the volume of the cavity and S_c is the cylindrical surface area (we can neglect the contributions from S_1 and S_2 as the cavity is made thinner and thinner), then

$$\phi_d(\vec{r}) = \frac{1}{4\pi\epsilon_0} \int_{S_c + S_{ext}} dS \frac{\sigma_P(\vec{r}')}{|\vec{r} - \vec{r}'|} + \frac{1}{4\pi\epsilon_0} \int_{V_D - V_1} d^3\vec{r}' \frac{\rho_P(\vec{r}')}{|\vec{r} - \vec{r}'|} \quad (5.3.5)$$

where S_{ext} is the exterior surface of the dielectric (which should also be included). Now if the orientation of our cavity were chosen to lie parallel to the electric field in the dielectric, the surface charge density on S_c would be vanishing. Taking $V_1 \rightarrow 0$ we get

$$\phi_d(\vec{r}) = \frac{1}{4\pi\epsilon_0} \int_{S_{ext}} dS \frac{\sigma_P(\vec{r}')}{|\vec{r} - \vec{r}'|} + \frac{1}{4\pi\epsilon_0} \int_{V_D} d^3\vec{r}' \frac{\rho_P(\vec{r}')}{|\vec{r} - \vec{r}'|} \quad (5.3.6)$$

which, not surprisingly, is the same as we had before for points outside of the dielectric.

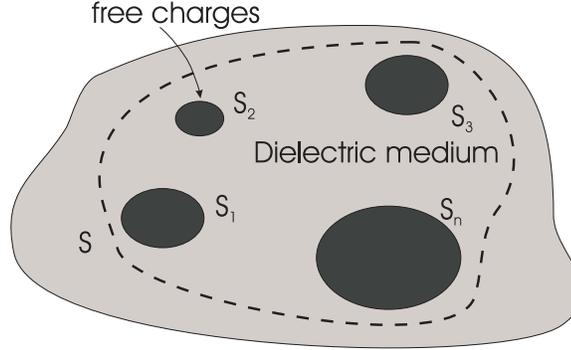


Figure 5.4: Gauss' Law inside a Dielectric

5.4 Gauss' Law inside Dielectrics

A more general situation is obtained by imagining that we introduce some free charge into the dielectric and that this charge moves freely within certain regions that we can approximate as conductors. Consider a Gaussian surface inside the dielectric and let the surface enclose a certain number of these regions (or conductors), say $S_1, S_2 \dots S_n$. Applying Gauss' law to the surface

$$\oint_S \vec{E} \cdot d\vec{S} = \frac{q_{in}}{\epsilon_0} = \frac{q_f}{\epsilon_0} + \frac{q_p}{\epsilon_0} \quad (5.4.1)$$

where q_p refers to the polarization charge,

$$q_p = \int_{V-V_1-V_2 \dots -V_n} d^3\vec{r}' [-\vec{\nabla}' \cdot \vec{P}(\vec{r}')] + \int_{S_1+S_2 \dots +S_n} d^2\vec{r}' [\hat{n} \cdot \vec{P}(\vec{r}')] \quad (5.4.2)$$

where we have used the polarization charge densities obtained earlier and $V - V_1 \dots - V_n$ is the volume of the dielectric enclosed by the Gaussian surface and $S_1 \dots + S_n$ is the sum of all the conducting surfaces within this volume. Notice that we cannot include the Gaussian surface itself because there is no surface charge density there. The volume integral can be transformed into a surface integral using Gauss' theorem

$$\int_{V-V_1 \dots -V_n} d^3\vec{r}' [-\vec{\nabla}' \cdot \vec{P}(\vec{r}')] = - \int_{S+S_1 \dots +S_n} d^2\vec{r}' \hat{n} \cdot \vec{P}(\vec{r}') \quad (5.4.3)$$

where, now, all surfaces bounding the volume V are included. The net contribution is therefore

$$q_p = \int_S d^2\vec{r}' \hat{n} \cdot \vec{P}(\vec{r}'), \quad (5.4.4)$$

due only to the contribution from the distribution of charge on the Gaussian surface. We find

$$\oint_S \left[\vec{E} + \frac{\vec{P}}{\epsilon_0} \right] \cdot d\vec{S} = \frac{q_f}{\epsilon_0} \quad (5.4.5)$$

which has the form of (a modified) Gauss' law with only the free charge as a source. Define the electric displacement by $\vec{D} = \epsilon_0 \vec{E} + \vec{P}$ then Gauss' law in a dielectric medium will take the form

$$\oint_S \vec{D} \cdot d\vec{S} = q_f \quad (5.4.6)$$

Obviously we can put this in differential form by reexpressing the total free charge in terms of the free charge density and using Gauss' theorem

$$\begin{aligned} \oint_S \vec{D} \cdot d\vec{S} &= \int_V \vec{\nabla} \cdot \vec{D} \, d^3\vec{r} = q_f = \int_V \rho_f(\vec{r}) \, d^3\vec{r} \\ \Rightarrow \vec{\nabla} \cdot \vec{D} &= \rho_f \end{aligned} \quad (5.4.7)$$

which is the form of Gauss' law in a dielectric medium. In the absence of polarization, $\vec{D} = \epsilon_0 \vec{E}$. In a dielectric medium, the electric field is made up of two parts, the first due to the free charges and the second due to the electric polarization, \vec{P} , of the medium. The electric field continues conserved

$$\vec{\nabla} \times \vec{E} = 0, \quad (5.4.8)$$

but \vec{D} is not necessarily conservative ($\vec{\nabla} \times \vec{D} = \vec{\nabla} \times \vec{P} \neq 0$) and therefore \vec{D} is not derivable from a potential. Equations (5.4.7) and (5.4.8) constitute the analogues of the vacuum electrostatic equations in (2.3.18) and (2.3.1) respectively.

We cannot really proceed with problem solving without having some relation between \vec{D} and \vec{E} . This is equivalent to knowing the the electric polarization of the medium. The polarization occurs via the the application of the external electric field, so generally speaking $\vec{P} = \vec{P}(\vec{E})$. However, while theoretical models may be able to determine $\vec{P}(\vec{E})$ for certain materials and under certain conditions, in general $\vec{P}(\vec{E})$ must be obtained experimentally. Phenomenologically, one could write an expression of the form

$$P_i(\vec{E}) = \sum_j \chi_{ij}(\vec{E}) E_j \quad (5.4.9)$$

where $\chi_{ij}(\vec{E})$ are the components of the so-called electric susceptibility tensor. If the material is *isotropic* then the susceptibility tensor is determined by just one function $\chi(\vec{E})$, and

$$\chi_{ij}(\vec{E}) = \chi(\vec{E}) \delta_{ij} \Rightarrow \vec{P}(\vec{E}) = \chi(\vec{E}) \vec{E} \quad (5.4.10)$$

implying that \vec{P} is in the same direction as \vec{E} . $\chi(\vec{E})$ is simply called the *electric susceptibility* of the medium. With this function, we can write

$$\vec{D} = [\epsilon_0 + \chi(\vec{E})]\vec{E} = \epsilon(\vec{E})\vec{E} \quad (5.4.11)$$

where $\epsilon(\vec{E}) = \epsilon_0 + \chi(\vec{E})$ is the *electric permittivity* of the dielectric medium (ϵ_0 may be thought of as the electric permittivity of the vacuum). A dielectric medium is said to be *linear* if the electric susceptibility (equivalently the electric permittivity) is constant. In that case

$$\vec{P} = \chi\vec{E}, \quad \vec{D} = \epsilon\vec{E} \quad (5.4.12)$$

where ϵ is constant. Calling $\epsilon = K\epsilon_0$ we find, in terms of the electric susceptibility that

$$K = \left[1 + \frac{\chi}{\epsilon_0} \right]. \quad (5.4.13)$$

K is called the *dielectric constant*. The dielectric constant of the vacuum is unity and it is infinite for a conductor. For a linear dielectric medium, therefore, one can combine (5.4.7) and (5.4.8) to obtain Poisson's equation exactly as we did before. The fact that \vec{E} is conservative implies the existence of a corresponding electric potential

$$\epsilon\vec{E}(\vec{r}) = -\vec{\nabla}\phi_D(\vec{r}) = \vec{D} \quad (5.4.14)$$

Gauss' law then provides the equation for $\phi(\vec{r})$,

$$\vec{\nabla} \cdot \vec{D} = -\vec{\nabla}^2\phi_D = \rho_f \quad (5.4.15)$$

(note that, as $\vec{D} = \epsilon\vec{E}$ the electric potential ϕ is related to the displacement potential ϕ_D by $\phi_D = \epsilon\phi$). In the absence of charges both potentials obey Laplace's equations.

It should be evident that methods for solving all problems involving linear, isotropic media will be identical to those of the previous chapters except that the electric fields produced by the given free charges are reduced by a factor of ϵ/ϵ_0 . This reduction in the effective field can be understood as a consequence of the fact that the dipole electric fields of the atoms or molecules of the medium oppose the external field. If the same dielectric medium does not fill space, which is usually the case, or if there are dielectric media with different responses in the region occupied by the electric fields we must consider the boundary conditions on the displacement and electric fields at the boundaries.

5.5 Boundary Conditions

What boundary conditions are appropriate at the interface of two dielectrics? Consider two dielectric media of different responses whose interface is denoted by the surface S in

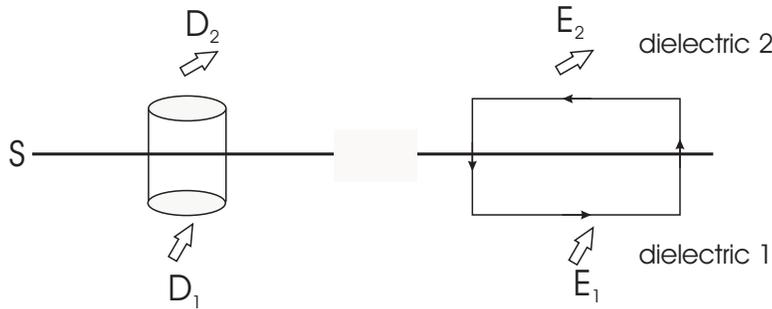


Figure 5.5: Boundary Conditions at a Dielectric Interface

figure (5.5) We begin with the pillbox, whose flat cylindrical surface can be made as small as desired. Therefore, we need to consider only the flat surfaces of the pill box, which are respectively in medium 1 and medium 2. By Gauss' law,

$$\int_{S_1+S_2} \vec{D} \cdot d\vec{S} = \int_{S_1} \vec{D} \cdot d\vec{S} + \int_{S_2} \vec{D} \cdot d\vec{S} = q_f \quad (5.5.1)$$

where σ is the surface density of the charge on the interface. For small surface areas, we can express this condition in the form

$$(\vec{D}_1 \cdot \hat{n}_1 + \vec{D}_2 \cdot \hat{n}_2)\Delta S = \sigma_f \Delta S \quad (5.5.2)$$

where $\vec{D}_{1,2}$ and $\hat{n}_{1,2}$ refer to the electric displacement and unit normals to the surfaces $S_{1,2}$ respectively. Because \hat{n} is a unit outgoing normal, $\hat{n}_1 = -\hat{n}_2$. If we define $D_{i,n} = \vec{D}_i \cdot \hat{n}_2$ then

$$D_{2,n} - D_{1,n} = \sigma_f \quad (5.5.3)$$

says that the normal component of the electric displacement is generally discontinuous across an interface between two dielectric media, the discontinuity being proportional to the surface density of free charge on the surface.¹ In the absence of free charge, the normal component of the displacement vector is continuous.

Next consider the rectangular closed curve on the right of figure (5.5) and integrate \vec{E} around it. Since \vec{E} is conservative, and since we can make the curve as thin as we want,

$$\oint_C \vec{E} \cdot d\vec{r} = \vec{E}_1 \cdot \hat{t}_1 + \vec{E}_2 \cdot \hat{t}_2 \quad (5.5.4)$$

¹It is important to keep account of the sign convention: the unit normal chosen to define the normal component of both displacement vectors is \hat{n}_2 , which is directed *into* medium 2

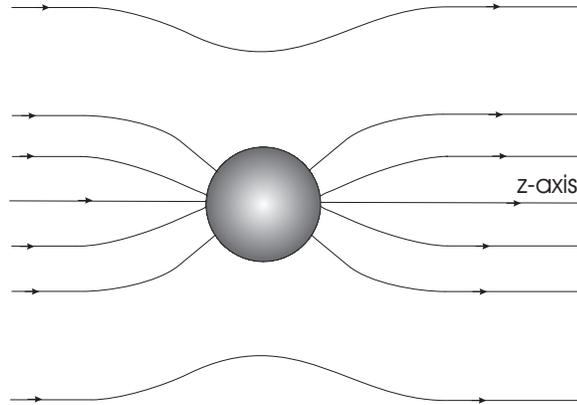


Figure 5.6: Dielectric sphere in a uniform electric field.

where \hat{t}_1 and \hat{t}_2 are the unit tangent vectors to the two sides of the curve shown in the figure. We have ignored the other two sides which can be made arbitrarily small by making the rectangle as needle-like as we wish. Then because $\hat{t}_1 = -\hat{t}_2$ we have

$$E_{2,t} - E_{1,t} = 0 \quad (5.5.5)$$

where, in the same notation we used earlier, $E_{i,t} = \vec{E}_i \cdot \hat{t}_2$. This says that the tangential component of the electric field must be continuous across the interface. Two points are worth noting

- In the case of linear, isotropic media we have seen that the electric displacement is directly proportional to the electric field, $\vec{D} = \epsilon\vec{E}$. Therefore, if the interface lies between two linear and isotropic dielectrics that both the tangential component and the normal component of \vec{D} (equivalently \vec{E}) are continuous across the interface. Thus no free surface charge develops on the interface between two linear, isotropic dielectrics.
- If one of the media present at the interface is a conductor, the electric field in the other can only be perpendicular to the interface. This follows from the fact that $\vec{E} = 0$ in the conductor and that E_t is continuous. For the same reason, $D_n = \sigma$ is the normal component of \vec{D} in the dielectric.

5.6 Examples

5.6.1 Dielectric sphere in a uniform electric field

We have considered earlier how a uniform electric field in a vacuum is modified by the presence of a spherical conductor. We will now replace the conductor by a spherical dielectric medium that is both linear and isotropic for simplicity and see how it modifies the electric field.

We are interested in solving Laplace's equation in spherical coordinates, whose origin we situate at the center of the dielectric sphere. The boundary conditions we wish to apply are

$$\vec{E}_\infty = \lim_{r \rightarrow \infty} \vec{E}(\vec{r}) = E_0 \hat{z} \Rightarrow \phi(\vec{r}) = -E_0 z = -E_0 r \cos \theta + \text{const.} \quad (5.6.1)$$

Without loss of generality, we can take the constant in the expression for the asymptotic form of $\phi(\vec{r})$ to be zero. We also require that $\phi(\vec{r})$ is well behaved at $r = 0$ because there is no charge there.

As before, the proper dependence on $\cos \theta$ near infinity requires that

$$\phi_{\text{out}}(\vec{r}) = a_1 r \cos \theta + \frac{b_0}{r} + \frac{b_1}{r^2} \cos \theta + \frac{b_2}{r^3} (3 \cos^2 \theta - 1) + \dots \quad (5.6.2)$$

with $a_1 = -E_0$. Inside the sphere, regularity of the potential at the origin implies that

$$\phi_{\text{in}}(\vec{r}) = \sum_l a'_l P_l(\cos \theta) = a'_0 + a'_1 r \cos \theta + a'_2 r^2 (3 \cos^2 \theta - 1) + \dots \quad (5.6.3)$$

where no φ dependence has been considered because the system is azimuthally symmetric. Continuity of the potential across the boundary gives

$$-E_0 R \cos \theta + \frac{b_0}{R} + \frac{b_1}{R^2} \cos \theta + \frac{b_2}{R^3} (3 \cos^2 \theta - 1) + \dots = a'_0 + a'_1 R \cos \theta + a'_2 R^2 (3 \cos^2 \theta - 1) + \dots \quad (5.6.4)$$

which, term by term in $\cos \theta$ gives

$$\begin{aligned} \frac{b_0}{R} &= a'_0 - a'_2 R^2 - \dots \\ -E_0 R + \frac{b_1}{R^2} &= a'_1 R \\ \frac{b_2}{R^3} &= a'_2 R^2, \quad \text{etc.} \end{aligned} \quad (5.6.5)$$

Next, by considering the continuity of \vec{D} on the surface of the dielectric (no free surface charge is present) we find,

$$D_{2,n}|_R = D_{1,n}|_R \Rightarrow \epsilon_0 E_{\text{out},n}|_R = \epsilon E_{\text{in},n}|_R \Rightarrow E_{\text{out},n}|_R = K E_{\text{out},n}|_R \quad (5.6.6)$$

where K is the dielectric constant. Taking the normal derivative of $\phi(\vec{r})$ gives

$$-E_0 \cos \theta - \frac{b_0}{R^2} - \frac{2b_1}{R^3} \cos \theta - \frac{3b_2}{R^4} (\cos^2 \theta - 1) + \dots = K[a'_1 \cos \theta + 2a'_2 R(3 \cos^2 \theta - 1) + \dots] \quad (5.6.7)$$

We see that $b_0 = 0$ and

$$\begin{aligned} -E_0 - \frac{2b_1}{R^3} &= K a'_1 \\ \frac{3b_2}{R^4} &= 2K a'_2 R, \text{ etc.}, \end{aligned} \quad (5.6.8)$$

the last of which is compatible with the last of (5.6.5) if and only if $b_2 = b_3 \dots = 0 = a'_2 = a'_3 = \dots$. We are left with just two constants to evaluate: *viz.*, b_1 and a'_1 . Again, dividing the second equation in (5.6.5) by R , multiplying it by two and adding the result to (5.6.8) gives

$$-3E_0 = (K + 2)a'_1 \Rightarrow a'_1 = -\frac{3E_0}{K + 2} \quad (5.6.9)$$

Inserting this result into (5.6.8) gives

$$b_1 = \left(\frac{K - 1}{K + 2} \right) E_0 R^3 \quad (5.6.10)$$

Thus we have found

$$\phi_{\text{out}} = -E_0 r \cos \theta \left[1 - \left(\frac{K - 1}{K + 2} \right) \frac{R^3}{r^3} \right] \quad (5.6.11)$$

and

$$\phi_{\text{in}} = -\frac{3E_0}{K + 2} r \cos \theta = -\frac{3E_0 z}{K + 2} \quad (5.6.12)$$

The electric field inside the dielectric is constant and points in the z direction,

$$\vec{E}_{\text{in}} = -\partial_z \phi_{\text{in}} = \frac{3}{K + 2} E_0 \hat{z}, \quad (5.6.13)$$

approaching zero inside a conductor ($K \rightarrow \infty$). It is smaller in magnitude than the external field if $K > 1$ and points in the same direction. The polarization of the medium

$$\vec{P} = \chi \vec{E} = 3\epsilon_0 \frac{K - 1}{K + 2} E_0 \hat{z} \quad (5.6.14)$$

also points in the same direction as the external electric field in this case ($K > 1$). The volume density of the polarization charge vanishes identically. Its surface charge density is

$$\sigma_P = \hat{n} \cdot \vec{P} \equiv \hat{r} \cdot \vec{P} = 3\epsilon_0 \left(\frac{K - 1}{K + 2} \right) E_0 \cos \theta \quad (5.6.15)$$

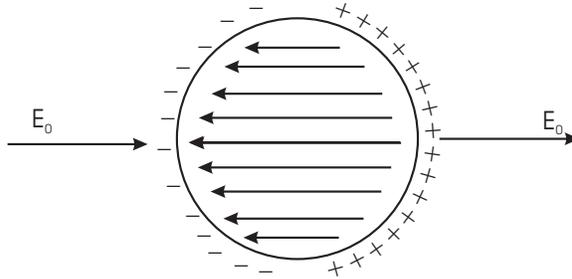


Figure 5.7: Polarization charge density on a dielectric sphere in a uniform electric field.

which is shown in figure (5.7). This has the effect of producing an internal electric field directed opposite to the external, applied field and so reducing the electric field inside the medium.

The problem of a spherical cavity inside a dielectric medium can be handled in precisely the same way as the problem we have considered. Examination of the boundary conditions suggests that the result can be obtained from the results above by simply replacing $K \rightarrow 1/K$. Thus the electric field inside the cavity would be given by

$$\vec{E}_{\text{in}} = \frac{3K}{1+2K} E_0 \hat{z} \quad (5.6.16)$$

which is clearly *larger* than the external electric field when $K > 1$. Similar the polarization of the dielectric is given by

$$\vec{P} = 3\epsilon_0 \frac{1-K}{1+2K} E_0 \hat{z} \quad (5.6.17)$$

which is directed opposite to the applied field and yields a surface charge

$$\sigma_P = \hat{r} \cdot \vec{P} = 3\epsilon_0 \frac{1-K}{1+2K} E_0 \cos \theta \quad (5.6.18)$$

which, for $K > 1$ is arranged oppositely from that shown in figure (5.7) and is such as to enhance the electric field inside the cavity.

5.6.2 Images: Two dielectric media

Consider the following problem: two dielectric media are separated by a plane boundary and a charge q is embedded in one of them, say dielectric 1 (see figure (5.8), in which the boundary is taken to be the $y-z$ plane). We wish to determine the electric field at all points in the two dielectrics. Introduce an image charge situated at a distance $-d$ from the boundary on the x -axis in dielectric 2 and imagine that the potentials in dielectric

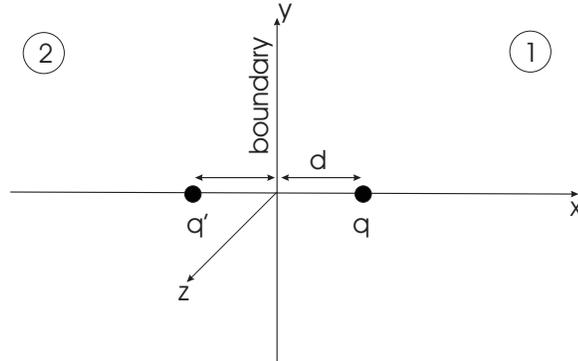


Figure 5.8: Plane boundary between two dielectric media

medium 1 is obtained as a superposition of the electric potentials due to these two charges, q and q' . The potential in medium 1 is then

$$\phi_1(\vec{r}) = \frac{1}{4\pi\epsilon_1} \left[\frac{q}{r} + \frac{q'}{r'} \right]. \quad (5.6.19)$$

In region 2 we imagine that the potential is due to yet another image charge, q'' , located on the x -axis at $x = d$. The potential in region 2 is then

$$\phi_2(\vec{r}) = \frac{q''}{4\pi\epsilon_2 r} \quad (5.6.20)$$

where we have used the notation

$$\begin{aligned} r &= \sqrt{(x-d)^2 + y^2 + z^2} \\ r' &= \sqrt{(x+d)^2 + y^2 + z^2}. \end{aligned} \quad (5.6.21)$$

The two unknowns, *viz.* q' , q'' must be obtained from our boundary conditions, which must be that

- ϕ must be continuous at $x = 0$,
- The normal component of \vec{D} must be continuous across $x = 0$ and
- the tangential component of \vec{E} must be continuous.

The first condition is satisfied if we require

$$\frac{(q + q')}{4\pi\epsilon_1 \sqrt{d^2 + x^2 + z^2}} = \frac{q''}{4\pi\epsilon_2 \sqrt{d^2 + x^2 + z^2}} \quad (5.6.22)$$

which gives the first constraint,

$$\frac{q''}{\epsilon_2} = \frac{q + q'}{\epsilon_1} \quad (5.6.23)$$

Let $r_d = \sqrt{d^2 + x^2 + z^2}$. The normal component of \vec{D} can be taken to be its x -component, since the x -axis is normal to the interface,

$$\vec{D}_{1,x} = -\epsilon_1 \partial_x \phi_1(\vec{r})|_{x=0} = \frac{q'd - qd}{4\pi r_d^3} \quad (5.6.24)$$

and

$$D_{2,x} = -\epsilon_2 \partial_x \phi_2(\vec{r})|_{x=0} = -\frac{q''d}{4\pi r_d^3} \quad (5.6.25)$$

Equating them gives $q - q' = q''$, which is a second (independent) relation between the unknowns. This is sufficient to solve for q' and q'' . Inserting the last relation into the first we find

$$\epsilon_1(q - q') = \epsilon_2(q + q') \Rightarrow q' = \frac{\epsilon_1 - \epsilon_2}{\epsilon_1 + \epsilon_2} q \quad (5.6.26)$$

and therefore

$$q'' = \frac{2\epsilon_2}{\epsilon_1 + \epsilon_2} q \quad (5.6.27)$$

Finally requiring the continuity of the tangential component of \vec{E} , which we can take to be either its y - or z - component, gives nothing new, for (taking the y -components)

$$\frac{(q + q')y}{4\pi\epsilon_1 r_d^3} = \frac{q''y}{4\pi\epsilon_2 r_d^3} \quad (5.6.28)$$

gives the same constraint as did the continuity of $\phi(\vec{r})$.

5.7 Electrostatic Energy of a Charge distribution

We now turn to a key property of charge distributions that we have completely ignored so far, the energy. We have pretended that a static charge distribution was given to us, but not asked the question of how the distribution may come into being. Collecting charges in a distribution involves performing a certain amount of work. Charges exert forces on one another and this alone tells us that every charge in a distribution will have a certain potential energy, equal to the work required to bring the charge from some standard fixed point to its final location within the distribution. To make the situation as general as possible, let us consider arbitrary distributions of charge placed in an external electric field, which we will denote by the superscript “ext”. Such a distribution has two sources of energy, attributable to

- the external field and
- the internal field of the other charges in the distribution.

The first is easily calculated. If $\phi^{\text{ext}}(\vec{r})$ represents the electrostatic potential of the external field then an infinitesimal charge placed in this field would have a potential energy

$$dU^{\text{ext}}(\vec{r}') = dq(\vec{r}')\phi^{\text{ext}}(\vec{r}') \quad (5.7.1)$$

Expressing $dq(\vec{r}')$ in terms of the density of charge, $\rho(\vec{r}')$, and summing over all charges in the distribution we have

$$U^{\text{ext}} = \int_V d^3\vec{r}' \rho(\vec{r}')\phi^{\text{ext}}(\vec{r}') \quad (5.7.2)$$

where the integration is carried out over the entire volume of the distribution. If the distribution is discrete, the analogous expression for the potential energy is

$$U^{\text{ext}} = \sum_j q_j \phi^{\text{ext}}(\vec{r}_j) \quad (5.7.3)$$

The calculation of the internal potential energy (or self energy) of a distribution is more complicated because every time a charge is placed at a given location it changes the electric field of the distribution itself.

The self-energy of a distribution will be constructed relative to the potential energy of the charges when all are separated by an infinite distance from each other.

5.7.1 Electrostatic self-energy of a discrete distribution

Imagine that we bring the charges from infinity to their final locations within the distribution one by one. No work is expended in bringing in the first charge. However, to bring in the second charge we must work against the electric field due to the first, thus

$$U_2 = \frac{q_1 q_2}{4\pi\epsilon_0 |\vec{r}_2 - \vec{r}_1|} \quad (5.7.4)$$

While bringing in the third charge we work against the combined field of the first and the second,

$$U_3 = \frac{1}{4\pi\epsilon_0} \left[\frac{q_1 q_3}{|\vec{r}_3 - \vec{r}_1|} + \frac{q_2 q_3}{|\vec{r}_3 - \vec{r}_2|} \right] \quad (5.7.5)$$

If we go on in this way, the energy expended in bringing in the k^{th} charge is

$$U_k = \sum_{j=1}^{k-1} \frac{q_j q_k}{4\pi\epsilon_0 |\vec{r}_k - \vec{r}_j|} \quad (5.7.6)$$

Summing up the contributions to the energy expended in bringing in each and every charge to its specified location within the distribution gives the self energy of the distribution

$$U = \sum_{k=1}^N \sum_{j=1}^{k-1} \frac{q_j q_k}{4\pi\epsilon_0 |\vec{r}_k - \vec{r}_j|} = \sum_{k=1}^N \sum_{j=1}^{k-1} U_{jk} \quad (5.7.7)$$

The matrix U_{jk} is symmetric, so we can write

$$U = \frac{1}{2} \sum_{j,k=1, j \neq k}^N U_{jk} = \frac{1}{2} \sum_{j,k=1, j \neq k}^N \frac{q_j q_k}{4\pi\epsilon_0 |\vec{r}_k - \vec{r}_j|} \quad (5.7.8)$$

where we omit the term $j = k$. Again, since

$$\phi_k(\vec{r}_k) = \sum_{j \neq k} \frac{q_j}{4\pi\epsilon_0 |\vec{r}_k - \vec{r}_j|} \quad (5.7.9)$$

is the electric potential due to all the other charges (excluding q_k) at the location of charge q_k , we can write this expression as

$$U = \frac{1}{2} \sum_{k=1}^N q_k \phi_k \quad (5.7.10)$$

Although we have imagined constructing a distribution of charges in a vacuum, this final expression holds even if the distribution is created in a region of space that is filled with one or more dielectric media. In that case, ϕ_k is the electric potential subject to whatever boundary conditions may be present.

5.7.2 Electrostatic energy of a system of conductors

If we replace the point charges with conductors the electrostatic energy of the system of isolated conductors is just

$$U = \frac{1}{2} \sum_j Q_j \phi_j \quad (5.7.11)$$

where Q_j is the total charge on the j^{th} conductor and ϕ_j is its constant electrostatic potential. The electric potential of any conductor is a linear function of all the charges present in the system. We can express this statement as

$$\phi_i = \sum_j p_{ij} Q_j \quad (5.7.12)$$

where p_{ij} are called the coefficients of potential and depend in a complicated way on the distribution and geometry of the conductors, but they are independent of the charges. This statement follows quite trivially from the linearity of Laplace's equation, which implies that its solutions are superposable. The coefficients of potential satisfy the following three conditions: (a) $p_{ij} > 0 \forall i, j$, (b) $p_{ii} \geq p_{ij} \forall j$ and (c) $p_{ij} = p_{ji}$ (the matrix \hat{p} is symmetric). We will now prove these statements.

- $p_{ii} \geq p_{ij} > 0 \forall i, j$. Consider a situation in which only one of the conductors, say i , is charged and all the other conductors are uncharged. Let the charge on i be Q_i (which we take to be positive, for convenience). The charged conductor acts as a source for the potential in all of space. In particular the potentials on any conductor will be

$$\phi_j = p_{ji}Q_i \quad (5.7.13)$$

We want to argue that ϕ_j is positive. Now since all potentials are measured with respect to a zero potential reservoir and since only conductor i is charged, every field line of displacement will originate on it and flow into the zero potential reservoir, possibly by way of other conductors. Thus $\phi_i > 0$ which implies that $p_{ii} > 0$. On the other hand, every line of electric displacement that enters a conductor other than i will have originated on i and every line leaving that conductor will be flow into the zero potential reservoir, again possibly by way of some other conductor. Thus $\phi_j > 0 \forall j$, or $p_{ii} > p_{ij} > 0 \forall j$.

Suppose j in the above argument is completely shielded by another conductor, say k . This can occur if j lies wholly within the conducting *shell* that is k . Then the electric field in the region between the conductors is identically zero by Gauss' law, implying that the two conductors are at the same potential, *i.e.*, $p_{ij} = p_{ik}$. In particular, if the shielding conductor is i itself then $p_{ii} = p_{ij}$. The argument is independent of the number of charged conductors and, combined with the result of the previous paragraph, proves the statement $p_{ii} \geq p_{ik} > 0$.

- $p_{ij} = p_{ji}$. This follows directly from the energy of the system of conductors

$$U = \frac{1}{2} \sum_{i,j=1}^N p_{ij}Q_iQ_j \quad (5.7.14)$$

Let us imagine changing the charge on conductor m by dQ_m . The change in energy would be

$$dU = \frac{1}{2} \sum_{j=1}^N (p_{mj} + p_{jm})Q_jdQ_m \quad (5.7.15)$$

But we can compute this change in energy by calculating the work that must be done to bring the charge dQ_m from a zero potential reservoir to the conductor m . This is just

$$dU = dQ_m \phi_m = \sum_{j=1}^N p_{mj} Q_j dQ_m \quad (5.7.16)$$

We find that

$$\frac{1}{2}(p_{mj} + p_{jm}) = p_{mj} \Rightarrow p_{jm} = p_{mj} \quad (5.7.17)$$

The conditions satisfied by p_{ij} are sufficient to ensure that the determinant of \hat{p} does not vanish, therefore the matrix \hat{p} may be inverted to obtain the charges as linear functions of the potentials,

$$Q_i = \sum_j C_{ij} \phi_j \quad (5.7.18)$$

The C_{ij} are called the coefficients of capacitance. The diagonal elements are called the capacitances and the off diagonal elements are called the coefficients of electrostatic induction. The matrix \hat{C} will inherit its properties from the matrix \hat{p} .² The energy of the system of conductors can be written in terms of the coefficients of capacitance as

$$U = \frac{1}{2} \sum_{jk} C_{jk} \phi_j \phi_k \quad (5.7.19)$$

As an example, consider the description of an electronic device called the “capacitor”. It consists of a pair of conductors with equal and opposite charges, Q and $-Q$ at potentials V_1 and V_2 respectively. We should have

$$\begin{bmatrix} Q \\ -Q \end{bmatrix} = \begin{bmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{bmatrix} \begin{bmatrix} V_1 \\ V_2 \end{bmatrix} \quad (5.7.20)$$

or

$$\begin{aligned} V_1 &= -\frac{Q(C_{11} + C_{21})}{C_{11}C_{22} - C_{12}^2} \\ V_2 &= \frac{Q(C_{12} + C_{22})}{C_{11}C_{22} - C_{12}^2} \end{aligned} \quad (5.7.21)$$

so that

$$\Delta V = V_2 - V_1 = \frac{Q(C_{11} + 2C_{12} + C_{22})}{C_{11}C_{22} - C_{12}^2} \quad (5.7.22)$$

²Problem: Determine the properties of the coefficients of capacitance from those of the coefficients of potential.

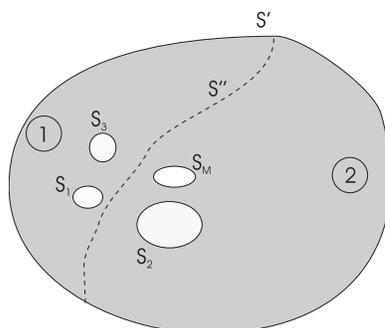


Figure 5.9: Dielectric media with conducting surfaces

and

$$\frac{Q}{\Delta V} = \frac{C_{11}C_{22} - C_{12}^2}{(C_{11} + C_{22} + 2C_{12})} = C \quad (5.7.23)$$

is often called the “capacitance” of the capacitor.³

5.7.3 Electrostatic energy of a continuous distribution

Obviously the same conditions would apply to a continuous distribution. Imagine that a continuous distribution is constructed from scratch inside a region that is filled by one or more dielectric media and containing one or more conductors [see figure (5.9)]. Following similar reasoning and putting together infinitesimal charges, we arrive at

$$U = \frac{1}{2} \int_V d^3\vec{r}' \rho_f(\vec{r}')\phi(\vec{r}') + \frac{1}{2} \oint_S d^2\vec{r}' \sigma_f(\vec{r}')\phi(\vec{r}') \quad (5.7.24)$$

where $\rho_f(\vec{r}')$ and $\sigma_f(\vec{r}')$ are just the volume and surface densities of the *free* charges in the distribution, V is the total volume of the dielectric media (excluding the conductors) and S is the surface bounding V . If there are M conductors embedded within the dielectrics, $S = S' + S_1 + S_2 + \dots S_M$. $\phi(\vec{r}')$ is the electrostatic potential due to the entire charge distribution, bound and free, and subject to whatever boundary conditions are appropriate for the system. Note however that the expression only accounts for the work that is done on the *free* charge, and does not account for work that is necessary to produce a certain state of polarization in the medium.

We will now make two simplifying assumptions, *viz.*,

³Problem: Compute the energy of the capacitor.

- all the free surface charges lie on interfaces with conductors that may be present in the dielectrics (such interfaces will be called *conducting* surfaces as opposed to the non-conducting surfaces which would be interfaces between the dielectrics), and
- the free charge is bounded.

With the first assumption,

$$U = \frac{1}{2} \int_V d^3 \vec{r}' \rho_f(\vec{r}') \phi(\vec{r}') + \frac{1}{2} \oint_{S_1+S_2+\dots} d^2 \vec{r}' \sigma_f(\vec{r}') \phi(\vec{r}') \quad (5.7.25)$$

because no other surface will contribute to the energy integral. On the other hand, with help from Gauss' law for dielectrics, the expression for the volume density of free charge

$$\rho_f = \vec{\nabla} \cdot \vec{D} \quad (5.7.26)$$

and the expression for the surface density of free charge,

$$\sigma_f = -\hat{n} \cdot \vec{D} \quad (5.7.27)$$

where \hat{n} is the unit normal vector normal to the conducting surface and directed *out* of the dielectric (hence the negative sign), we may rewrite the volume charge density and the surface charge density on the conducting surfaces in terms of the electric displacement vector

$$U = \frac{1}{2} \left[\int_V d^3 \vec{r}' [\vec{\nabla}' \cdot \vec{D}(\vec{r}')] \phi(\vec{r}') - \oint_{S_1+S_2+\dots} d^2 \vec{r}' [\hat{n}' \cdot \vec{D}(\vec{r}')] \phi(\vec{r}') \right] \quad (5.7.28)$$

In the first integral, we could replace the divergence of \vec{D} by

$$(\vec{\nabla}' \cdot \vec{D})\phi = \vec{\nabla}' \cdot (\phi \vec{D}) - (\vec{\nabla}' \phi) \cdot \vec{D} = \vec{\nabla}' \cdot (\phi \vec{D}) + \vec{E} \cdot \vec{D} \quad (5.7.29)$$

and use Gauss' theorem to write the volume integral of the divergence as a surface integral, with the result

$$\int_V d^3 \vec{r}' \vec{\nabla}' \cdot [\phi(\vec{r}') \vec{D}(\vec{r}')] = \int_{S'+S_1+S_2+\dots} d^2 \vec{r}' [\hat{n}' \cdot \vec{D}(\vec{r}')] \phi(\vec{r}') \quad (5.7.30)$$

since Gauss' theorem requires a unit *outgoing* normal. It should be evident that we will be left only with a surface integral over the non-conducting outer surface S' , plus the contribution from the second term in (5.7.29). If our charge distribution is bounded the electric potential falls off inversely as the distance from the distribution, *i.e.*, as r^{-1} and the electric displacement vector as r^{-2} whereas the area grows as r^2 , the surface integral

over S' will fall off as r^{-1} and, if this outer surface is taken to infinity, the integral over S' also vanishes. Under these conditions,

$$U = \frac{1}{2} \int_V d^3\vec{r}' \vec{E}(\vec{r}') \cdot \vec{D}(\vec{r}') \quad (5.7.31)$$

where the integration is over all dielectrics (external to the conductors), but may be extended to all of space because the electric field inside a conductor vanishes. We are led therefore to the concept of the energy density of the electrostatic field, which we define as

$$u = \frac{1}{2} \vec{D} \cdot \vec{E} \quad (5.7.32)$$

This expression is valid only for *linear* dielectrics. Otherwise the energy of the final configuration must be calculated as follows: consider an infinitesimal *change* in energy due to some change in the macroscopic free charge density ρ existing in all of space. The work done to accomplish this change is

$$\begin{aligned} \delta W &= \int d^3\vec{r}' \delta\rho_f(\vec{r}')\phi(\vec{r}') = \int d^3\vec{r}' \vec{\nabla} \cdot \delta\vec{D}(\vec{r}')\phi(\vec{r}') \\ &= \int d^3\vec{r}' \left[\vec{\nabla} \cdot \{\delta\vec{D}(\vec{r}')\phi(\vec{r}')\} + \delta\vec{D}(\vec{r}') \cdot \vec{E}(\vec{r}') \right] \end{aligned} \quad (5.7.33)$$

The first term is a surface integral which may be made to vanish by taking the surface out to infinity. The total electrostatic energy can be written formally by allowing \vec{D} to be brought the initial value $\vec{D} = 0$ to its final value \vec{D} :

$$W = \int d^3\vec{r}' \int_0^D \vec{E}(\vec{r}') \cdot \delta\vec{D}(\vec{r}') \quad (5.7.34)$$

If the medium is not linear, this integral will, in general, depend on the process by which \vec{D} is brought from zero to its final value, *i.e.*, it will depend on the past history of the system (hysteresis).

5.8 Energy of a linear dielectric in an external field

An interesting problem is of the change in energy when a linear dielectric object is placed in an external electric field with fixed sources. For simplicity, we assume that the dielectric is absent initially and the medium is just the vacuum in which there is an initial electric field is \vec{E}_0 , sourced by charges which are described by the density ρ_0 . The initial energy of the system is then

$$W_0 = \frac{1}{2} \int d^3\vec{r}' (\vec{E}_0 \cdot \vec{D}_0) \quad (5.8.1)$$

where $\vec{D}_0 = \epsilon_0 \vec{E}$. When the dielectric, of volume V , is introduced into the initial vacuum, the electric field gets modified to \vec{E} and the displacement to $\vec{D} = \epsilon(\vec{r})\vec{E}$. We have included a dependence of ϵ on \vec{r} to indicate that it changes outside the dielectric. We will assume that $\epsilon(\vec{r})$ is a smooth function, changing rapidly but continuously at the boundary of the dielectric. This is only to avoid mathematical complications that are not intrinsic to the physical problem at hand. The energy is now

$$W = \frac{1}{2} \int d^3\vec{r} (\vec{E} \cdot \vec{D}) \quad (5.8.2)$$

and the change in energy is

$$\begin{aligned} W - W_0 &= \frac{1}{2} \int d^3\vec{r} (\vec{E} \cdot \vec{D} - \vec{E}_0 \cdot \vec{D}_0) \\ &= \frac{1}{2} \int d^3\vec{r} (\vec{E} + \vec{E}_0) \cdot (\vec{D} - \vec{D}_0) + \frac{1}{2} \int d^3\vec{r} (\vec{E} \cdot \vec{D}_0 - \vec{E}_0 \cdot \vec{D}) \end{aligned} \quad (5.8.3)$$

The first integral vanishes because, by superposition and the conservation of \vec{E} and \vec{E}_0 , we can set $\vec{E} + \vec{E}_0 = \vec{\nabla}\phi$ giving for the first term

$$\frac{1}{2} \int d^3\vec{r} \vec{\nabla}\phi \cdot (\vec{D} - \vec{D}_0) = -\frac{1}{2} \int d^3\vec{r} \phi \vec{\nabla} \cdot (\vec{D} - \vec{D}_0) \quad (5.8.4)$$

where we have ignored the surface term in going to the right hand expression because we take the integration surface to infinity. This expression vanishes if no new free charge has been introduced. The energy change is therefore

$$W - W_0 = \Delta W = \frac{1}{2} \int d^3\vec{r} (\vec{E} \cdot \vec{D}_0 - \vec{E}_0 \cdot \vec{D}) = \int d^3\vec{r} (\epsilon - \epsilon_0) \vec{E} \cdot \vec{E}_0 \quad (5.8.5)$$

While the integration appears to be over all of space it is really only over the region of space where $\epsilon \neq \epsilon_0$, *i.e.*, over V , the volume of the dielectric that was introduced. It follows that

$$\Delta W = -\frac{1}{2} \int_V d^3\vec{r} \vec{P} \cdot \vec{E}_0 \quad (5.8.6)$$

where \vec{P} is the polarization of the dielectric. The energy density of the electric field has changed by

$$w = -\frac{1}{2} \vec{P} \cdot \vec{E}_0 \quad (5.8.7)$$

which must be the potential energy of the dielectric. This should be compared with the corresponding result for a dipole in an external electric field (2.3.16). It differs by a factor of $\frac{1}{2}$, which accounts for the work done by the field to polarize the medium in the first

place. The result shows that the dielectric will move toward regions of increasing field E_0 , provided $\epsilon > \epsilon_0$.

A consequence of this discussion is that dielectric media will experience forces, tending to move toward regions of increasing field. If the sources of the applied external field are maintained fixed there is no external source of energy and the change in energy can be interpreted as the potential energy acquired by the dielectric medium. The force acting on a dielectric medium is then simply

$$\vec{F} = - \vec{\nabla}(\Delta W) \Big|_Q \quad (5.8.8)$$

where the subscript is used to indicate that the sources are held fixed.

5.9 Multipoles: The multipole expansion

In describing dielectric media, it was sufficient to consider only the electric potential due to dipoles. This is of course just an approximation in which the molecules of the material are presumed neutral and the dipole approximation was used as the leading macroscopic effect of distorting the molecules of the medium by the application of an external field. Higher order effects also exist and the purpose of this chapter is to classify and examine them in some detail.

Begin by considering the expression for the electrostatic potential due to some distribution of charge when no non-trivial boundaries are present,

$$\phi(\vec{r}) = \frac{1}{4\pi\epsilon_0} \int d^3\vec{r}' \frac{\rho(\vec{r}')}{|\vec{r} - \vec{r}'|} \quad (5.9.1)$$

where

$$|\vec{r} - \vec{r}'|^{-1} = [r^2 + r'^2 - 2\vec{r} \cdot \vec{r}']^{-1/2} \quad (5.9.2)$$

Let the origin of coordinates be located “near” the charge distribution and let the observation take place far from the source distribution, which is assumed to be localized so that $r \gg r'$ for all points \vec{r}' locating an element of charge, then

$$|\vec{r} - \vec{r}'|^{-1} = \frac{1}{r} \left[1 + \left(\frac{r'}{r} \right)^2 - \frac{2\vec{r} \cdot \vec{r}'}{r^2} \right]^{-1/2} \quad (5.9.3)$$

Using the generalized binomial expansion

$$(1 + x)^a = 1 + ax + \frac{a(a-1)}{2!}x^2 + \frac{a(a-1)(a-2)}{3!}x^3 + \dots \quad (5.9.4)$$

or, equivalently

$$(1+x)^{-1/2} = 1 - \frac{1}{2}x + \frac{3}{8}x^2 - \frac{5}{16}x^3 \dots \quad (5.9.5)$$

In the case we are examining, $x = (r'/r)^2 - 2(\vec{r} \cdot \vec{r}')/r^2$, so we find

$$\begin{aligned} \frac{1}{|\vec{r} - \vec{r}'|} &= \frac{1}{r} \left[1 - \frac{1}{2} \{ (r'/r)^2 - 2(\vec{r} \cdot \vec{r}')/r^2 \} + \frac{3}{8} \{ (r'/r)^2 - 2(\vec{r} \cdot \vec{r}')/r^2 \}^2 + \dots \right] \\ &= \frac{1}{r} + \frac{\vec{r} \cdot \vec{r}'}{r^3} + \frac{1}{2r^5} [3(\vec{r} \cdot \vec{r}')^2 - r'^2 r^2] + \dots \mathcal{O}(r^{-4}) \end{aligned} \quad (5.9.6)$$

and the electrostatic potential (5.9.1) becomes

$$\phi(\vec{r}) = \frac{1}{4\pi\epsilon_0} \int d^3\vec{r}' \rho(\vec{r}') \left[\frac{1}{r} + \frac{\vec{r} \cdot \vec{r}'}{r^3} + \frac{1}{2r^5} [3(\vec{r} \cdot \vec{r}')^2 - r'^2 r^2] + \dots \mathcal{O}(r^{-4}) \right] \quad (5.9.7)$$

The first term can be written as

$$\phi_0(\vec{r}) = \frac{Q}{4\pi\epsilon_0 r} \quad (5.9.8)$$

where $Q = \int d^3\vec{r}' \rho(\vec{r}')$ is the total charge of the distribution. To this order the solution behaves as a point charge located at the origin of the coordinates. This is called the “monopole?” contribution. The second term in the expansion is

$$\phi_1(\vec{r}) = \frac{1}{4\pi\epsilon_0 r^3} \vec{r} \cdot \int d^3\vec{r}' \rho(\vec{r}') \vec{r}' \quad (5.9.9)$$

so, taking into account the definition of the electric dipole moment \vec{p} in (2.2.4), it can be written as

$$\phi_1(\vec{r}) = \frac{\vec{p} \cdot \vec{r}}{r^3} \quad (5.9.10)$$

which will be recognized as the electric potential due to an electric dipole located at the origin. It is called the electric “dipole” contribution. The third term can be rewritten as

$$\phi_2(r) = \frac{1}{8\pi\epsilon_0 r^5} \sum_{ij} x_i x_j \int d^3\vec{r}' \rho(\vec{r}') [3x'_i x'_j - r'^2 \delta_{ij}] = \frac{1}{8\pi\epsilon_0 r^5} \sum_{ij} Q_{ij} x_i x_j \quad (5.9.11)$$

where the integral

$$Q_{ij} = \int d^3\vec{r}' \rho(\vec{r}') [3x'_i x'_j - r'^2 \delta_{ij}] \quad (5.9.12)$$

is called the quadrupole moment tensor of the distribution. One can continue in this fashion to obtain higher moments of the distribution, each correction getting suppressed by an additional factor of $1/r$.

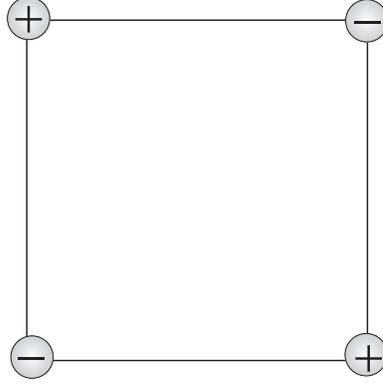


Figure 5.10: Quadrupole moment distribution.

It is interesting to understand the physical significance of this expansion. At very large distances from the distribution (*i.e.*, distances much larger than the characteristic dimension of the distribution itself), it acts as a point charge to zeroth order. The first correction to the effective point charge is the dipole term, which can be interpreted as the potential due to two charges separated by some distance that is characteristic of the distribution. The next correction is the quadrupole term and so on.

We will now show that the quadrupole term can be interpreted as the lowest order electric potential far from a distribution of four charges of equal magnitude located on the vertices of a square of characteristic length a as shown in figure (5.10). Locating the origin of coordinates at the lower left charge, the two positive charges are, say, at positions $\vec{r}_1 = (a, 0, 0)$ and $\vec{r}_2 = (0, a, 0)$ and the diagonally opposite (negative) charge is at $\vec{r}_3 = (a, a, 0)$. The electric potential at any point \vec{r} is then

$$\phi(\vec{r}) = \frac{q}{4\pi\epsilon_0} \left[-\frac{1}{|\vec{r}|} + \frac{1}{|\vec{r} - \vec{r}_1|} + \frac{1}{|\vec{r} - \vec{r}_2|} - \frac{1}{|\vec{r} - \vec{r}_3|} \right] \quad (5.9.13)$$

Expanding the right hand side for $a/r \ll 1$,

$$|\vec{r} - \vec{r}_i|^{-1} = r^{-1} \left[1 + \frac{r_i^2}{r^2} - 2\frac{\vec{r} \cdot \vec{r}_i}{r^2} \right]^{-1/2} \approx r^{-1} \left[1 + \frac{\vec{r} \cdot \vec{r}_i}{r^2} + \frac{1}{2r^4} \{3(\vec{r} \cdot \vec{r}_i)^2 - r_i^2 r^2\} + \dots \right] \quad (5.9.14)$$

where we have retained only the first three terms. The first is the monopole term, the second is the dipole term and so on. It can be checked that the monopole term and the dipole term vanish identically, giving the lowest order contribution from the third term

(the quadrupole term) on the right hand side,

$$\phi(\vec{r}) = \frac{1}{8\pi\epsilon_0 r^5} \sum_i q_i [3(\vec{r} \cdot \vec{r}_i)^2 - r_i^2 r^2] \quad (5.9.15)$$

All the moments can all be interpreted in this way as the lowest order contributions to the electric potentials due to 2^n charges.

The moments of a distribution depend upon the manner in which the charges are distributed about the origin. For example, consider a single charge located away from the origin at \vec{r}_0 . It admits all multipoles as long as $\vec{r}_0 \neq 0$. In general the multipole expansion depends on the choice of origin, only the the lowest order multipole being independent of it.

Chapter 6

Currents and the Magnetic Field

Up until now we have examined only static charges and charge distributions. When charges are in motion the two laws of electrostatics,

$$\begin{aligned}\vec{\nabla} \cdot \vec{D} = \rho_f &\Leftrightarrow \oint_S \vec{D} \cdot d\vec{S} = q_f, \quad (\text{Gauss}) \\ \vec{\nabla} \times \vec{E} = 0 &\Leftrightarrow \oint \vec{E} \cdot d\vec{r} = 0\end{aligned}\tag{6.0.1}$$

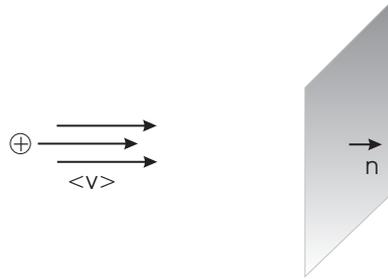
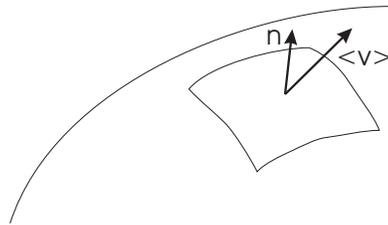
where q_f is the free charge, will be modified and enhanced by effects that are peculiar to the motion of the charges.

6.1 Current and Current Density

Moving charges constitute a current, which is defined in the following way. Consider some arbitrary surface S and let ΔQ represent the charge that is transported across S in time Δt , then

$$i = \lim_{\Delta t \rightarrow 0} \frac{\Delta Q}{\Delta t} = \frac{dQ}{dt}\tag{6.1.1}$$

is called the current across S and is measured in **amperes**: a current of 1 Ampere is said to have flowed when a charge of 1 Coulomb is transported across S in 1 second. Charge in materials is usually transported by electrons (carrying a negative charge of 1.6×10^{-19} C) and/or “holes” *i.e.*, atoms that have lost an electron in the outermost shell and are therefore positively charged. We will consider such objects as “charge carriers”. Let there be n charge carriers per unit volume of the material, each with a charge q , and let $\langle \vec{v} \rangle$ represent the average (drift) velocity of the carriers. Of course, individual carriers will move more or less randomly about this mean velocity. The net effect of the random motion will be ignorable and only $\langle \vec{v} \rangle$ will be relevant. Consider first, a surface δS , whose

Figure 6.1: Normal to S is parallel to $\langle \vec{v} \rangle$.Figure 6.2: Normal to S is not parallel to $\langle \vec{v} \rangle$.

normal is parallel to $\langle \vec{v} \rangle$, as in figure (6.1). The total number of charge carriers that are transported across the surface δS in a time Δt is then the number of carriers contained in a volume $\langle \vec{v} \rangle \Delta t \delta S$, *i.e.*,

$$\Delta N = n \langle \vec{v} \rangle \Delta t \delta S \quad (6.1.2)$$

The number of carriers crossing the surface δS per unit time is then

$$\delta \nu = \frac{\Delta N}{\Delta t} = n \langle \vec{v} \rangle \delta S \quad (6.1.3)$$

and the current across δS is consequently $\delta i = nq \langle \vec{v} \rangle \delta S$. More generally, if the normal \hat{n} , of δS , is not parallel to $\langle \vec{v} \rangle$, as in figure (6.2), then

$$\delta \nu = \frac{\Delta N}{\Delta t} = n (\langle \vec{v} \rangle \cdot \hat{n}) \delta S \Rightarrow \delta i = nq (\langle \vec{v} \rangle \cdot \hat{n}) \delta S \quad (6.1.4)$$

and, most generally, if there are also several types of charge carriers,

$$\delta i = \sum_i n_i q_i \langle \vec{v}_i \rangle \cdot \hat{n} \delta S \quad (6.1.5)$$

where q_i is the charge carried by carrier type i with number density n_i and drift velocity $\langle \vec{v}_i \rangle$. The quantity

$$\vec{j} = \sum_i n_i q_i \langle \vec{v}_i \rangle, \quad (6.1.6)$$

is called the current density, in terms of which one might express the current

$$i = \int_S \vec{j} \cdot d\vec{S}. \quad (6.1.7)$$

In applications, the current density is the more convenient of the two quantities measuring the transport of charge.

Consider now a closed surface S . The net charge contained in the volume enclosed by S is simply given by

$$Q = \int_V d^3\vec{r} \rho(\vec{r}, t) \quad (6.1.8)$$

where ρ is the volume density of charge. Now the rate of change of the charge contained in S

$$\frac{dQ}{dt} = \int_V d^3\vec{r} \frac{\partial \rho}{\partial t} \quad (6.1.9)$$

must be equal to the rate at which charge flows into or out of S . This is expressed by

$$\int_V d^3\vec{r} \frac{\partial \rho}{\partial t} = - \oint_S \vec{j} \cdot d\vec{S} = - \int_V d^3\vec{r} \vec{\nabla} \cdot \vec{j} \quad (6.1.10)$$

The minus sign appears because if charge is *leaving* the bounded volume, then $\langle \vec{v} \rangle$ and hence \vec{j} is outgoing, *i.e.*, the surface integral of \vec{j} over S is positive while the rate of change of Q is negative and vice versa. We therefore obtain the continuity equation, which expresses the conservation of charge

$$\frac{\partial \rho}{\partial t} + \vec{\nabla} \cdot \vec{j} = 0 \quad (6.1.11)$$

This simple equation will shortly become crucial in determining the consistency of the equations of the electromagnetic field.

6.2 The Magnetic Field

Charges that are in motion with respect to the observer give rise to a new field called the magnetic field and denoted by \vec{B} . The magnetic field and the electric field will later be combined into one field called the electromagnetic field, but at present we will treat them separately. The presence of a magnetic field implies that other charges in motion

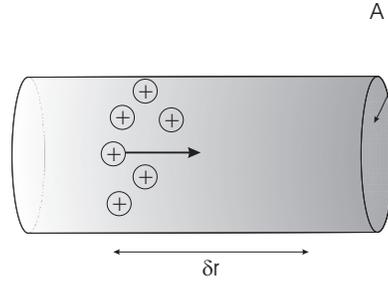


Figure 6.3: Charges flowing in a conducting wire.

experience a Newtonian force. This force can be used to define the magnetic field, just as the electrostatic force was earlier used to define the electrostatic field. A charge q with velocity \vec{v} relative to the laboratory frame in which there is a magnetic field, denoted by \vec{B} , experiences a force given by

$$\vec{F}_m = q\vec{v} \times \vec{B}. \quad (6.2.1)$$

If there is both an electric and a magnetic field present, the total force acting on q is

$$\vec{F} = q(\vec{E} + \vec{v} \times \vec{B}) \quad (6.2.2)$$

This is called the “Lorentz force” on q .

6.3 Force and torque on a current carrying conductor

Consider a typical situation in which many charge carriers, say n per unit volume, with drift velocity $\langle \vec{v} \rangle$ with respect to the Lab frame in which there is a magnetic field \vec{B} . We will think of the charges as flowing through some conducting element of a circuit (figure (6.3)), then the net charge contained in a volume $\delta V = A|\delta\vec{r}|$ is

$$\delta q = qn|\delta\vec{r}|A \quad (6.3.1)$$

and the force acting on the length $\delta\vec{r}$ of the wire will be given by

$$\delta\vec{F} = qnA|\delta\vec{r}|\langle \vec{v} \rangle \times \vec{B} \quad (6.3.2)$$

If, now, we choose the direction of $\delta\vec{r}$ to be the same as the direction of $\langle \vec{v} \rangle$ we find

$$\delta\vec{F} = \underbrace{qnA|\langle \vec{v} \rangle|}_{i} \delta\vec{r} \times \vec{B} = i\delta\vec{r} \times \vec{B} \quad (6.3.3)$$

Integrating over the circuit we would find the total force

$$\vec{F}_{\text{circ}} = i \oint_C d\vec{r} \times \vec{B} \quad (6.3.4)$$

Further, if \vec{r} is the position vector of the circuit element, the torque experienced by it in the presence of the magnetic field is

$$\delta\vec{\tau} = \vec{r} \times \delta\vec{F} = i\vec{r} \times (\delta\vec{r} \times \vec{B}) \quad (6.3.5)$$

Integrating over the entire circuit gives the net torque

$$\vec{\tau}_{\text{circ}} = i \oint_C \vec{r} \times (d\vec{r} \times \vec{B}) \quad (6.3.6)$$

This expression can be greatly simplified if the magnetic field is uniform at least over the area enclosed by the current loop. Suppose, without loss of generality, that the magnetic field points in the \hat{z} direction and consider the identity

$$\vec{r} \times (\delta\vec{r} \times \vec{B}) + \vec{B} \times (\vec{r} \times \delta\vec{r}) + \delta\vec{r} \times (\vec{B} \times \vec{r}) = 0 \quad (6.3.7)$$

(cyclic permutations), which implies that

$$\oint_C \vec{r} \times (d\vec{r} \times \vec{B}) = \oint_C (\vec{r} \times d\vec{r}) \times \vec{B} + \oint_C (\vec{B} \times \vec{r}) \times d\vec{r} \quad (6.3.8)$$

Now

$$(\vec{r} \times d\vec{r}) \times \vec{B} = B(zdx - xdz, zdy - ydz, 0) \quad (6.3.9)$$

and because

$$\oint_C d(xz) = 0 = \oint_C zdx + \oint_C xdz \Rightarrow -\oint_C xdz = \oint_C zdx \quad (6.3.10)$$

we have

$$\oint_C (\vec{r} \times d\vec{r}) \times \vec{B} = 2B \left(\oint_C zdx, \oint_C zdy, 0 \right) \quad (6.3.11)$$

With the same argument,

$$\oint_C (\vec{B} \times \vec{r}) \times d\vec{r} = B \left(-\oint_C zdx, -\oint_C zdy, 0 \right) \quad (6.3.12)$$

showing that in this case

$$\oint_C (\vec{B} \times \vec{r}) \times d\vec{r} = -\frac{1}{2} \oint_C (\vec{r} \times d\vec{r}) \times \vec{B} \quad (6.3.13)$$

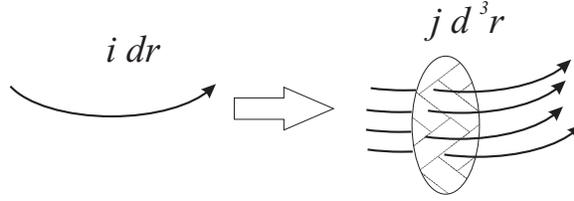


Figure 6.4: Circuits replaced by current carrying regions.

Therefore,

$$\vec{\tau}_{\text{circ}} = \frac{i}{2} \oint_C (\vec{r} \times d\vec{r}) \times \vec{B} = i\vec{A} \times \vec{B} = \vec{\mu} \times \vec{B} \quad (6.3.14)$$

where

$$\vec{A} = \frac{1}{2} \oint_C \vec{r} \times d\vec{r} \quad (6.3.15)$$

is the area enclosed by the circuit. The quantity $\vec{\mu} = i\vec{A}$ is called the magnetic moment of the current carrying loop.

6.4 Magnetostatics: The Biot-Savart Law

Our next question is obviously to determine how the magnetic field itself depends upon the motion of charges. After a series of experiments, the French mathematicians André Marie Ampère and Jean-Baptiste Biot with his assistant Savart, which has since come to be known as the Biot-Savart Law. This law may be stated as follows: if C_1 and C_2 represent two circuits in which flow currents i_1 and i_2 respectively, then the force exerted on circuit “2” by circuit “1” is given by

$$\vec{F}_{1 \rightarrow 2} = \frac{\mu_0}{2\pi} i_1 i_2 \oint_{C_1} \oint_{C_2} \frac{d\vec{r}_2 \times [d\vec{r}_1 \times (\vec{r}_2 - \vec{r}_1)]}{|\vec{r}_2 - \vec{r}_1|^3} \quad (6.4.1)$$

We may now compare this complicated force law with our expression (6.3.4) and determine the magnetic field at the point labeled “2” on circuit C_2 due to the current flowing through C_1 . It is simply

$$\vec{B}(\vec{r}) = \frac{\mu_0}{2\pi} i_1 \oint_{C_1} \frac{d\vec{r}_1 \times (\vec{r} - \vec{r}_1)}{|\vec{r} - \vec{r}_1|^3} \quad (6.4.2)$$

(where we have simply replaced \vec{r}_2 by \vec{r}). Now we can rid ourselves of the explicit reference to circuits and conductors, since it is really the current *i.e.*, the charges in motion that

generates the magnetic field, by replacing $id\vec{r} \rightarrow \vec{j}d^3\vec{r}$ as depicted in figure (6.4), which gives

$$B(\vec{r}) = \frac{\mu_0}{2\pi} \int_V d^3\vec{r}' \frac{\vec{j}(\vec{r}') \times (\vec{r} - \vec{r}')}{|\vec{r} - \vec{r}'|^3} \quad (6.4.3)$$

An important consequence is the following law (Gauss' law for the magnetic field)

$$\vec{\nabla} \cdot \vec{B} \equiv 0 \quad (6.4.4)$$

This is very easy to prove. Simply take the divergence¹

$$\begin{aligned} \vec{\nabla} \cdot \vec{B} &= \frac{\mu_0}{2\pi} \int_V d^3\vec{r}' \vec{\nabla} \cdot \frac{\vec{j}(\vec{r}') \times (\vec{r} - \vec{r}')}{|\vec{r} - \vec{r}'|^3} \\ &= -\frac{\mu_0}{2\pi} \int_V d^3\vec{r}' \vec{j}(\vec{r}') \cdot \vec{\nabla} \times \frac{\vec{r} - \vec{r}'}{|\vec{r} - \vec{r}'|^3} \equiv 0 \end{aligned} \quad (6.4.5)$$

because the last integral is seen to be the curl of the gradient of a potential.²

A second consequence is difficult to prove, but it may be done quite rigorously: let C be a closed curve through which there is a non-vanishing and *steady* current, i , then

$$\oint_C \vec{B} \cdot d\vec{r} = \mu_0 i \quad (6.4.6)$$

This is Ampère's law. It may be put into differential form quite easily by exploiting Stoke's theorem

$$\begin{aligned} \oint_C \vec{B} \cdot d\vec{r} &= \int_S \vec{\nabla} \times \vec{B} \cdot d\vec{S} = \mu_0 i = \mu_0 \int_S \vec{j} \cdot d\vec{S} \\ \Rightarrow \quad \vec{\nabla} \times \vec{B} &= \mu_0 \vec{j} \end{aligned} \quad (6.4.7)$$

To summarize then we have four equations that seem to define the electromagnetic field in steady state, *viz.*,

$$\begin{aligned} \vec{\nabla} \cdot \vec{E} &= \frac{\rho}{\epsilon_0} \\ \vec{\nabla} \times \vec{E} &= 0 \end{aligned}$$

¹Also make use of the identity (for any two vectors \vec{a} and \vec{b}): $\vec{\nabla} \cdot (\vec{a} \times \vec{b}) = (\vec{\nabla} \times \vec{a}) \cdot \vec{b} - \vec{a} \cdot (\vec{\nabla} \times \vec{b})$, taking $\vec{a} \equiv \vec{j}$ and $\vec{b} \equiv (\vec{r} - \vec{r}')/|\vec{r} - \vec{r}'|^3$.

²Notice that

$$\frac{(\vec{r} - \vec{r}')}{|\vec{r} - \vec{r}'|^3} = -\vec{\nabla} \phi, \quad \phi = \frac{1}{|\vec{r} - \vec{r}'|}$$

$$\begin{aligned}\vec{\nabla} \cdot \vec{B} &= 0 \\ \vec{\nabla} \times \vec{B} &= \mu_0 \vec{j}\end{aligned}\tag{6.4.8}$$

Of these, the second and fourth will be modified in situations that are not steady. The second equation was modified by Faraday through careful experimentation and the last by Maxwell through purely theoretical arguments involving consistency. It is worth noting however that the first and third equations are unaffected and therefore final. The third equation implies that the magnetic field may be expressed as the rotation of another vector field, which we will call the magnetic vector potential, as

$$\vec{B} = \vec{\nabla} \times \vec{A}\tag{6.4.9}$$

The magnetic vector potential is the magnetic analogue of the electric scalar potential. One can obtain an expression for the magnetic vector potential in terms of the current density by exploiting the expression for the magnetic field in (6.4.3):

$$\begin{aligned}B(\vec{r}) &= \frac{\mu_0}{2\pi} \int_V d^3\vec{r}' \frac{\vec{j}(\vec{r}') \times (\vec{r} - \vec{r}')}{|\vec{r} - \vec{r}'|^3} = -\frac{\mu_0}{2\pi} \int_V d^3\vec{r}' \vec{j}(\vec{r}') \times \vec{\nabla} \frac{1}{|\vec{r} - \vec{r}'|} \\ &= \frac{\mu_0}{2\pi} \vec{\nabla} \times \int_V d^3\vec{r}' \frac{\vec{j}(\vec{r}')}{|\vec{r} - \vec{r}'|}\end{aligned}\tag{6.4.10}$$

where we have used $\vec{\nabla} \times (\phi \vec{a}) = \vec{\nabla} \phi \times \vec{a} + \phi \vec{\nabla} \times \vec{a}$. It follows from its definition that the most general form of the vector potential is

$$\vec{A}(\vec{r}) = \frac{\mu_0}{2\pi} \int_V d^3\vec{r}' \frac{\vec{j}(\vec{r}')}{|\vec{r} - \vec{r}'|} + \vec{\nabla} \Psi,\tag{6.4.11}$$

where $\Psi(\vec{r})$ is an arbitrary scalar field. The freedom to add the gradient of an arbitrary scalar field is a consequence of the fact that the rotation of a gradient vanishes, so the second term gives no contribution to physical magnetic field. This implies an arbitrariness in the definition of the magnetic field since for a given \vec{B} the vector potential can be freely transformed according to

$$\vec{A} \rightarrow \vec{A} + \vec{\nabla} \Psi\tag{6.4.12}$$

This is called a *gauge transformation*. The vector potential satisfies the equation

$$\vec{\nabla} \times \vec{B} = -\vec{\nabla}^2 \vec{A} + \vec{\nabla}(\vec{\nabla} \cdot \vec{A}) = \mu_0 \vec{j}\tag{6.4.13}$$

Taking the divergence of (6.4.11) gives

$$\vec{\nabla} \cdot \vec{A} = \frac{\mu_0}{2\pi} \int_V d^3\vec{r}' \vec{j}(\vec{r}') \cdot \vec{\nabla} \frac{1}{|\vec{r} - \vec{r}'|} + \vec{\nabla}^2 \Psi$$

$$\begin{aligned}
&= -\frac{\mu_0}{2\pi} \int_V d^3\vec{r}' \vec{j}(\vec{r}') \cdot \vec{\nabla}' \frac{1}{|\vec{r} - \vec{r}'|} + \vec{\nabla}^2 \Psi \\
&= -\frac{\mu_0}{2\pi} \int_V d^3\vec{r}' \vec{\nabla}' \cdot \frac{\vec{j}(\vec{r}')}{|\vec{r} - \vec{r}'|} + \frac{\mu_0}{2\pi} \int_V d^3\vec{r}' \frac{\vec{\nabla}' \cdot \vec{j}(\vec{r}')}{|\vec{r} - \vec{r}'|} + \vec{\nabla}^2 \Psi \\
&= -\frac{\mu_0}{2\pi} \oint_S dS \frac{\vec{j}(\vec{r}') \cdot \hat{n}}{|\vec{r} - \vec{r}'|} + \frac{\mu_0}{2\pi} \int_V d^3\vec{r}' \frac{\vec{\nabla}' \cdot \vec{j}(\vec{r}')}{|\vec{r} - \vec{r}'|} + \vec{\nabla}^2 \Psi. \tag{6.4.14}
\end{aligned}$$

Now the second term vanishes by virtue of the fact that $\vec{\nabla}' \cdot \vec{j}(\vec{r}') = 0$. If further there are no currents on the bounding surface, S , we find that any choice of Ψ that satisfies Laplace's equation will ensure

$$\vec{\nabla} \cdot \vec{A} = 0. \tag{6.4.15}$$

On the other hand, if the first term does not vanish we let

$$\vec{\nabla}^2 \Psi = \frac{\mu_0}{2\pi} \oint_S dS \frac{\vec{j}(\vec{r}') \cdot \hat{n}}{|\vec{r} - \vec{r}'|}, \tag{6.4.16}$$

again giving $\vec{\nabla} \cdot \vec{A} = 0$.³ Therefore the vector potential may quite generally be taken to satisfy

$$\vec{\nabla}^2 \vec{A} = -\mu_0 \vec{j}. \tag{6.4.17}$$

It is clear that the solution presented in (6.4.11) is a particular solution of the magnetostatic problem in the same sense as (2.5.4) is a particular solution of the electrostatic problem.

Finally, in regions where $\vec{j}(\vec{r}) = 0$,

$$\vec{B} = -\mu_0 \vec{\nabla} \Phi^*. \tag{6.4.18}$$

solves $\vec{\nabla} \times \vec{B} = 0$. The scalar function $\Phi^*(\vec{r})$ is called the magnetic scalar potential. Applying $\vec{\nabla} \cdot \vec{B} = 0$ to (6.4.18) shows that the magnetic scalar potential satisfies Laplace's equation

$$\vec{\nabla}^2 \Phi^* = 0 \tag{6.4.19}$$

which is the exact analogue of vacuum electrostatics. Therefore much of the work we have done in electrostatics is directly applicable here. However, care must be taken with the boundary conditions as we will shortly see.

³The choice $\vec{\nabla} \cdot \vec{A} = 0$ is called the Coulomb gauge.

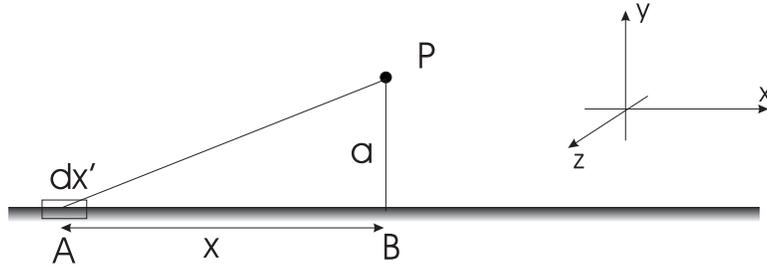


Figure 6.5: An infinite current carrying conductor

6.5 Simple Applications

6.5.1 An infinitely long current carrying wire

Suppose we desire the magnetic field close to a current carrying conductor. By “close” we mean that the perpendicular distance from the conductor is small compared with the length of the conductor itself. In this case we may consider the conductor to be effectively an infinite line of current. Let the point P , at which we desire the field be a perpendicular distance a from the the line of current as shown in figure (6.5) If \vec{r} represents the vector \overrightarrow{BP} , and if \vec{r}' is the vector \overrightarrow{BA} , then $\vec{r} - \vec{r}'$ is the vector from the element $d\vec{r}'$ of the conductor to P and we can write (two dimensions suffice)

$$\vec{r} = (0, a), \quad \vec{r}' = (x, 0), \quad \vec{r} - \vec{r}' = (-x, a) \quad (6.5.1)$$

so that the Biot-Savart law (6.4.2) takes the form

$$\vec{B}(\vec{r}) = \frac{\mu_0 i}{4\pi} \int_{-\infty}^{\infty} \frac{a dx}{(x^2 + a^2)^{3/2}} \hat{z} \quad (6.5.2)$$

The integral is solved by the substitution $x = a \tan \eta$; we find

$$\vec{B} = \frac{\mu_0 i}{2\pi a} \hat{z} \quad (6.5.3)$$

In fact, since our choice of $x - y$ plane is completely arbitrary it should be clear that

$$\vec{B} = -\frac{\mu_0 i}{2\pi a} \hat{\theta} \quad (6.5.4)$$

in spherical coordinates, where $\hat{\theta}$ is the unit vector in the direction of increasing polar angle.

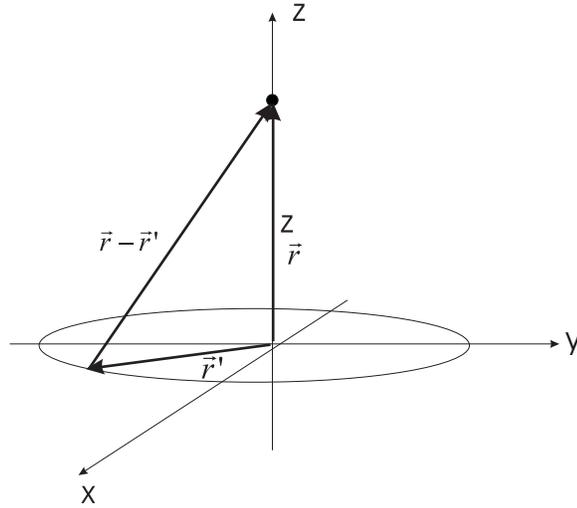


Figure 6.6: A circular loop

6.5.2 A current loop

Consider a current loop of radius a and a point, situated on its axis of symmetry at a distance z from its plane as shown in figure (6.6). Imagine that the current is flowing in a counter-clockwise fashion around the loop and that we are interested in the magnetic field at the point P . Writing,

$$\begin{aligned}
 \vec{r} &= (0, 0, z) \\
 \vec{r}' &= (a \cos \varphi, a \sin \varphi, 0) \\
 \vec{r} - \vec{r}' &= (-a \cos \varphi, -a \sin \varphi, z) \\
 d\vec{r}' &= (-a \sin \varphi, a \cos \varphi, 0) d\varphi \\
 d\vec{r}' \times (\vec{r} - \vec{r}') &= (az \cos \varphi, az \sin \varphi, a^2)
 \end{aligned} \tag{6.5.5}$$

gives

$$\vec{B} = \frac{\mu_0 i}{4\pi} \oint_C \frac{(az \cos \varphi, az \sin \varphi, a^2)}{(a^2 + z^2)^{3/2}} d\varphi \tag{6.5.6}$$

The first two components integrate to zero over a closed loop and the last component gives 2π when integrated, so that

$$\vec{B} = \frac{\mu_0 i}{2} \frac{a^2 \hat{z}}{(a^2 + z^2)^{3/2}} \tag{6.5.7}$$

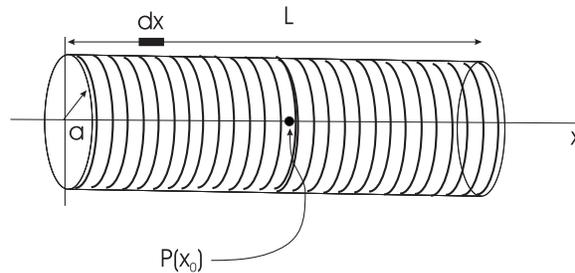


Figure 6.7: A solenoid

6.5.3 A solenoid

The field on the axis of a solenoid is produced by a number of loops that are tightly packed. We imagine that the loops are of negligible thickness (*i.e.*, the thickness of the loop is small compared with the length of the solenoid). Let L be its length and let N be the total number of loops so that $n = N/L$ is the (linear) density of loops in the solenoid. This means that there are ndx loops in a length dx of the solenoid. If i is the current flowing through the solenoid, the current flowing in a length dx will therefore be $nidx$ and will produce a field

$$\vec{B}(x_0) = \frac{\mu_0 n i}{2} \int_0^L \frac{a^2 \hat{x} dx}{(a^2 + (x - x_0)^2)^{3/2}} \quad (6.5.8)$$

where we have used the result of the previous calculation, replacing $i \rightarrow nidx$. The integral is easy to solve⁴, and one gets

$$\vec{B}(x_0) = \frac{\mu_0 n i}{2} (\sin \theta_2 - \sin \theta_1) = \frac{\mu_0 n i}{2} \left[\frac{L - x_0}{\sqrt{a^2 + (L - x_0)^2}} + \frac{x_0}{\sqrt{a^2 + x_0^2}} \right] \hat{x} \quad (6.5.9)$$

Observe the symmetry $x_0 \leftrightarrow L - x_0$. In most practical applications, $a \ll L$ and the expression can be expanded as

$$\vec{B}(x_0) \approx \mu_0 n i \left[1 - \frac{a^2}{4x_0^2} - \frac{a^2}{4(L - x_0)^2} + \frac{3a^4}{16x_0^4} + \frac{3a^4}{16(L - x_0)^4} + \dots \right] \hat{x} \quad (6.5.10)$$

⁴Problem: Solve it by making the substitution

$$(x - x_0) = a \tan \theta$$

which gives limits

$$\theta_1 = -\tan^{-1} \left(\frac{x_0}{a} \right), \quad \theta_2 = \tan^{-1} \left(\frac{L - x_0}{a} \right)$$

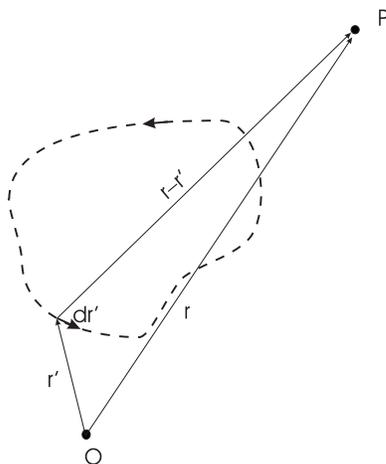


Figure 6.8: A Distant Circuit

If we are not too close to the edges so that $a \ll x_0$ and $a \ll (L - x_0)$, we can neglect all but the first term in square brackets to get the approximation

$$\vec{B}(x_0) \approx \mu_0 n i \hat{x}. \quad (6.5.11)$$

6.5.4 A Distant Circuit

Consider the magnetic field of an arbitrarily shaped circuit situated very far from the point P at which the magnetic field is being observed (in figure 6.8, this means that $|\vec{r}| \gg |\vec{r}'|$ and the typical dimension of the circuit itself.)

We can begin with the expression for $\vec{B}(\vec{r})$ as given by the Biot-Savart law,

$$\vec{B}(\vec{r}) = \frac{\mu_0 i}{4\pi} \oint_C \frac{d\vec{r}' \times (\vec{r} - \vec{r}')}{|\vec{r} - \vec{r}'|^3} \quad (6.5.12)$$

or the expression for the magnetic vector potential

$$\vec{A}(\vec{r}) = \frac{\mu_0 i}{4\pi} \oint_C \frac{d\vec{r}'}{|\vec{r} - \vec{r}'|} \quad (6.5.13)$$

where C represents the circuit. It is better to start with the vector potential. Since we are interested in distant points, consider the expansion

$$|\vec{r} - \vec{r}'| = \sqrt{r^2 + r'^2 - 2\vec{r} \cdot \vec{r}'}$$

$$\begin{aligned} \frac{1}{|\vec{r} - \vec{r}'|} &= r^{-1} \left[1 + \left(\frac{r'}{r} \right)^2 - \frac{2\vec{r} \cdot \vec{r}'}{r^2} \right]^{-1} \\ &= \frac{1}{r} \left[1 + \frac{\vec{r} \cdot \vec{r}'}{r^2} + \mathcal{O} \left(\frac{r'}{r} \right)^2 \right] \end{aligned} \quad (6.5.14)$$

Then inserting the expansion into the integral expression for the magnetic vector potential, we find

$$\vec{A}(\vec{r}) = \frac{\mu_0 i}{4\pi r} \left[\oint_C d\vec{r}' + \frac{1}{r^2} \oint_C d\vec{r}' (\vec{r} \cdot \vec{r}') + \dots \right] \quad (6.5.15)$$

The first integral computed around a closed curve vanishes exactly. The only contribution is from the second integral (up to the present order). We can simplify the integration by noting that

$$(\vec{r}' \times d\vec{r}') \times \vec{r} = (\vec{r}' \cdot \vec{r}) d\vec{r}' - (\vec{r}' \cdot d\vec{r}') \vec{r}' \quad (6.5.16)$$

and further that

$$d[\vec{r}' (\vec{r} \cdot \vec{r}')] = d\vec{r}' (\vec{r} \cdot \vec{r}') + \vec{r}' (\vec{r} \cdot d\vec{r}') \quad (6.5.17)$$

where we have used the fact that \vec{r} is fixed. Adding the two we get

$$(\vec{r}' \times d\vec{r}') \times \vec{r} + d[\vec{r}' (\vec{r} \cdot \vec{r}')] = 2(\vec{r}' \cdot \vec{r}') d\vec{r}' \quad (6.5.18)$$

The right hand side is, of course, what we want to integrate. On the left, the integral of the total derivative around a closed loop gives zero. We conclude that

$$\vec{A}(\vec{r}) \approx \frac{\mu_0 i}{4\pi r^3} \oint_C (\vec{r}' \times d\vec{r}') \times \vec{r} = \frac{\mu_0}{4\pi r^3} (\vec{\mu} \times \vec{r}) \quad (6.5.19)$$

where $\vec{\mu}$ is the magnetic moment of the circuit. The magnetic field \vec{B} is the curl of \vec{A} ,

$$\vec{B}(\vec{r}) \approx \frac{\mu_0}{4\pi} \left[\vec{\mu} \left(\vec{\nabla} \cdot \frac{\vec{r}}{r^3} \right) - (\vec{\mu} \cdot \vec{\nabla}) \frac{\vec{r}}{r^3} \right] \quad (6.5.20)$$

But the first term vanishes and the expression may be simplified

$$\vec{B}(\vec{r}) = \frac{\mu_0}{4\pi} \left[-\frac{\vec{\mu}}{r^3} + \frac{3(\vec{\mu} \cdot \vec{r})\vec{r}}{r^5} \right] \quad (6.5.21)$$

6.6 Faraday's Law

Recall that, for an electric field we have so far two laws, both derived from Coulomb's law, *viz.*,

$$\vec{\nabla} \cdot \vec{D} = \rho \quad (\text{Gauss' Law})$$

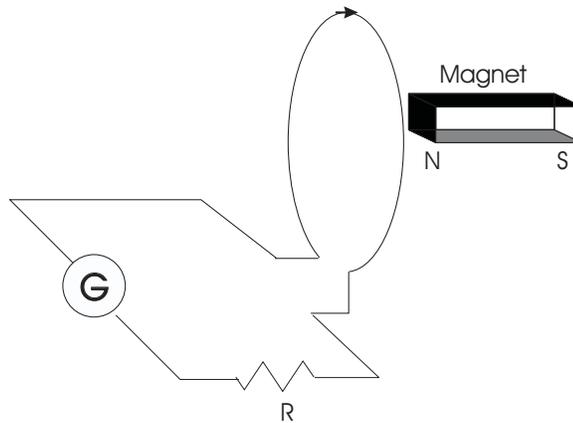


Figure 6.9: Faraday Experiment I

$$\vec{\nabla} \times \vec{E} = 0 \quad (\text{The electrostatic force is conservative}) \quad (6.6.1)$$

Coulomb's law is obtained from data involving only static charges. When charges are in motion, the second law does not generally hold. Writing this law in integral form,

$$\vec{\nabla} \times \vec{E} = 0 \Leftrightarrow \oint_C \vec{E} \cdot d\vec{r} = 0 \quad (6.6.2)$$

and the quantity $\oint \vec{E} \cdot d\vec{r}$ is the net work done in moving a unit charge around a closed loop (circuit). It is called the **electromotive force** (which is a misnomer as it has units of work/unit charge). In the presence of moving charges, Coulomb's law breaks down and we must define the electric field by the Lorentz force law

$$\vec{F} = q(\vec{E} + \vec{v} \times \vec{B}). \quad (6.6.3)$$

Consider the crude version of Faraday's first experiment depicted in Figure (6.9) In the figure, \textcircled{G} represents a galvanometer. As long as the bar magnet is held stationary, no deflection of the galvanometer needle is observed, indicating that no current is flowing through the loop. When the magnet is moved so that there is a changing magnetic flux across the surface enclosed by the loop a current is induced. The current is such that

- It changes direction as the current is moved “into” the loop and “out” of the loop, indicating that the current is sensitive to whether the flux is increasing or decreasing.
- It changes direction if the north pole is substituted for the south pole and vice versa, indicating that it is sensitive to the orientation of the magnetic field.

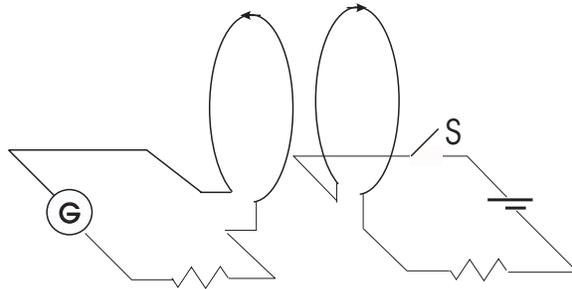


Figure 6.10: Faraday Experiment II

- There is no induced current when the magnet is held still.

Apparently, a changing magnetic flux through a current loop induces a current through it. Next, consider the set-up in figure (6.10)

- When switch S is in the off position, no current flows through either current loop.
- When switch S is turned on a current is produced in the loop on the right and a current is induced in the loop on the left in the opposite direction.
- If the switch is left on, no current is induced in the loop on the left once the current in the loop on the right has reached steady state.
- When the switch is turned off, a current is once again induced in the loop on the left, but in the opposite direction.

We can think of the right loop as a source for a magnetic field. As the switch is turned on the current steadily rises, causing the magnetic field to also increase in intensity. There is therefore a varying magnetic flux through the left loop which generates a current in it in such a direction as to decrease the net flux through the loop. When the switch is turned off the current in the right loop steadily decreases and so does the magnetic field and flux through the left loop. A current is induced in the left loop, again in such a direction as would hold the flux through it steady.

The results of a large number of experiments point to the fact that when there is a changing magnetic flux through a current loop, the electromotive force (e.m.f.) across the loop is not zero but depends on the rate of change of the flux through it. Specifically, if Φ_B is the magnetic flux through the loop then

$$\oint_C \vec{E} \cdot d\vec{r} = \mathcal{E} = -\frac{d\Phi_B}{dt} = -\frac{d}{dt} \int_S \vec{B} \cdot d\vec{S} \quad (6.6.4)$$

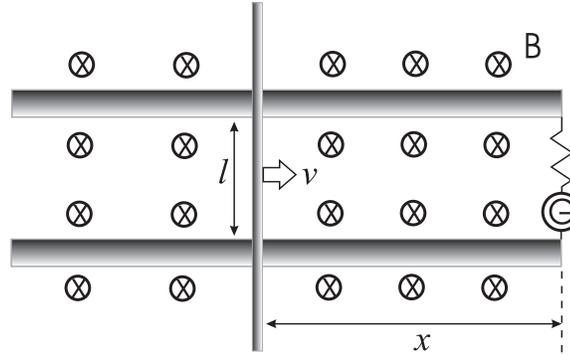


Figure 6.11: Sliding rod

We can use Stoke's theorem to write this equation as

$$\int_S (\vec{\nabla} \times \vec{E}) \cdot d\vec{S} = - \int_S \frac{\partial \vec{B}}{\partial t} \cdot d\vec{S} \quad (6.6.5)$$

or

$$\vec{\nabla} \times \vec{E} + \frac{\partial \vec{B}}{\partial t} = 0 \quad (6.6.6)$$

which is a modification of the second equation at the beginning of this section and is known as Faraday's Law of Induction. The negative sign implies that the induced e.m.f. is such as to generate a current which would oppose the change in magnetic flux. This statement is encapsulated in the so-called **Lenz's law**, which states: "Any change in a magnetic system will cause things to happen that **oppose** the change."

As an example of how this works, consider the simple example shown in figure (6.11) in which the circuit is closed by a rod that is able to slide freely along the rails. The magnetic field is constant and points into the plane of the page. As long as the rod is stationary, no current is detected in the galvanometer. If the rod has a velocity v then

$$\Phi_B = \oint_S \vec{B} \cdot d\vec{S} = Blx \quad (6.6.7)$$

assuming that the normal to S is taken into the plane of the paper. Thus $\mathcal{E} = -d\Phi_B/dt = -Blv$ is rate at which the flux through the current loop is changing. The induced current in the loop is

$$i = \frac{Blv}{R} \quad (6.6.8)$$

where R is the loop resistance. Lenz's law gives its direction as clockwise if v is to the right (decreasing flux) and counterclockwise if v is to the left (increasing flux).⁵

6.7 Inductance

We have seen how circuits can “talk” to each other, *i.e.*, be linked without actually being in contact via each other's magnetic fields. Thus changes in one circuit induce changes in its neighbor. This is the phenomenon of “Inductance”. This linking two circuits depends on the currents flowing through them as well as upon the geometry of the circuits. Assuming the circuits to be rigid, *i.e.*, the geometry stays time independent, the change in magnetic flux depends only on the currents. Label the circuits by $1, 2, i, j, \dots, N$ and let Φ_{ij} represent the flux through circuit i due to the current in circuit j . The net magnetic flux through i due to all circuits in the system will be

$$\Phi_i = \sum_{j=1}^N \Phi_{ij} \quad (6.7.1)$$

where the sum on the right **does not exclude** the circuit i itself. Thus

$$\Phi_i = \sum_{j=1}^N \oint_{S_i} \vec{B}_j(\vec{r}_i) \cdot d\vec{S}_i \quad (6.7.2)$$

where $B_j(\vec{r}_i)$ is the magnetic field due to the current through j at \vec{r}_i . From Faraday's Law we are interested in the rate at which the flux Φ_i changes,

$$\frac{d\Phi_i}{dt} = \sum_{j=1}^N \frac{d\Phi_{ij}}{dt} = \sum_{j=1}^N \frac{d\Phi_{ij}}{di_j} \frac{di_j}{dt} = \sum_j M_{ij} \frac{di_j}{dt} \quad (6.7.3)$$

where i_j is the current through circuit j and $M_{ij} = d\Phi_{ij}/di_j$ is called the **mutual inductance** of the circuits i and j . The **self-inductance** of circuit i is the diagonal element $L_i = M_{ii}$ of M_{ij} and expresses the self-interaction of the current in i with its own magnetic field.

From Faraday's law,

$$\mathcal{E}_i = -\frac{d\Phi_i}{dt} = -\sum_{j=1}^N M_{ij} \frac{di_j}{dt} \quad (6.7.4)$$

Let us look at a simple example.

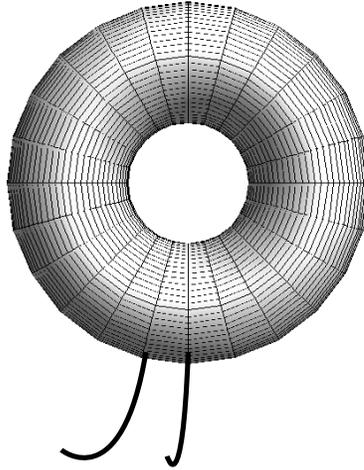


Figure 6.12: A toroidal coil

6.7.1 The toroidal coil

We will first calculate the self inductance of a toroidal coil carrying a current i , as shown in figure (6.12). Using the toroidal symmetry, we can apply Ampère's law to the curve C shown in the figure, passing through the center of the torus,

$$\oint_C \vec{B} \cdot d\vec{r} = \mu_0 N i \quad (6.7.5)$$

By the symmetry, \vec{B} on the curve will have constant magnitude and be directed tangent to it at all points. Applying the right hand rule $\vec{B} = B\hat{\phi}$, and we also have $d\vec{r} = b d\phi \hat{\phi}$, where b is the radius of the curve. Ampère's law gives

$$\vec{B} = \frac{\mu_0 N i}{2\pi b} \hat{\phi} \quad (6.7.6)$$

for N turns of wire. The magnetic flux through the toroidal coil is therefore

$$\Phi_B = \oint_S \vec{B} \cdot d\vec{S} \approx \frac{\mu_0 N^2 i (b-a)^2}{2b} \quad (6.7.7)$$

where we use $d\vec{S} = N dS \hat{\phi} = 2\pi N (b-r) dr$ and approximate \vec{B} by its value along the curve passing through the center of the torus, given in (6.7.6). This approximation is

⁵Convince yourself that these directions for the current generate magnetic fields that oppose the change in flux in each case

good so long as the torus thickness is small compared with its radius, *i.e.*, $b - a \ll b$. The self-inductance of the coil is therefore

$$L = \frac{d\Phi_B}{di} = \frac{\mu_0 N^2 i (b-a)^2}{2b} \quad (6.7.8)$$

Now imagine that the same configuration, but with a coil of N_2 turns carrying a current i_2 wound around it. The magnetic field due to the first coil is given approximately by

$$\vec{B}_1 = \frac{\mu_0 N_1 i_1}{2\pi b} \hat{\varphi} \quad (6.7.9)$$

, giving for the flux through coil 2 the expression

$$\Phi_{21} = \mu_0 N_1 N_2 i_1 \frac{(b-a)^2}{2b} \quad (6.7.10)$$

Likewise,

$$\Phi_{12} = \mu_0 N_1 N_2 i_2 \frac{(b-a)^2}{2b} \quad (6.7.11)$$

and the mutual inductances are equal,

$$M_{21} = \frac{d\Phi_{21}}{di_1} = \mu_0 N_1 N_2 \frac{(b-a)^2}{2b} = \frac{d\Phi_{12}}{di_2} = M_{12} \quad (6.7.12)$$

Comparing the mutual and self inductances, we find an interesting relationship between them:

$$M_{12} = M_{21} = \sqrt{L_1 L_2} \quad (6.7.13)$$

6.7.2 The Neumann Formula

We can get a general formula for the mutual inductance between two circuits of arbitrary shape in a collection of circuits as shown in figure (6.13). From the n circuits, focus on two of them, say i and j . Consider the flux through C_i due to the magnetic field produced by C_j ,

$$\Phi_{ij} = \int_{S_i} \vec{B}_j(\vec{r}_i) \cdot d\vec{S}_i \quad (6.7.14)$$

Using the expression for \vec{B}_j at \vec{r}_i ,

$$\vec{B}_j(\vec{r}_i) = \frac{\mu_0 i_j}{4\pi} \oint_{C_j} d\vec{r}_j \times \frac{(\vec{r}_i - \vec{r}_j)}{|\vec{r}_i - \vec{r}_j|^3} = \vec{\nabla}_i \times \vec{A}_j(\vec{r}_i) \quad (6.7.15)$$

where

$$\vec{A}_j(\vec{r}_i) = \frac{\mu_0 i_j}{4\pi} \oint_{C_j} \frac{d\vec{r}_j}{|\vec{r}_i - \vec{r}_j|} \quad (6.7.16)$$

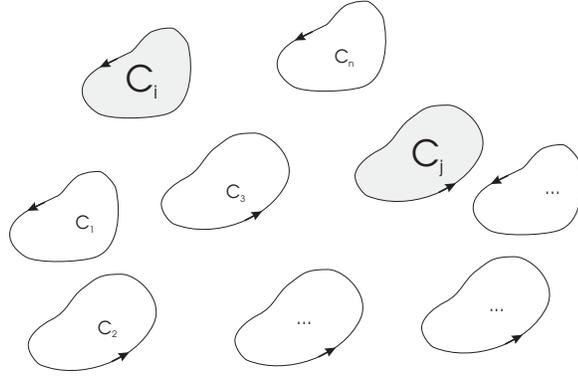


Figure 6.13: Two arbitrary circuits: the Neumann formula

Now we exploit Stoke's theorem,

$$\Phi_{ij} = \int_{S_i} \vec{\nabla}_i \times \vec{A}_j(\vec{r}_i) \cdot d\vec{r}_i = \oint_{C_i} \vec{A}_j(\vec{r}_i) \cdot d\vec{r}_i \quad (6.7.17)$$

which gives

$$\Phi_{ij} = \frac{\mu_0 i_j}{4\pi} \oint_{C_i} \oint_{C_j} \frac{d\vec{r}_i \cdot d\vec{r}_j}{|\vec{r}_i - \vec{r}_j|} \quad (6.7.18)$$

and hence the mutual inductance as

$$M_{ij} = \frac{d\Phi_{ij}}{di_j} = \frac{\mu_0}{4\pi} \oint_{C_i} \oint_{C_j} \frac{d\vec{r}_i \cdot d\vec{r}_j}{|\vec{r}_i - \vec{r}_j|} \quad (6.7.19)$$

This is the Neumann formula for the mutual inductance and shows quite generally that $M_{ij} = M_{ji}$ for any pair of circuits in a collection.

Chapter 7

Magnetic Media

When considering the microscopic structure of materials, one is led inevitably to the atom, which we model roughly as consisting of a central nucleus with electrons in motion about it. Electrons carry charge and their motion leads to electric currents, although there may be no bulk motion of the material or the charge. These currents are due to bound charges that circulate the nuclei, each circulation giving rise to a tiny current loop. Atomic currents can also occur by quantum mechanical spin. When dealing with the magnetic properties of materials, it is therefore convenient to separate whatever current is present into two distinct pieces: a macroscopic piece that is due to a bulk movement of charge in the material and a microscopic piece which averages over the current loops (orbital or spin) associated with each atom. The term “microscopic” here is intended to be such a volume element as is large compared with the typical atomic volume but small compared with any volume element that will be experimentally probed. Thus even such a “microscopic” element of the material is assumed to contain a large enough number of atoms so that the averaging process referred to above is meaningful. An approximation which exploits such an averaging is called the “continuum approximation” and therefore largely ignores the microscopic origin of phenomena.

We will concentrate on the microscopic currents. We know that current loops yield magnetic fields and if the latter are required only at large distances (compared with the typical size of the loop) from the loops, then each may be approximated by a magnetic dipole [see (6.5.19)]. In most (un-magnetized) materials thermal fluctuations cause the dipoles to orient randomly (see figure (7.1) with respect to one another, so that contributions of the individual dipoles to the bulk magnetic field cancel. However, if the current loops are, for some reason, aligned so as to add coherently then there will be a bulk field and the material is said to be magnetized. This can occur in several ways. In **Paramagnetic** materials, the torque experienced by the atomic dipoles when an external magnetic field is applied is sufficient to align them and the material develops a bulk magnetic field

of its own. However, in a purely paramagnetic material the atomic dipoles do not interact with one another and the effect disappears when the external magnetic field is turned off. For materials in which the dipoles do interact, they will generally order themselves in some way. When the effective interactions between them are such as to energetically favor alignment, the material is said to be **Ferromagnetic** (a permanent magnet) and for materials in which the interactions favor anti-alignment the material is **Anti-ferromagnetic**. Paramagnetic behavior is also observed in Ferromagnetic and Anti-ferromagnetic materials if the temperature is high enough (the Curie temperature for ferromagnetic materials and the Neel temperature for anti-ferromagnets). In atoms with no dipole moment, a dipole moment may be induced by applying a strong external magnetic field. This occurs because the Lorentz force on the electrons cause them to either speed up in their orbits or slow down. Consider two electrons moving in circular orbits, one clockwise and the other counterclockwise. The “average” dipole moment of our toy system is zero. Imagine that an external magnetic field perpendicular to and into the plane of their motion is switched on. The Lorentz force on the electron moving clockwise will cause it to speed up whereas it will cause the electron moving counterclockwise to slow down. This has the effect of generating a net dipole moment and therefore magnetic field in such a direction as opposes the external field (which is just Lenz’s law at work). When this occurs in materials, it is called **Diamagnetism**. Although all materials show a diamagnetic response, the induced magnetic moment is very small in most everyday materials and can hardly be observed when other forms of magnetic behavior are present.

We will not concern ourselves with these interactions here, but consider instead only the effect of having a magnetization.

7.1 Magnetization

Let us divide the entire volume of the material into macroscopically small but microscopically large pieces, in the sense indicated above. Let the volume of each piece be Δv and suppose that the magnetic dipole moment of the i^{th} atom in this volume element is \vec{m}_i .

Definition: The magnetization of the material is the magnetic dipole moment per unit volume of the material, *i.e.*,

$$\vec{M} = \lim_{\Delta v \rightarrow 0} \frac{1}{\Delta v} \sum_{i \in \Delta v} \vec{m}_i \quad (7.1.1)$$

We shall assume that \vec{M} is experimentally known and compute the material’s contribution to the bulk magnetic field. In the unmagnetized state $\sum_i \vec{m}_i = 0$ and so is \vec{M} as a result of random orientations of the magnetic moments. In the presence of an external magnetic field, torques on the atomic (molecular) current loops will align them and we expect \vec{M} to depend on \vec{B} apart from the properties of the material itself.

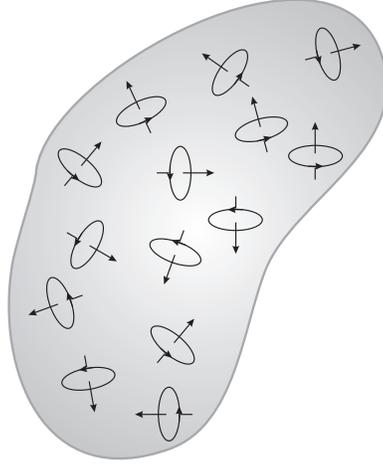


Figure 7.1: Randomly oriented magnetic dipoles in magnetic media

From the definition of the magnetization, the magnetic moment of a volume Δv of the material will be

$$\Delta \vec{m} = \vec{M} \Delta v \quad (7.1.2)$$

Using the expression we had for the magnetic field of a distant dipole, we write

$$\begin{aligned} \vec{A}(\vec{r}) &= \frac{\mu_0}{4\pi} \int_V d^3 \vec{r}' \frac{\vec{M}(\vec{r}') \times (\vec{r} - \vec{r}')}{|\vec{r} - \vec{r}'|^3} \\ &= \frac{\mu_0}{4\pi} \int_V d^3 \vec{r}' \vec{M}(\vec{r}') \times \vec{\nabla}' \left(\frac{1}{|\vec{r} - \vec{r}'|} \right) \end{aligned} \quad (7.1.3)$$

where the integral is over the entire volume of the material. Using a familiar vector identity, we can reexpress this integral as

$$\frac{\mu_0}{4\pi} \int_V d^3 \vec{r}' \left[\vec{\nabla}' \times \frac{\vec{M}(\vec{r}')}{|\vec{r} - \vec{r}'|} - \frac{\vec{\nabla}' \times \vec{M}(\vec{r}')}{|\vec{r} - \vec{r}'|} \right] \quad (7.1.4)$$

If we use the identity (\vec{a} is any vector field)

$$\int_V d^3 \vec{r} \vec{\nabla} \times \vec{a} = \oint_S \vec{a} \times d\vec{S} \quad (7.1.5)$$

on the first term, it transforms the expression for the magnetic vector potential due to the magnetization into

$$\vec{A}(\vec{r}) = \frac{\mu_0}{4\pi} \left[\oint_S \frac{\vec{M}(\vec{r}') \times \hat{n}'}{|\vec{r} - \vec{r}'|} + \int_V d^3 \vec{r}' \frac{\vec{\nabla}' \times \vec{M}(\vec{r}')}{|\vec{r} - \vec{r}'|} \right] \quad (7.1.6)$$

where \hat{n}' is the unit normal to the surface bounding the material. The vector potential is now given as a sum over a bulk term (the second) and a surface term (the first). Both terms have the standard form for the magnetic vector potential if we identify each numerator as a current density, *i.e.*,

$$\begin{aligned}\vec{J}_M(\vec{r}') &= \vec{\nabla}' \times M(\vec{r}') \quad (\text{volume magnetization current density}) \\ \vec{j}_M(\vec{r}') &= \vec{M}(\vec{r}') \times \hat{n}' \quad (\text{surface magnetization current density})\end{aligned}\quad (7.1.7)$$

in terms of which

$$\vec{A}(\vec{r}) = \frac{\mu_0}{4\pi} \left[\oint_S \frac{j_M(\vec{r}')}{|\vec{r} - \vec{r}'|} + \int_V d^3\vec{r}' \frac{J_M(\vec{r}')}{|\vec{r} - \vec{r}'|} \right] \quad (7.1.8)$$

In the end, however, we are interested in the expression for the magnetic field $\vec{B}(\vec{r})$. For this it is more convenient to take the curl of the expression in (7.1.3) for the vector potential

$$\begin{aligned}\vec{B}(\vec{r}) &= \vec{\nabla} \times \vec{A}(\vec{r}) = \frac{\mu_0}{4\pi} \int_V d^3\vec{r}' \vec{\nabla} \times \left[\vec{M}(\vec{r}') \times \frac{(\vec{r} - \vec{r}')}{|\vec{r} - \vec{r}'|^3} \right] \\ &= \frac{\mu_0}{4\pi} \int_V d^3\vec{r}' \left[\vec{M}(\vec{r}') \vec{\nabla} \cdot \frac{(\vec{r} - \vec{r}')}{|\vec{r} - \vec{r}'|^3} - (\vec{M}(\vec{r}') \cdot \vec{\nabla}) \frac{(\vec{r} - \vec{r}')}{|\vec{r} - \vec{r}'|^3} \right]\end{aligned}\quad (7.1.9)$$

We recognize the first term as just μ_0 times the magnetization at \vec{r} , because

$$\vec{\nabla} \cdot \frac{(\vec{r} - \vec{r}')}{|\vec{r} - \vec{r}'|^3} = 4\pi\delta^3(\vec{r} - \vec{r}') \quad (7.1.10)$$

The second term can be put into a more transparent form if we use the vector identity

$$\vec{\nabla}(\vec{a} \cdot \vec{b}) = (\vec{a} \cdot \vec{\nabla})\vec{b} + (\vec{b} \cdot \vec{\nabla})\vec{a} + \vec{a} \times (\vec{\nabla} \times \vec{b}) + \vec{b} \times (\vec{\nabla} \times \vec{a}) \quad (7.1.11)$$

taking \vec{a} to be $\vec{M}(\vec{r}')$ *i.e.*, independent of \vec{r} and $\vec{b} = (\vec{r} - \vec{r}')/|\vec{r} - \vec{r}'|^3$. Then all derivatives of \vec{a} vanish as does the rotation of \vec{b} . The second integral is reduced to

$$- \frac{\mu_0}{4\pi} \vec{\nabla} \int_V d^3\vec{r}' \frac{\vec{M}(\vec{r}') \cdot (\vec{r} - \vec{r}')}{|\vec{r} - \vec{r}'|^3} \quad (7.1.12)$$

which is the divergence of a scalar potential. Putting everything together gives the magnetic field due to a magnetized distribution of matter,

$$\vec{B}(\vec{r}) = \mu_0 \vec{M}(\vec{r}) - \mu_0 \vec{\nabla} \Phi_M^*(\vec{r}) \quad (7.1.13)$$

where we have defined

$$\Phi_M^*(\vec{r}) = \frac{1}{4\pi} \int_V d^3\vec{r}' \frac{\vec{M}(\vec{r}') \cdot (\vec{r} - \vec{r}')}{|\vec{r} - \vec{r}'|^3} \quad (7.1.14)$$

The potential Φ_M^* is the magnetic scalar potential due to the magnetic material: it has an interesting interpretation. Consider

$$\begin{aligned}
 \Phi_M^*(\vec{r}) &= \frac{1}{4\pi} \int_V d^3\vec{r}' \frac{\vec{M}(\vec{r}') \cdot (\vec{r} - \vec{r}')}{|\vec{r} - \vec{r}'|^3} \\
 &= \frac{1}{4\pi} \int_V d^3\vec{r}' \vec{M}(\vec{r}') \cdot \vec{\nabla}' \frac{(\vec{r} - \vec{r}')}{|\vec{r} - \vec{r}'|^3} \\
 &= \frac{1}{4\pi} \int_V d^3\vec{r}' \left[\vec{\nabla}' \cdot \frac{\vec{M}(\vec{r}')}{|\vec{r} - \vec{r}'|} - \frac{\vec{\nabla}' \cdot \vec{M}(\vec{r}')}{|\vec{r} - \vec{r}'|} \right]
 \end{aligned}
 \tag{7.1.15}$$

By Gauss' theorem, the first term can be written as an integral over the bounding surface and the scalar potential may be expressed as

$$\Phi_M^*(\vec{r}) = \frac{1}{4\pi} \left[\oint_S dS \frac{(\vec{M}(\vec{r}') \cdot \hat{n}')}{|\vec{r} - \vec{r}'|} + \int_V d^3\vec{r}' \frac{(-\vec{\nabla}' \cdot \vec{M}(\vec{r}'))}{|\vec{r} - \vec{r}'|} \right].
 \tag{7.1.16}$$

This, of course, is identical to the expression for the electric scalar potential in a dielectric medium if we identify

$$\begin{aligned}
 \rho_M(\vec{r}') &= -\vec{\nabla}' \cdot \vec{M}(\vec{r}') \\
 \sigma_M(\vec{r}') &= \vec{M}(\vec{r}') \cdot \hat{n}'
 \end{aligned}
 \tag{7.1.17}$$

as charges, called (respectively) the volume and surface the “magnetic pole densities”. Thus

$$\Phi_M^*(\vec{r}) = \frac{1}{4\pi} \left[\oint_S dS \frac{\sigma_M(\vec{r}')}{|\vec{r} - \vec{r}'|} + \int_V d^3\vec{r}' \frac{\rho_M(\vec{r}')}{|\vec{r} - \vec{r}'|} \right],
 \tag{7.1.18}$$

which expression is already of a familiar form.

If there is a macroscopic current density also present then we must include its contribution to the total magnetic field at any point. Doing this, we get

$$\vec{B}(\vec{r}) = \frac{\mu_0}{4\pi} \int_V d^3\vec{r}' \frac{j(\vec{r}') \times (\vec{r} - \vec{r}')}{|\vec{r} - \vec{r}'|^3} + \mu_0 \vec{M}(\vec{r}) - \mu_0 \vec{\nabla} \Phi_M^*(\vec{r})
 \tag{7.1.19}$$

The current $j(\vec{r})$ includes all transport currents that may be present in the system. The effect of the atomic (molecular) currents is found in the magnetization vector which appears in the last two terms. Thus, we must only have a knowledge of $j(\vec{r})$ and $M(\vec{r})$ everywhere in order to obtain $\vec{B}(\vec{r})$ anywhere. The problem is, as said earlier, that while

$\vec{j}(\vec{r})$ is easily determined or controlled, $\vec{M}(\vec{r}')$ depends also on $\vec{B}(\vec{r})$. We get around this by a field redefinition: let

$$\vec{H}(\vec{r}) = \frac{1}{\mu_0} \vec{B}(\vec{r}) - \vec{M}(\vec{r}) \quad (7.1.20)$$

then

$$\vec{H}(\vec{r}) = \frac{1}{4\pi} \int_V d^3\vec{r}' \frac{\vec{j}(\vec{r}') \times (\vec{r} - \vec{r}')}{|\vec{r} - \vec{r}'|^3} - \vec{\nabla} \Phi_M^*(\vec{r}) \quad (7.1.21)$$

Superficially it appears that we have gone nowhere, since \vec{M} still appears on the right hand side through $\Phi_M^*(\vec{r})$. However, the rotation of a gradient being zero, we have

$$\vec{\nabla} \times \vec{H}(\vec{r}) = \frac{1}{4\pi} \int_V d^3\vec{r}' \vec{\nabla} \times \frac{\vec{j}(\vec{r}') \times (\vec{r} - \vec{r}')}{|\vec{r} - \vec{r}'|^3} \quad (7.1.22)$$

which can also be written as

$$\vec{\nabla} \times \vec{H}(\vec{r}) = \frac{1}{4\pi} \int_V d^3\vec{r}' \left[\vec{j}(\vec{r}') \vec{\nabla} \cdot \frac{(\vec{r} - \vec{r}')}{|\vec{r} - \vec{r}'|^3} - (\vec{j}(\vec{r}') \cdot \vec{\nabla}) \frac{(\vec{r} - \vec{r}')}{|\vec{r} - \vec{r}'|^3} \right] \quad (7.1.23)$$

The first integral as before simply yields $\mu_0 \vec{j}(\vec{r})$. The second may be put into the form $-\vec{\nabla} \Phi_j^*(\vec{r})$, where¹

$$\begin{aligned} \Phi_j^*(\vec{r}) &= \frac{1}{4\pi} \int_V d^3\vec{r}' \frac{\vec{j}(\vec{r}') \cdot (\vec{r} - \vec{r}')}{|\vec{r} - \vec{r}'|^3} \\ &= \frac{1}{4\pi} \left[\int_V d^3\vec{r}' \frac{(-\vec{\nabla}' \cdot \vec{j}(\vec{r}'))}{|\vec{r} - \vec{r}'|} + \oint_S dS \frac{(\vec{j}(\vec{r}') \cdot \hat{n}')}{|\vec{r} - \vec{r}'|} \right] \end{aligned} \quad (7.1.24)$$

For steady flows $\vec{\nabla}' \cdot \vec{j}(\vec{r}') \equiv 0$, so the first term vanishes. The surface term can be made to vanish also if we take the bounding surface to be at infinity. This does not contradict the possibility that the transport currents may be confined to a small region, since we may simply take $\vec{j}(\vec{r}')$ to be vanishing outside that region. In any case, $\vec{j}(\vec{r}')$ is required to fall off sufficiently rapidly at infinity so that the second term vanishes.

Finally, we have

$$\begin{aligned} \vec{H}(\vec{r}) &= \frac{\vec{B}(\vec{r})}{\mu_0} - \vec{M}(\vec{r}) \\ \vec{\nabla} \times \vec{H}(\vec{r}) &= \vec{j}(\vec{r}) \quad (\text{for steady flows}) \end{aligned} \quad (7.1.25)$$

The equation $\vec{\nabla} \cdot \vec{B} = 0$ continues to hold, since it simply claims that there are no magnetic monopoles. This is easy to show by taking the divergence of the right hand side of the

¹Problem: Follow the steps after (7.1.9) to obtain the given result.

equation for $\vec{B}(\vec{r})$ in (7.1.19). The divergence of the first term vanishes as we have already shown, and we are left with

$$\vec{\nabla} \cdot \vec{B}(\vec{r}) = \mu_0(\vec{\nabla} \cdot \vec{M}(\vec{r}) - \vec{\nabla}^2 \Phi_M^*(\vec{r})) \quad (7.1.26)$$

This too vanishes. To show that it does, begin with the expression for $\Phi_M^*(\vec{r})$,

$$\begin{aligned} \Phi_M^*(\vec{r}) &= \frac{1}{4\pi} \left[\oint_S dS \frac{\sigma_M(\vec{r}')}{|\vec{r} - \vec{r}'|} + \int_V d^3\vec{r}' \frac{\rho_M(\vec{r}')}{|\vec{r} - \vec{r}'|} \right], \\ \Rightarrow \vec{\nabla}^2 \Phi_M^*(\vec{r}) &= \frac{1}{4\pi} \left[\oint_S dS \sigma_M(\vec{r}') \vec{\nabla}^2 \frac{1}{|\vec{r} - \vec{r}'|} + \int_V d^3\vec{r}' \rho_M(\vec{r}') \vec{\nabla}^2 \frac{1}{|\vec{r} - \vec{r}'|} \right] \end{aligned} \quad (7.1.27)$$

and taking the bounding surface out to infinity, we get simply

$$\vec{\nabla}^2 \Phi_M^*(\vec{r}) = -\rho_M(\vec{r}) = \vec{\nabla} \cdot \vec{M}(\vec{r}) \quad (7.1.28)$$

Let us end this section by writing out the electromagnetic field equations for steady flows:

$$\begin{aligned} \vec{\nabla} \cdot \vec{D} &= \rho_f \\ \vec{\nabla} \times \vec{E} + \frac{\partial \vec{B}}{\partial t} &= 0 \\ \vec{\nabla} \cdot \vec{B} &= 0 \\ \vec{\nabla} \times \vec{H} &= \vec{j}_f \end{aligned} \quad (7.1.29)$$

where j_f is the transport current, and

$$\begin{aligned} \vec{D} &= \epsilon_0 \vec{E} - \vec{P} \quad (\text{the electric displacement vector}) \\ \vec{H} &= \frac{\vec{B}}{\mu_0} - \vec{M} \quad (\text{the magnetic intensity vector}) \end{aligned} \quad (7.1.30)$$

Notice

- the symmetry in the equations,
- that only the first and last equations are sensitive to the presence of sources,
- that the middle two equations are written in terms of the vacuum fields \vec{E} and \vec{B} and have the same form, no matter what the sources.

The middle two equations are really constraints. They tell us that the electric and magnetic fields are not entirely independent of one another and are in fact manifestations of a single field, the electromagnetic field, which consists of a single vector potential and a single scalar potential. These equations can be understood more generally as geometric identities *viz.*, the so-called Bianchi identities of gauge theories.

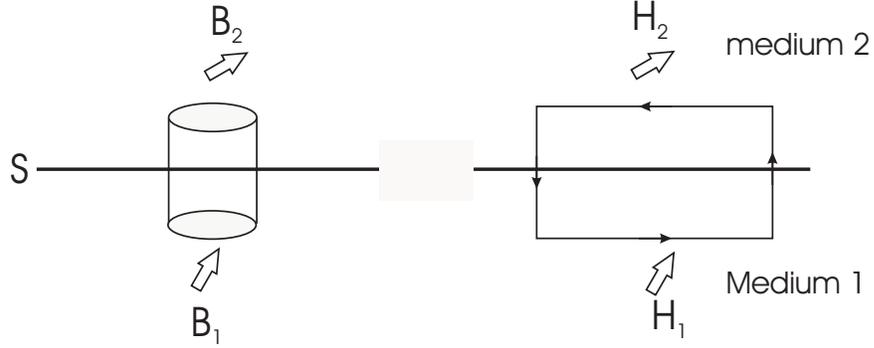


Figure 7.2: Boundary conditions at the interface of two magnetic media

7.2 Boundary Conditions

Consider two magnetic media of different responses in contact with each other and whose interface is denoted by the surface S in figure (7.2). As before, begin with the pillbox on the left, whose cylindrical surface can be made as small as we desire. Then considering only the flat surfaces of the pillbox, which respectively in medium 1 and medium 2. By Gauss' law for the magnetic field

$$\int_{S_1+S_2} \vec{B} \cdot d\vec{S} = 0 \Rightarrow B_{1n} = B_{2n} \quad (7.2.1)$$

i.e., the normal component of the magnetic field is continuous.

Turning to the rectangular closed curve on the right of figure (7.2) we have

$$\oint_C \vec{H} \cdot d\vec{r} = \int_S (\vec{\nabla} \times \vec{H}) \cdot d\vec{S} = \int_S \vec{j}_f \cdot d\vec{S} \quad (7.2.2)$$

and, referring to the figure,

$$\Delta l (\vec{H}_1 - \vec{H}_2) \cdot \hat{t} = \mathfrak{J}_f \cdot (\hat{t} \times \hat{n}) \Delta l \Delta A \quad (7.2.3)$$

where we have introduced the current density per unit length of the surface, \mathfrak{J}_f , hence the extra factor of Δl on the right. The factor A is just the area of the rectangular closed curve and we have exploited the fact that $\hat{t} \times \hat{n}$ gives the normal to the surface bounded by the rectangular closed curve. Call $\mathfrak{J}_f \Delta A = \vec{l}_f$, *i.e.*, the transport current per unit length along the surface, then using the symmetries of the triple product on the right the above equation can be put in the form

$$(\vec{H}_1 - \vec{H}_2) = \hat{n} \times \vec{l}_f \quad (7.2.4)$$

Thus there is a discontinuity of the horizontal component of the magnetic intensity vector if surface currents are present at the interface. If not, the tangential components of the magnetic intensity vector are also continuous across the interface (as are the normal components of \vec{B}).

We have seen that the magnetization vector \vec{M} generally depends on the external applied field \vec{B} , but it is convenient to write a phenomenological relationship of the form

$$\vec{M}_i(\vec{H}) = \sum_j \chi_{ij}^m(\vec{H}) H_j \quad (7.2.5)$$

instead of a relationship directly in terms of the magnetic field, \vec{B} . A magnetic medium is **isotropic** if

$$\chi_{ij}^m(\vec{B}) = \chi^m(\vec{H}) \delta_{ij} \quad (7.2.6)$$

in which case the magnetic intensity and the magnetic field vectors are oriented in the same direction. $\chi^m(\vec{B})$ is the *magnetic susceptibility* of the medium. We then have

$$\vec{H} = \frac{\vec{B}}{\mu_0} - \vec{M} = \frac{\vec{B}}{\mu_0} - \chi^m(\vec{H}) \vec{H} \Rightarrow \vec{B} = \mu_0 [1 + \chi^m(\vec{H})] \vec{H} \quad (7.2.7)$$

Define the *magnetic permeability* of the material by $\vec{B} = \mu(\vec{H}) \vec{H}$, *i.e.*,

$$\mu(\vec{H}) = \mu_0 [1 + \chi^m(\vec{H})] \Rightarrow \chi^m(\vec{H}) = \frac{\mu(\vec{H})}{\mu_0} - 1 \quad (7.2.8)$$

The quantity $K_m = \chi^m + 1 = \mu/\mu_0$ is called the *relative permeability*. A medium is said to be **linear** if $\chi^m(\vec{H})$ is independent of \vec{H} . We consider only linear, isotropic magnetic media henceforth.

7.3 Examples

Let us now apply the techniques learned in electrostatics to solve problems in magnetostatics.

7.3.1 Magnetic sphere in a uniform magnetic field

Consider a sphere of radius a , made of some linear, isotropic magnetic material and placed in a uniform magnetic field \vec{B} . We wish to find the magnetic field everywhere. As we have seen, where the transport currents vanish identically,

$$\vec{\nabla} \times \vec{H} = 0 \Rightarrow \vec{H} = -\vec{\nabla} \Phi_M^* \quad (7.3.1)$$

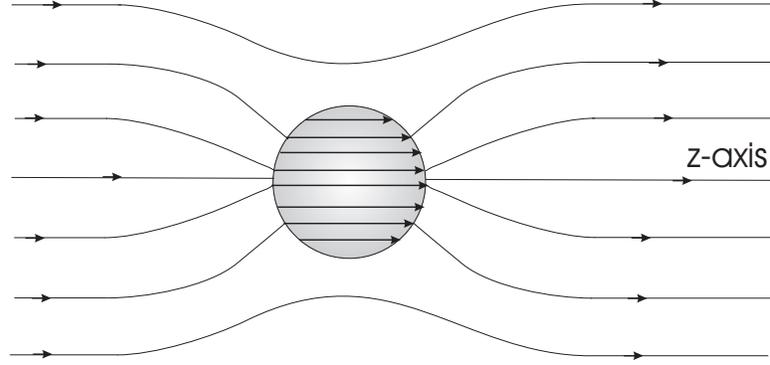


Figure 7.3: Magnetic field inside and out of a magnetic medium

and because the material is isotropic we expect that

$$\vec{H} = \frac{\vec{B}}{\mu} \quad (7.3.2)$$

everywhere. In particular this means that $\vec{B} = -\mu\vec{\nabla}\Phi_M^* \Rightarrow \vec{\nabla}^2\Phi_M^* = 0$. Let us take the uniform magnetic field to be oriented in the z direction so that $\vec{B} = B_0\hat{z}$. The problem is therefore to find a solution to Laplace's equation in spherical coordinates and subject to

$$\vec{B}_\infty = \lim_{r \rightarrow \infty} \vec{B}(\vec{r}) = B_0\hat{z} \Rightarrow \Phi_M^* = -B_0z = -B_0r \cos\theta + \text{const.} \quad (7.3.3)$$

The general solution of Laplace's equation in the exterior of the sphere may be written in the form

$$\Phi_{M,\text{ext}}^* = A + \frac{B}{r} + \left(C_1r + \frac{C_2}{r^2}\right) \cos\theta + \left(D_1r^2 + \frac{D_2}{r^3}\right) (3\cos^2\theta - 1) + \dots \quad (7.3.4)$$

At infinity, all terms in r^{-n} , $n \geq 1$ are vanishing, so

$$\Phi_{M,\text{ext}}^*(r \rightarrow \infty) = A + C_1r \cos\theta + D_1r^2(3\cos^2\theta - 1) + \dots \quad (7.3.5)$$

but, because of the behavior of Φ_M^* at infinity, we see that $A = 0 = D_1 = E_1 = \dots$ and

$$\Phi_{M,\text{ext}}^*(r, \theta) = \frac{B}{r} - \left(\frac{B_0}{\mu_0}r + \frac{C_2}{r^2}\right) \cos\theta + \frac{D_2}{r^3}(3\cos^2\theta - 1) + \dots \quad (7.3.6)$$

Again, since there are no transport currents in the interior, we must likewise have the general solution

$$\Phi_{M,\text{int}}^*(r, \theta) = \bar{A} + \frac{\bar{B}}{r} + \left(\bar{C}_1r + \frac{\bar{C}_2}{r^2}\right) \cos\theta + \left(\bar{D}_1r^2 + \frac{\bar{D}_2}{r^3}\right) (3\cos^2\theta - 1) + \dots \quad (7.3.7)$$

but regularity at the center requires $\bar{B} = 0 = \bar{C}_2 = \bar{D}_2 = \dots$ and therefore

$$\Phi_{M,\text{int}}^*(r, \theta) = \bar{A} + \bar{C}_1 r \cos \theta + \bar{D}_1 r^2 (3 \cos^2 \theta - 1) + \dots \quad (7.3.8)$$

The boundary conditions at the surface of the sphere tell us that the normal component of \vec{B} and the tangential components of \vec{H} are continuous there. The normal component of \vec{B} is computed from $\mu(\hat{r} \cdot \vec{\nabla} \Phi_M^*)$ in each region, thus

$$\mu_0 \frac{\partial}{\partial r} \Phi_{M,\text{ext}}^* \Big|_{r=a} = \mu \frac{\partial}{\partial r} \Phi_{M,\text{int}}^* \Big|_{r=a} \quad (7.3.9)$$

This gives the following relationships (comparing terms with the same associated Legendre polynomials),

$$\begin{aligned} B &= 0 \\ \mu_0 \left(\frac{B_0}{\mu_0} - \frac{2C_2}{a^3} \right) &= \mu \bar{C}_1 \\ -\frac{3\mu_0 D_2}{a^4} &= 2\mu \bar{D}_1 a \\ \dots &\dots \end{aligned} \quad (7.3.10)$$

Continuity of the potential across the surface gives

$$\begin{aligned} \frac{B}{a} &= \bar{A} \\ -\frac{B_0}{\mu_0} a + \frac{C_2}{a^2} &= \bar{C}_1 a \\ \frac{D_2}{a^3} &= \bar{D}_1 a^2 \\ \dots &\dots \end{aligned} \quad (7.3.11)$$

These are compatible with the first set only if

$$\begin{aligned} \bar{A} &= 0 \\ D_2 &= \bar{D}_1 = 0 \\ E_2 &= \bar{E}_1 \\ \dots &\dots \end{aligned} \quad (7.3.12)$$

and the two equations

$$K_m \bar{C}_1 = -\frac{B_0}{\mu_0} - \frac{2C_2}{a^3}, \quad \bar{C}_1 = -\frac{B_0}{\mu_0} + \frac{C_2}{a^3} \quad (7.3.13)$$

hold simultaneously. This system can be solved for C_2 and \bar{C}_1 . We find

$$\begin{aligned} C_2 &= \frac{B_0 a^3}{\mu_0} \left[\frac{K_m - 1}{K_m + 2} \right] \\ \bar{C}_1 &= -\frac{3B_0}{\mu_0(K_m + 2)} \end{aligned} \quad (7.3.14)$$

Thus, we have found the desired potential everywhere:

$$\begin{aligned} \Phi_{M,\text{ext}}^* &= -\frac{B_0}{\mu_0} r \cos \theta \left(1 - \frac{a^3}{r^3} \left[\frac{K_m - 1}{K_m + 2} \right] \right) \\ \Phi_{M,\text{int}}^* &= -\frac{3B_0}{\mu_0(K_m + 2)} r \cos \theta \end{aligned} \quad (7.3.15)$$

From the potential, we immediately calculate $\vec{B}_{\text{ext/int}}$ as²

$$\begin{aligned} \vec{B}_{\text{ext}} &= -\mu_0 \vec{\nabla} \Phi_{M,\text{ext}}^* = B_0 \hat{z} + B_0 \frac{a^3}{r^3} \left[\frac{K_m - 1}{K_m + 2} \right] (2 \cos \theta \hat{r} + \sin \theta \hat{\theta}) \\ \vec{B}_{\text{int}} &= -\mu \vec{\nabla} \Phi_{M,\text{int}}^* = \frac{3K_m B_0}{(K_m + 2)} \hat{z} \end{aligned} \quad (7.3.16)$$

This field is shown in figure (7.3)

²Problem: Recall that $\vec{B} = -\mu \vec{\nabla} \Phi_M^*$ and, in spherical coordinates:

$$\vec{\nabla} = \hat{r} \frac{\partial}{\partial r} + \frac{\hat{\theta}}{r} \frac{\partial}{\partial \theta} + \frac{\hat{\varphi}}{r \sin \theta} \frac{\partial}{\partial \varphi}$$

Also recall the non-vanishing derivatives of the unit vectors in spherical coordinates:

$$\frac{\partial \hat{r}}{\partial \theta} = \hat{\theta}, \quad \frac{\partial \hat{\theta}}{\partial \theta} = -\hat{r}, \quad \frac{\partial \hat{\varphi}}{\partial \varphi} = -\hat{r} \sin \theta - \hat{\theta} \cos \theta$$

Combining these, show that

$$B_r = -\mu \frac{\partial \Phi_M^*}{\partial r}, \quad B_\theta = \frac{\mu}{r} \frac{\partial \Phi_M^*}{\partial \theta}, \quad B_\varphi = -\frac{\mu}{r \sin \theta} \frac{\partial \Phi_M^*}{\partial \varphi}$$

7.3.2 Uniformly magnetized sphere

Now consider the problem of the magnetic field in all of space due to a uniformly magnetized sphere of radius a and magnetization \vec{M} . Let the magnetization be in the \hat{z} direction, so that $\vec{M} = M\hat{z}$ (M is constant). Once again, let us examine the solution of $\vec{\nabla}^2\Phi^* = 0$, as there are no transport currents in this problem.

In spherical coordinates, the solution will be of the form

$$\Phi^*(\vec{r}) = A + \frac{B}{r} + \left(C_1 r + \frac{C_2}{r^2}\right) \cos \theta + \left(D_1 r^2 + \frac{D_2}{r^3}\right) (3 \cos^2 \theta - 1) + \dots \quad (7.3.17)$$

everywhere. As before, let the coefficients be barred for the interior solution and unbarred for the exterior solution. Requiring that the potential vanish at infinity eliminates all non-negative powers of r , so

$$\Phi_{\text{ext}}^* = \frac{B}{r} + \frac{C_2}{r^2} \cos \theta + \frac{D_2}{r^3} (3 \cos^2 \theta - 1) + \dots \quad (7.3.18)$$

Likewise, all terms that approach infinity as $r \rightarrow 0$ must be set equal to zero. This gives

$$\Phi_{\text{int}}^* = \bar{A} + \bar{C}_1 r \cos \theta + \bar{D}_1 r^2 (3 \cos^2 \theta - 1) + \dots \quad (7.3.19)$$

With no surface currents, our boundary conditions are simply that the normal component of \vec{B} , the tangential components of \vec{H} and the potential are continuous. Consider first the normal component of \vec{B} . If we write

$$\vec{B} = \mu_0(\vec{H} + \vec{M}) \quad (7.3.20)$$

then continuity of the normal component of \vec{B} implies that

$$H_{\text{int},n} + M_n = H_{\text{ext},n}, \quad (M_{\text{ext},n} = 0) \quad (7.3.21)$$

Now $M_n = M(\hat{z} \cdot \hat{r}) = M \cos \theta$, therefore

$$-\frac{\partial}{\partial r} \Phi_{\text{ext}}^* \Big|_{r=a} = \left(-\frac{\partial}{\partial r} \Phi_{\text{int}}^* + M \cos \theta \right)_{r=a} \quad (7.3.22)$$

or

$$\frac{B}{a^2} + \frac{2C_2}{a^3} \cos \theta + \frac{3D_2}{a^4} (3 \cos^2 \theta - 1) \dots = -(\bar{C}_1 \cos \theta + 2\bar{D}_1 a (3 \cos^2 \theta - 1) \dots) + M \cos \theta \quad (7.3.23)$$

which gives

$$B = 0, \quad \frac{2C_2}{a^3} = -\bar{C}_1 + M, \quad \frac{3D_2}{a^4} = -2\bar{D}_1 a, \quad \dots \quad (7.3.24)$$

Again, continuity of Φ^* across the surface of the magnetized sphere implies that

$$\bar{A} = \frac{B}{a}, \quad \bar{C}_1 a = \frac{C_2}{a^2}, \quad \bar{D}_1 a^2 = \frac{D_2}{a^3} \dots \quad (7.3.25)$$

Comparing the two sets of relations, we find that $D_2 = E_2 = \dots = 0$ and $\bar{D}_1 = \bar{E}_1 = \dots = 0$. Then we are left with

$$\begin{aligned} C_2 &= \bar{C}_1 a^3 \\ C_2 &= -\frac{\bar{C}_2 a^3}{2} + \frac{M a^3}{2} \end{aligned} \quad (7.3.26)$$

which together give

$$C_2 = \frac{1}{3} M a^3, \quad \bar{C}_1 = \frac{1}{3} M \quad (7.3.27)$$

and thus we have solved the problem:³

$$\begin{aligned} \Phi_{\text{ext}}^*(r, \theta) &= \frac{1}{3} \frac{M a^3}{r^2} \cos \theta \\ \Phi_{\text{int}}^*(r, \theta) &= \frac{1}{3} M r \cos \theta \end{aligned} \quad (7.3.28)$$

7.3.3 Magnetic shielding

Consider a cylindrically symmetric system in which there is an applied magnetic field along the \hat{x} direction and an (effectively) infinite cylindrical shell of inner radius a and outer radius b situated along the z -axis, as shown in figure (7.4). There are three regions to consider, *viz.*, $r < a$ (inside the shell), $a < r < b$ (in the shell itself) and $r > b$ (outside the shell). Assume that there are no transport currents, so that $\vec{\nabla}^2 \Phi^*$ everywhere. From the cylindrically symmetric electrostatic problem we have solved before, we know that the general form of the potential will be

$$\Phi^*(\rho, \varphi) = A + B \ln \rho + \sum_{n=-\infty, n \neq 0}^{\infty} \rho^n (C_n \cos n\varphi + D_n \sin n\varphi) \quad (7.3.29)$$

³Problem: Show from the given potentials that

$$\begin{aligned} \vec{B}_{\text{ext}} &= \mu_0 \left[\frac{2}{3} \frac{M a^3}{r^3} \cos \theta \hat{r} - \frac{1}{3} \frac{M a^3}{r^3} \sin \theta \hat{\theta} \right] \\ \vec{B}_{\text{int}} &= -\frac{1}{3} \mu M \hat{z} \end{aligned}$$

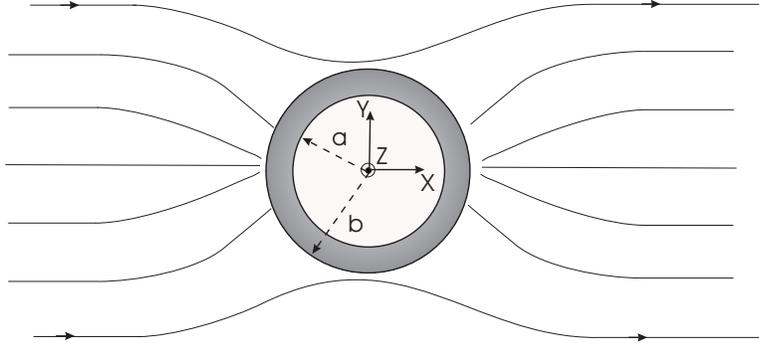


Figure 7.4: Magnetic field inside and out of a magnetic medium

We require that Φ^* be an even function of φ . With the given external field, it must also behave as

$$\lim_{\rho \rightarrow \infty} \Phi^*(\rho, \varphi) = -B_0 x = -B_0 \rho \cos \varphi \quad (7.3.30)$$

This means that $D_n = 0$ for all n . In the exterior ($\rho > b$) we are left with

$$\Phi_{\text{ext}}^*(\rho, \varphi) = \sum_{n=1}^{\infty} \rho^{-n} C_{-n} \cos n\varphi - B_0 \rho \cos \varphi \quad (7.3.31)$$

whereas in the interior ($\rho < a$), regularity at the center requires that

$$\Phi_{\text{int}}^*(\rho, \varphi) = \bar{A} + \sum_{n=1}^{\infty} \rho^n \bar{C}_n \cos n\varphi \quad (7.3.32)$$

Inside the shell, we must in principle include all terms (even in φ), so that

$$\Phi_{\text{shell}}^*(\rho, \varphi) = A' + B' \ln \rho + \sum_{n=1}^{\infty} (C'_n \rho^n + C'_{-n} \rho^{-n}) \cos n\varphi \quad (7.3.33)$$

Matching the shell and exterior solutions gives

$$\begin{aligned} A' &= 0 = B' \\ -B_0 + \frac{C_{-1}}{b^2} &= C'_1 + \frac{C'_{-1}}{b^2} \\ \frac{C_{-n}}{b^n} &= C'_n b^n + \frac{C'_{-n}}{b^n}, \quad n > 1 \end{aligned} \quad (7.3.34)$$

Likewise, matching the shell and interior solutions gives

$$\begin{aligned}\bar{A} &= A' = 0 \\ \bar{C}_n &= C'_n + \frac{C'_{-n}}{a^{2n}}, \quad n \geq 1\end{aligned}\tag{7.3.35}$$

Let us now apply the continuity of the normal components of \vec{B} . For the exterior-shell boundary we find

$$\begin{aligned}\mu_0(-B_0 - \frac{C_{-1}}{b^2}) &= \mu(C'_1 - \frac{C'_{-1}}{b^2}) \\ -\mu_0 \frac{C_{-n}}{b^{n+1}} &= \mu(C'_n b^{n-1} - \frac{C'_{-n}}{b^{n+1}}), \quad n > 1\end{aligned}\tag{7.3.36}$$

Solving the system in (7.3.34) and (7.3.35) gives

$$\begin{aligned}C'_1 &= \frac{1}{2b^2\mu} [C_{-1}(\mu - \mu_0) - b^2 B_0(\mu + \mu_0)] \\ C'_{-1} &= \frac{1}{2\mu} [C_{-1}(\mu + \mu_0) - b^2 B_0(\mu - \mu_0)] \\ C'_n &= \frac{\mu - \mu_0}{2b^{2n}\mu} C_{-n}, \quad n > 1 \\ C'_{-n} &= \frac{\mu + \mu_0}{2\mu} C_{-n}, \quad n > 1\end{aligned}\tag{7.3.37}$$

Again, applying the continuity of the normal component of \vec{B} across the shell-interior boundary, we have

$$\mu_0 \bar{C}_n = \mu \left(C'_n - \frac{C'_{-n}}{a^{2n}} \right)\tag{7.3.38}$$

Combining these equations with (7.3.35) gives

$$\begin{aligned}C'_n &= \frac{\bar{C}_n(\mu + \mu_0)}{2\mu} \\ C'_{-n} &= \frac{\bar{C}_n a^{2n}(\mu - \mu_0)}{2\mu}\end{aligned}\tag{7.3.39}$$

These conditions can be satisfied if we take $\bar{C}_n = 0$ for all $n > 1$. Then $C'_n = 0 = C'_{-n}$ and therefore $C_{-n} = 0$ for all $n > 1$. We are left with the relations:

$$C'_1 = \frac{\bar{C}_1(\mu + \mu_0)}{2\mu} = \frac{1}{2b^2\mu} [C_{-1}(\mu - \mu_0) - b^2 B_0(\mu + \mu_0)]$$

$$C'_{-1} = \frac{\bar{C}_1 a^2 (\mu - \mu_0)}{2\mu} = \frac{1}{2\mu} [C_{-1}(\mu + \mu_0) - b^2 B_0 (\mu - \mu_0)] \quad (7.3.40)$$

These equations in turn yield ($K_m = \mu/\mu_0$)

$$\bar{C}_1 = -\frac{4K_m B_0}{(K_m + 1)^2 - a^2/b^2 (K_m - 1)^2} \quad (7.3.41)$$

and

$$\begin{aligned} C'_1 &= -\frac{2B_0(K_m + 1)}{(K_m + 1)^2 - a^2/b^2 (K_m - 1)^2} \\ C'_{-1} &= -\frac{2a^2 B_0(K_m - 1)}{(K_m + 1)^2 - a^2/b^2 (K_m - 1)^2} \\ C_{-1} &= \frac{(b^2 - a^2)B_0(K_m^2 - 1)}{(K_m + 1)^2 - a^2/b^2 (K_m - 1)^2} \end{aligned} \quad (7.3.42)$$

The “shielding” is apparent from the fact that the field inside the cylinder is given by $\vec{H}_{\text{int}} = \bar{C}_1 \hat{x}$ and \bar{C}_1 is a decreasing function of K_m , behaving as $[(1 - a^2/b^2)K_m]^{-1}$ when $K_m \gg 1$. Considerable shielding from the external magnetic field can be achieved within the cylinder if, in addition, the thickness of the shell can be made significant.

7.4 Magnetic Energy

When we examined the electrostatic energy of a charge distribution, we considered the work that was necessary to produce the given distribution. Electric charges are sources of the electrostatic field, whereas electric currents are sources of the magnetic field. The magnetic energy of a collection of electric circuits will just be the work that must be done to create the given distribution.

Consider first a single circuit with a source voltage of V applied. The circuit may be in a time varying magnetic field \vec{B} , which will also induce an e.m.f. in it according to Faraday’s law. If the resistance of the circuit is R , we can write the following equation (Kirchoff’s voltage law)

$$V + \mathcal{E} = iR = R \frac{dq}{dt}. \quad (7.4.1)$$

Now the work done by the source in moving an electric charge dq through the circuit is

$$dW = Vdq = -\mathcal{E}dq + iRdq = (-\mathcal{E}i + i^2 R)dt = id\Phi_B + i^2 Rdt \quad (7.4.2)$$

where we have put $d\Phi_B = -\mathcal{E}dt$. The term $i^2 Rdt$ represents the energy dissipated into the environment by resistance in the form of heat. The first term is the work done against

the induced e.m.f. in the circuit. This is the work required to change the magnetic field and we can write

$$dW_B = id\Phi_B \quad (7.4.3)$$

For a rigid, stationary circuit, showing no energy loss other than resistive, dW_B is the change in magnetic energy of the circuit.

Now if there are N circuits,

$$dW_B = \sum_{j=1}^N i_j d\Phi_j \quad (7.4.4)$$

where Φ_j is the magnetic flux through circuit j . If the changes in the external magnetic field are produced by changes in the currents only, then

$$d\Phi_j = \sum_{k=1}^N \frac{d\Phi_j}{di_k} di_k = \sum_{k=1}^N M_{jk} di_k \quad (7.4.5)$$

where M_{jk} represents the mutual inductance between circuits j and k . It follows that

$$dW_B = \sum_{j,k=1}^N M_{jk} i_j di_k \quad (7.4.6)$$

Now for linear media, M_{jk} is independent of the currents, so the integration may be done quite easily. Let us imagine that all the currents are brought from zero to their maximum value in concert, *i.e.*, $i_j = \alpha i_j^{\max}$, $\alpha \in [0, 1]$, then

$$W_B = \sum_{j,k=1}^N M_{jk} i_j^{\max} i_k^{\max} \int_0^1 \alpha d\alpha = \frac{1}{2} \sum_{j,k=1}^N M_{jk} i_j^{\max} i_k^{\max} \quad (7.4.7)$$

This gives the energy of the magnetic field of N static coupled circuits in linear magnetic media as

$$U = \frac{1}{2} \sum_{j,k=1}^N M_{jk} i_j i_k \quad (7.4.8)$$

where we have now suppressed the superscript “max”. An alternative expression for the energy of the magnetic field may be given in terms of the field itself; let us note that expression (7.4.8) can also be written in terms of the magnetic flux by using

$$\Phi_k = \sum_{j=1}^N M_{kj} i_j \Rightarrow U = \frac{1}{2} \sum_k \Phi_k i_k \quad (7.4.9)$$

Now

$$\Phi_k = \int_{S_k} \vec{B} \cdot d\vec{S} = \int_{S_k} (\vec{\nabla} \times \vec{A}) \cdot d\vec{S} = \oint_{C_k} \vec{A} \cdot d\vec{r} \Rightarrow U = \frac{1}{2} \sum_k i_k \oint_{C_k} \vec{A} \cdot d\vec{r} \quad (7.4.10)$$

The expression can be made more general by noting that if we had instead volumetric current distributions, defined by densities, then the substitution

$$\sum_k i_k \oint_{C_k} d\vec{r} \rightarrow \int_V d^3\vec{r}' j(\vec{r}') \quad (7.4.11)$$

yields

$$U = \frac{1}{2} \int_V d^3\vec{r}' j(\vec{r}') \cdot \vec{A}(\vec{r}') \quad (7.4.12)$$

However, since $\vec{\nabla} \times \vec{H} = \vec{j}$,

$$U = \frac{1}{2} \int_V d^3\vec{r}' (\vec{\nabla}' \times \vec{H}) \cdot \vec{A} \quad (7.4.13)$$

But, applying the identity

$$(\vec{\nabla}' \times \vec{H}) \cdot \vec{A} = (\vec{\nabla}' \times \vec{A}) \cdot \vec{H} - \vec{\nabla}' \cdot (\vec{A} \times \vec{H}) = \vec{B} \cdot \vec{H} - \vec{\nabla}' \cdot (\vec{A} \times \vec{H}) \quad (7.4.14)$$

we find

$$U = \frac{1}{2} \int_V d^3\vec{r}' (\vec{B} \cdot \vec{H}) - \frac{1}{2} \oint_S d\vec{S} \cdot (\vec{A} \times \vec{H}) \quad (7.4.15)$$

where we have used Gauss' law. If we take the volume to be all of space and the surface S to be at infinity, and if the fields fall off rapidly enough, the surface integral will vanish and we will be left with

$$U = \frac{1}{2} \int d^3\vec{r}' (\vec{B} \cdot \vec{H}). \quad (7.4.16)$$

The energy density of the magnetic field has an expression that is similar to that of the electrostatic field,

$$u_B = \frac{1}{2} (\vec{B} \cdot \vec{H}). \quad (7.4.17)$$

For linear magnetic media $\vec{H} = \vec{B}/\mu$ in which case $u_B = \vec{B}^2/2\mu = \mu\vec{H}^2/2$.

7.5 Maxwell's equations

When all is said and done, the electromagnetic equations for stationary currents are given by (7.1.29), which are supplemented with the boundary conditions in (7.2.1) and (7.2.4).

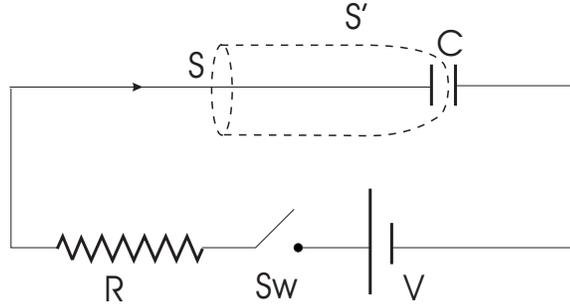


Figure 7.5: Circuit showing the failure of the stationary electromagnetic equations

Unfortunately, it is easy to argue that the equations are inconsistent when the currents are not stationary, *i.e.*, when $\partial\rho/\partial t \neq 0$.

To see why this is so, consider the circuit shown in figure (7.5). When the switch Sw is turned on, a current flows through the wire, charging up the capacitor, C . We might apply Ampère's law to the surface closed loop bounding two surfaces S and S' in the figure,

$$\oint_C \vec{H} \cdot d\vec{r} = i \quad (7.5.1)$$

where i represents the current flowing through either S or S' . The problem is that the current depends on which surface is considered – it is vanishing for surface S' , but not for surface S . This ambiguity appears while the capacitor is charging (or discharging when the switch is turned off), *i.e.*, when $\partial\rho/\partial t \neq 0$ on the capacitor, but not in the stationary state.

Applying the continuity equation (6.1.11) to the right hand side of the first of the electromagnetic field equations,

$$\frac{\partial\rho_f}{\partial t} + \vec{\nabla} \cdot \vec{j}_f = 0 \Rightarrow \frac{\partial}{\partial t} \vec{\nabla} \cdot \vec{D} + \vec{\nabla} \cdot \vec{j}_f = \vec{\nabla} \cdot \left(\frac{\partial\vec{D}}{\partial t} + \vec{j}_f \right) = 0 \quad (7.5.2)$$

Therefore, Maxwell proposed that the last of the electromagnetic equations should be modified according to

$$\vec{\nabla} \times \vec{H} = \frac{\partial\vec{D}}{\partial t} + \vec{j}_f$$

or

$$\vec{\nabla} \times \vec{H} - \frac{\partial\vec{D}}{\partial t} = \vec{j}_f \quad (7.5.3)$$

To summarize, we give the complete, self-consistent set of electromagnetic field equations

$$\vec{\nabla} \cdot \vec{D} = \rho_f$$

$$\begin{aligned}\vec{\nabla} \times \vec{E} + \frac{\partial \vec{B}}{\partial t} &= 0 \\ \vec{\nabla} \cdot \vec{B} &= 0 \\ \vec{\nabla} \times \vec{H} - \frac{\partial \vec{D}}{\partial t} &= \vec{j}_f\end{aligned}\tag{7.5.4}$$

They are known today as Maxwell's equations, although it should be noted that they are the product of more than a century of careful experimentation and deep thinking by many. We now turn to examining these equations in detail.

Chapter 8

Maxwell's Equations

8.1 A brief review

Let's begin with a brief revision of Maxwell's equations (we will always use the MKS system). Maxwell's system consists of the four equations which you have seen in the following form

$$\begin{aligned}\vec{\nabla} \cdot \vec{D} &= \rho_f \\ \vec{\nabla} \times \vec{E} + \frac{\partial \vec{B}}{\partial t} &= 0 \\ \vec{\nabla} \cdot \vec{B} &= 0 \\ \vec{\nabla} \times \vec{H} - \frac{\partial \vec{D}}{\partial t} &= \vec{j}_f\end{aligned}\tag{8.1.1}$$

where the suffix “*f*” represents “free”, *i.e.*, only the unbound (transport) charges and currents appear on the r.h.s of the above equations. The fields \vec{E} and \vec{B} are the electric and magnetic fields respectively and the fields \vec{D} and \vec{H} are respectively the electric displacement and the magnetic intensity. The latter two fields are phenomenological, *i.e.*, they are defined according to what is actually relevant to the actual measurement of electric and magnetic phenomena in materials. They are different from the electric and magnetic fields only in dielectric or magnetic media.

Recall that the electric displacement is defined according to

$$\vec{D} = \epsilon_o \vec{E} + \vec{P},\tag{8.1.2}$$

where \vec{P} is the electric polarization vector of the dielectrics present. If we think of the dielectric as consisting of a very large number of microscopic electric dipoles and consider

a volume element Δv which contains a large number of these dipoles (Δv is a volume element that is considered to be macroscopically infinitesimal but microscopically large) then the total electric dipole moment of Δv is defined by

$$\Delta \vec{p} = \sum_{i \in \Delta v} \vec{p}_i, \quad (8.1.3)$$

where i labels the dipoles inside Δv and \vec{p}_i is the dipole moment of dipole i . The polarization vector, \vec{P} , is defined as the electric dipole moment per unit volume

$$\vec{P} = \lim_{\Delta v \rightarrow 0} \frac{\Delta \vec{p}}{\Delta v}, \quad (8.1.4)$$

This concept is useful only if we are interested in the fields on scales that are much larger than the typical dipole length of the dielectric.

Likewise, the magnetic intensity vector is defined by

$$\vec{H} = \frac{\vec{B}}{\mu_0} - \vec{M}. \quad (8.1.5)$$

The magnetization, \vec{M} , is defined in a manner similar to our definition of the electric polarization. Imagine that the magnetic medium is made up of a large number of magnetic dipoles (microscopic current-loops), each of which has magnetic dipole moment, \vec{m}_i . The total magnetic dipole moment of the volume element Δv is just

$$\Delta \vec{m} = \sum_{i \in \Delta v} \vec{m}_i, \quad (8.1.6)$$

and the magnetization vector \vec{M} is defined as the magnetic dipole moment per unit volume

$$\vec{M} = \lim_{\Delta v \rightarrow 0} \frac{\Delta \vec{m}}{\Delta v}. \quad (8.1.7)$$

Again, the same comment we made before holds: \vec{M} is a useful concept only if we are interested in effects on scales that are much larger than the typical size of the magnetic dipoles.

Notes:

- In addition to Maxwell's equations we assume that the continuity equation of matter holds, *i.e.*,

$$\frac{\partial \rho}{\partial t} + \vec{\nabla} \cdot \vec{j} = 0. \quad (8.1.8)$$

- The right hand sides of Maxwell's equations as given in (8.1.1) involve only free charge and current densities. That is because the microscopic effects of the bound charges and currents (the electric and magnetic dipoles) are taken into account in an average way through the polarization vector, \vec{P} , and magnetization vector, \vec{M} . For this reason, Maxwell's equations in the form given in (8.1.1) are usually called the macroscopic Maxwell's equations.
- It follows that \vec{D} and \vec{H} are *derived* fields. As said before, they are defined so as to phenomenologically account for the contribution of the atomic charges and currents in an average way (this is why Δv must be microscopically *large*). A complete description of Maxwell's equations in the form given in (8.1.1) must, therefore, include relations between \vec{D} and \vec{E} and between \vec{H} and \vec{B} .

One could also write Maxwell's equations in the following way:

$$\begin{aligned}
 \vec{\nabla} \cdot \vec{E} &= \frac{\rho}{\epsilon_0} \\
 \vec{\nabla} \times \vec{E} + \frac{\partial \vec{B}}{\partial t} &= 0 \\
 \vec{\nabla} \cdot \vec{B} &= 0 \\
 \vec{\nabla} \times \vec{B} - \epsilon_0 \mu_0 \frac{\partial \vec{E}}{\partial t} &= \mu_0 \vec{j}
 \end{aligned} \tag{8.1.9}$$

but, now, the r.h.s. of the above equations no longer represent just the free charge and current densities but *all* charge and current densities, bound and unbound. These are usually called the “vacuum form” of Maxwell's equations. One can say that this is the “fundamental” form of Maxwell's equations.

Notice that the central two equations remain unchanged and uncoupled to the charge and current distributions. These two equations are not dynamical. They are actually geometric identities (called Bianchi identities), common to all so-called “gauge theories”. In practice they serve as constraints, reducing the number of independent fields as we see in the following section.

8.2 An alternative description

Consider the third equation in (8.1.9) and note that

$$\vec{\nabla} \cdot \vec{B} = 0 \rightarrow \vec{B} = \vec{\nabla} \times \vec{A}. \tag{8.2.1}$$

The vector \vec{A} is a vector function (field) and is called the vector potential. It is a “potential” in the sense that the magnetic field, which is related to the magnetic force on a charge, is

a derivative of this vector. In the same way,

$$\begin{aligned}
 \vec{\nabla} \times \vec{E} + \frac{\partial \vec{B}}{\partial t} &= 0 \\
 \rightarrow \vec{\nabla} \times \left[\vec{E} + \frac{\partial \vec{A}}{\partial t} \right] &= 0 \\
 \rightarrow \vec{E} = -\vec{\nabla} \phi - \frac{\partial \vec{A}}{\partial t}, & \quad (8.2.2)
 \end{aligned}$$

where ϕ is a scalar function (field). It is called the scalar potential. Thus we see that instead of the original six components of \vec{E} and \vec{B} , we have only four functions, *viz.*, ϕ and the three components of \vec{A} . The dynamical equations are the first and last ones in (8.1.9). They can be written in terms of (ϕ, \vec{A}) by direct substitution of the above expressions for the electric and magnetic fields. This gives

$$\vec{\nabla} \cdot \vec{E} = -\vec{\nabla}^2 \phi - \frac{\partial}{\partial t} \vec{\nabla} \cdot \vec{A} = \frac{\rho}{\epsilon_o} \quad (8.2.3)$$

and

$$\vec{\nabla} \times (\vec{\nabla} \times \vec{A}) - \epsilon_o \mu_o \frac{\partial}{\partial t} \left(-\vec{\nabla} \phi - \frac{\partial \vec{A}}{\partial t} \right) = \mu_o \vec{j}. \quad (8.2.4)$$

The product $\epsilon_o \mu_o$ will appear throughout. It has a very special meaning as the inverse square speed of light in the vacuum as we will soon see. Let us therefore simplify our notation by using “ $1/c^2$ ” for this quantity. Thus we write (8.2.4) as

$$\frac{1}{c^2} \frac{\partial^2 \vec{A}}{\partial t^2} - \vec{\nabla}^2 \vec{A} + \vec{\nabla} \left(\vec{\nabla} \cdot \vec{A} + \frac{1}{c^2} \frac{\partial \phi}{\partial t} \right) = \mu_o \vec{j}. \quad (8.2.5)$$

The two equations given in (8.2.3) and (8.2.5) are far from elegant. We will now look for ways to (a) simplify these equations and (b) solve them. To do so we must first understand some of their properties. The first of these is the so-called “gauge invariance”, which is our next topic.

8.3 Gauge invariance

The dynamical equations are greatly simplified if we recognize that only the fields \vec{E} and \vec{B} are measured classically. (The situation is a bit more complicated when quantum effects are accounted for.) But the electric and magnetic fields are derivatives of the scalar and vector potentials. One easily checks that the transformation

$$\vec{A} \rightarrow \vec{A}' = \vec{A} + \vec{\nabla} \Lambda, \quad (8.3.1)$$

for *any* scalar function of space and time, Λ , leaves the magnetic field unchanged (the curl of a gradient is identically zero). It does not, however, leave the electric field unchanged: notice that as $\vec{A} \rightarrow \vec{A}' = \vec{A} + \vec{\nabla}\Lambda$ the electric field transforms as

$$\vec{E} \rightarrow \vec{E}' = -\vec{\nabla}\phi - \frac{\partial \vec{A}'}{\partial t} = -\vec{\nabla}\phi - \frac{\partial \vec{A}}{\partial t} - \vec{\nabla}\frac{\partial \Lambda}{\partial t} = \vec{E} - \vec{\nabla}\frac{\partial \Lambda}{\partial t}. \quad (8.3.2)$$

However, we have said nothing about the transformation of ϕ . Consider the simultaneous transformations

$$\begin{aligned} \vec{A} &\rightarrow \vec{A}' = \vec{A} + \vec{\nabla}\Lambda \\ \phi &\rightarrow \phi' = \phi - \frac{\partial \Lambda}{\partial t}. \end{aligned} \quad (8.3.3)$$

Clearly \vec{B} is unaffected by the above transformations as we saw earlier. But now, because ϕ also is transformed, even \vec{E} remains unaffected as one can explicitly check:

$$\vec{E} \rightarrow \vec{E}' = -\vec{\nabla}\phi' - \frac{\partial \vec{A}'}{\partial t} = \vec{E}. \quad (8.3.4)$$

The transformations in (8.3.3) are called “gauge transformations”. The electric and magnetic fields are invariant under gauge transformations. This means that Maxwell’s equations are also invariant under gauge transformations, because they are written in terms of the fields \vec{E} and \vec{B} . Notice that Λ is completely arbitrary. The freedom we have in choosing Λ is called a “gauge freedom”. We will now see how a judicious choice of Λ can simplify the dynamical equations in (8.2.3) and (8.2.5).

8.4 Choice of gauge

It should be clear that choosing a particular Λ is equivalent to choosing a particular condition on the fields \vec{A} and ϕ . A few examples will clarify this statement.

8.4.1 The Lorentz gauge

Consider the condition

$$\vec{\nabla} \cdot \vec{A} + \frac{1}{c^2} \frac{\partial \phi}{\partial t} = 0. \quad (8.4.1)$$

(This is called the Lorentz condition.) Let us convince ourselves that this amounts to choosing a particular Λ . For suppose that we have a solution (ϕ, \vec{A}) that does *not* satisfy the condition. Can we find a Λ such that the transformed pair (ϕ', \vec{A}') (related therefore to (ϕ, \vec{A}) by a gauge transformation) does obey the condition? If we can, because the

dynamical equations themselves are invariant under these transformations, then (ϕ', \vec{A}') will certainly obey Maxwell's equations and, moreover, the electric and magnetic fields derivable from (ϕ', \vec{A}') will be identical to those derivable from (ϕ, \vec{A}) .

By hypothesis,

$$\vec{\nabla} \cdot \vec{A}' + \frac{1}{c^2} \frac{\partial \phi'}{\partial t} = 0 \quad (8.4.2)$$

and

$$\begin{aligned} \vec{A}' &= \vec{A} + \vec{\nabla} \Lambda \\ \phi' &= \phi - \frac{\partial \Lambda}{\partial t}. \end{aligned} \quad (8.4.3)$$

Therefore,

$$\vec{\nabla} \cdot \vec{A} + \vec{\nabla} \cdot \vec{\nabla} \Lambda + \frac{1}{c^2} \frac{\partial \phi}{\partial t} - \frac{1}{c^2} \frac{\partial^2 \Lambda}{\partial t^2} = 0, \quad (8.4.4)$$

which implies that

$$\vec{\nabla}^2 \Lambda - \frac{1}{c^2} \frac{\partial^2 \Lambda}{\partial t^2} = - \left(\vec{\nabla} \cdot \vec{A} + \frac{1}{c^2} \frac{\partial \phi}{\partial t} \right). \quad (8.4.5)$$

But, as we know the solution pair (ϕ, \vec{A}) of Maxwell's equations, the r.h.s. is known. What we have, therefore, is an equation for Λ . If the equation admits solutions (subject to appropriate boundary conditions about which we shall speak later) then such a Λ can be found. Now, because (ϕ', \vec{A}') obey Maxwell's equations, then imposing the condition in (8.4.1) it follows that the dynamical equations for (ϕ', \vec{A}') are

$$\begin{aligned} \frac{1}{c^2} \frac{\partial^2 \vec{A}'}{\partial t^2} - \vec{\nabla}^2 \vec{A}' &= \mu_o \vec{j} \\ \frac{1}{c^2} \frac{\partial^2 \phi'}{\partial t^2} - \vec{\nabla}^2 \phi' &= \frac{\rho}{\epsilon_o}. \end{aligned} \quad (8.4.6)$$

Thus the condition (8.4.1) could be imposed from the start and the equations (8.4.6) used. Maxwell's equations in Lorentz gauge can be summarized as follows:

$$\begin{aligned} \frac{1}{c^2} \frac{\partial^2 \phi}{\partial t^2} - \vec{\nabla}^2 \phi &= \frac{\rho}{\epsilon_o} \\ \frac{1}{c^2} \frac{\partial^2 \vec{A}}{\partial t^2} - \vec{\nabla}^2 \vec{A} &= \mu_o \vec{j} \\ \frac{1}{c^2} \frac{\partial \phi}{\partial t} + \vec{\nabla} \cdot \vec{A} &= 0. \end{aligned} \quad (8.4.7)$$

Notes:

- The equations for ϕ and \vec{A} are the same. The Lorentz gauge treats ϕ and \vec{A} on an equal footing.
- The vector and scalar fields decouple in this gauge.
- Even for potentials that satisfy the Lorentz gauge condition there is a degree of arbitrariness that survives. Thus, for example, the r.h.s. of our equation for Λ would turn into zero and our equation for Λ would read

$$\vec{\nabla}^2 \Lambda - \frac{1}{c^2} \frac{\partial^2 \Lambda}{\partial t^2} = 0, \quad (8.4.8)$$

and a transformation by a Λ that obeys this equation preserves the Lorentz condition. All potentials that satisfy the Lorentz gauge condition are said to belong to the Lorentz gauge.

- We will see in a short time that the Lorentz gauge condition is invariant under the transformations of the special theory of relativity. This is very convenient, because if it holds in one inertial frame it will hold in all inertial frames.

8.4.2 The Coulomb gauge

Another gauge condition that is often used is the so-called ‘‘Coulomb’’ gauge (or radiation gauge or transverse gauge). The condition is simply

$$\vec{\nabla} \cdot \vec{A} = 0. \quad (8.4.9)$$

Let us see if it is a good gauge condition. Following the argument in the previous section, let (ϕ, \vec{A}) be a solution pair (of Maxwell’s equations) that does *not* obey the condition. Can we find a Λ that would take the pair $(\phi, \vec{A}) \rightarrow (\phi', \vec{A}')$ such that the new pair (ϕ', \vec{A}') does satisfy the condition? By hypothesis

$$\vec{\nabla} \cdot \vec{A}' = 0 \rightarrow \vec{\nabla} \cdot \vec{A} + \vec{\nabla}^2 \Lambda = 0. \quad (8.4.10)$$

This means that

$$\vec{\nabla}^2 \Lambda = -\vec{\nabla} \cdot \vec{A} \quad (8.4.11)$$

and, once again, we know the r.h.s. of the equation. Provided a solution to this equation can be found, subject to appropriate boundary conditions, we can find a Λ that takes the pair (ϕ, \vec{A}) to a pair (ϕ', \vec{A}') that obeys the Coulomb gauge condition. It is a good exercise to show that in this gauge, the dynamical equations are

$$\vec{\nabla}^2 \phi = -\frac{\rho}{\epsilon_0}$$

$$\frac{1}{c^2} \frac{\partial^2 \vec{A}}{\partial t^2} - \vec{\nabla}^2 \vec{A} = \mu_0 \vec{j} - \frac{1}{c^2} \vec{\nabla} \frac{\partial \phi}{\partial t}. \quad (8.4.12)$$

The equation for the scalar potential is exceedingly simple.¹ Using

$$\vec{\nabla}^2 \frac{1}{|\vec{r} - \vec{r}'|} = -4\pi \delta^3(\vec{r} - \vec{r}'), \quad (8.4.13)$$

one can immediately write

$$\phi(\vec{r}, t) = \int \frac{d^3 \vec{r}'}{4\pi \epsilon_0} \frac{\rho(\vec{r}', t)}{|\vec{r} - \vec{r}'|}. \quad (8.4.14)$$

This is precisely what we wrote down in electrostatics for a static charge distribution described by the density $\rho(\vec{r})$. This is why the gauge is called the Coulomb gauge. Note, however, that here the density may vary in time.

The quantity $\vec{\nabla} \phi$ that appears on the r.h.s. of the equation for the vector potential, \vec{A} , certainly has vanishing curl and this inspires hopes that it may be able to cancel an irrotational portion of the current density vector which also occurs on the r.h.s. This is actually so. Consider the continuity equation

$$\frac{\partial \rho}{\partial t} + \vec{\nabla} \cdot \vec{j} = 0. \quad (8.4.15)$$

Now, using the equation for ϕ and taking a time derivative we get²

$$\vec{\nabla}^2 \frac{\partial \phi}{\partial t} = -\frac{1}{\epsilon_0} \frac{\partial \rho}{\partial t} = \frac{1}{\epsilon_0} \vec{\nabla} \cdot \vec{j} = \frac{1}{\epsilon_0} \vec{\nabla} \cdot \vec{j}_l. \quad (8.4.16)$$

It follows that

$$\vec{\nabla} \frac{\partial \phi}{\partial t} = \frac{1}{\epsilon_0} \vec{j}_l, \quad (8.4.17)$$

which, when inserted into the equation for \vec{A} in (8.4.12), gives

$$\frac{1}{c^2} \frac{\partial^2 \vec{A}}{\partial t^2} - \vec{\nabla}^2 \vec{A} = \mu_0 \vec{j}_t. \quad (8.4.18)$$

Thus the field equations in Coulomb gauge can be summarized as

$$\vec{\nabla}^2 \phi = -\frac{\rho}{\epsilon_0}$$

¹see section 9.

²Problem: Show that *any* vector can always be written as the sum of a “longitudinal” part, j_l , and a “transverse” part, j_t , *i.e.*, $\vec{j} = \vec{j}_l + \vec{j}_t$, where: $\vec{\nabla} \times j_l = 0$ and $\vec{\nabla} \cdot j_t = 0$.

$$\begin{aligned}\frac{1}{c^2} \frac{\partial^2 \vec{A}}{\partial t^2} - \vec{\nabla}^2 \vec{A} &= \mu_o \vec{j}_t \\ \vec{\nabla} \cdot \vec{A} &= 0.\end{aligned}\tag{8.4.19}$$

Notes:

- Because it is only the transverse part, \vec{j}_t , that acts as the source for \vec{A} (and not the entire current \vec{j}), this gauge is also called the “transverse” gauge.
- It is often used in the vacuum, *i.e.*, when no sources are present. In that case the condition $\vec{\nabla} \cdot \vec{A} = 0$ can be supplemented by $\phi = 0$. Then

$$\vec{E} = -\frac{\partial \vec{A}}{\partial t}, \quad \vec{B} = \vec{\nabla} \times \vec{A}.\tag{8.4.20}$$

- The gauge condition is not invariant under the transformations of the special theory of relativity. This means that if it holds in one inertial frame it will not in general hold in any other.

Other gauge conditions are used, depending upon the needs of the problem to be solved.³ Here we shall emphasize the above two.

8.5 The homogeneous wave equation: a first look

Let's consider the dynamical equations in the Lorentz gauge. All the fields (ϕ and the three components of \vec{A}) obey an equation of the form

$$\left[\frac{1}{c^2} \frac{\partial^2}{\partial t^2} - \vec{\nabla}^2 \right] \psi(\vec{r}, t) = f(\vec{r}, t).\tag{8.5.1}$$

where $\psi \sim \{\phi, A_i\}$ and $f \sim \{\rho, j_i\}$. We will study this equation in some detail during this course, but first let's get a feeling for some of its simplest solutions. Take, for example, $f(\vec{r}, t) = 0$ (no sources, so we're asking for *vacuum* solutions). Any equation of this type admits solutions of the form

$$\psi = \psi(\omega t \pm \vec{k} \cdot \vec{r}),\tag{8.5.2}$$

where ω and \vec{k} are arbitrary (even complex) constants and ψ is an *arbitrary* function of its argument. To convince ourselves of this, take $u = \omega t \pm \vec{k} \cdot \vec{r}$ so that $\psi = \psi(u)$. Then

$$\vec{\nabla} \psi = \pm \vec{k} \psi'$$

³Problem: Show that one can choose $A_3 = 0$ as a gauge condition. This is the “axial gauge” and is quite useful at times.

$$\frac{\partial \psi}{\partial t} = \omega \psi', \quad (8.5.3)$$

where $\psi' = d\psi/du$. Thus

$$\begin{aligned} \vec{\nabla}^2 \psi &= \vec{k}^2 \psi'' \\ \frac{\partial^2 \psi}{\partial t^2} &= \omega^2 \psi''. \end{aligned} \quad (8.5.4)$$

Plug this into the wave-equation and find

$$\left(\frac{\omega^2}{c^2} - \vec{k}^2 \right) \psi'' = 0, \quad (8.5.5)$$

which means that, either

$$|\vec{k}| = \frac{|\omega|}{c}, \quad (8.5.6)$$

or $\psi'' = 0$. But $\psi'' = 0$ yields $\psi = au + b$. Ignoring this last possibility, consider the following special function of the variable $u = \omega t \pm \vec{k} \cdot \vec{r}$:

$$\psi(\vec{r}, t) = \psi_o e^{-i(\omega t \pm \vec{k} \cdot \vec{r})} \quad (8.5.7)$$

or, explicitly in terms of the electromagnetic fields

$$\begin{aligned} \phi(\vec{r}, t) &= \phi_o e^{-i(\omega t \pm \vec{k} \cdot \vec{r})} \\ \vec{A}(\vec{r}, t) &= \vec{A}_o e^{-i(\omega t \pm \vec{k} \cdot \vec{r})}. \end{aligned} \quad (8.5.8)$$

Consider the special case in which ω and \vec{k} are both real.⁴ The solution is harmonic and represents plane waves propagating through space. The constant phase surfaces are given by $\omega t \pm \vec{k} \cdot \vec{r} = \text{const}$. At any given time ($t = t_o$, fixed) we see that

$$\vec{k} \cdot \vec{r} = \text{const} \mp \omega t_o. \quad (8.5.9)$$

The r.h.s. is a constant, so this is the equation of a plane whose normal is \hat{k} (the unit vector in the direction of \vec{k}). Thus, the vector \vec{k} is perpendicular to the constant phase planes (see figure 1). It is called the wave-vector. The modulus $|\vec{k}|$ is called the wave number and ω is the angular frequency. The wavelength is simply

$$\lambda = \frac{2\pi}{|\vec{k}|} = \frac{2\pi c}{\omega}. \quad (8.5.10)$$

⁴We will take \vec{k} and ω to be real, but \vec{A}_o and ϕ_o may be complex. The physical fields are the real parts of these solutions.

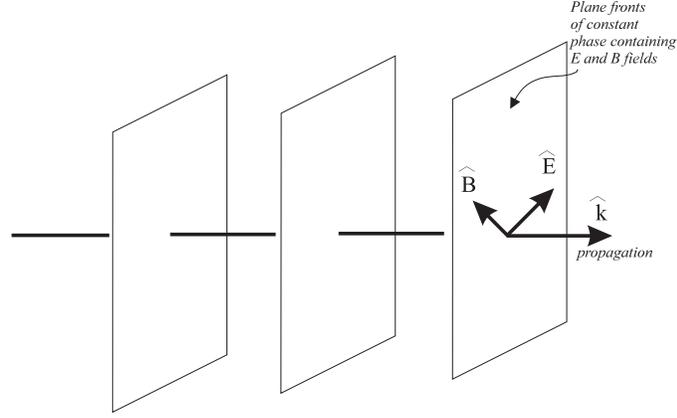


Figure 8.1: Plane monochromatic waves

Thus we find that $\lambda f = c$, where $f = \omega/2\pi$ is the frequency, *i.e.*, the wave travels at the speed c .

Let us now inquire about the measurable fields \vec{E} and \vec{B} . We have

$$\begin{aligned}\vec{B} &= \vec{\nabla} \times \vec{A} = \mp i\vec{k} \times \vec{A} \\ \vec{E} &= -\vec{\nabla}\phi - \frac{\partial \vec{A}}{\partial t} = \pm i\vec{k}\phi + i\omega\vec{A},\end{aligned}\quad (8.5.11)$$

and we should not forget that these fields are in the Lorentz gauge. The gauge condition supplements the above equations and reads

$$\mp i\vec{k} \cdot \vec{A} - \frac{i\omega}{c^2}\phi = 0. \quad (8.5.12)$$

We can eliminate ϕ using this equation,

$$\phi = \mp \frac{c^2}{\omega}(\vec{k} \cdot \vec{A}). \quad (8.5.13)$$

and substitute into (8.5.11) to find

$$\begin{aligned}\vec{E} &= -i\frac{c^2}{\omega}\vec{k}(\vec{k} \cdot \vec{A}) + i\omega\vec{A} = -\frac{ic^2}{\omega}\vec{k} \times (\vec{k} \times \vec{A}) \\ \vec{B} &= \mp i\vec{k} \times \vec{A},\end{aligned}\quad (8.5.14)$$

so the equations are written entirely in terms of \vec{A} . A straightforward calculation reveals

$$\vec{E} \times \vec{B} = \mp \frac{c^2}{\omega} \vec{B}^2 \vec{k}$$

$$\begin{aligned}\vec{B} \times \vec{k} &= \mp \frac{\omega}{c^2} \vec{E} \\ \vec{k} \times \vec{E} &= \mp \omega \vec{B}.\end{aligned}\tag{8.5.15}$$

This means that the unit vectors $(\hat{E}, \hat{B}, \mp \hat{k})$ form a right handed basis. Because the wave-vector is perpendicular to the planes of constant phase, the electric and magnetic fields oscillate *within* these planes while the wave itself propagates perpendicular to them (see figure 1).⁵ Our simple solution therefore represents plane, monochromatic electromagnetic waves (light) propagating in a vacuum. We will build on this solution in the future.

Another simple solution, which also happens to be of great interest, is obtained by examining the wave equation in a linear, isotropic conducting medium, *i.e.*, by applying Ohm's law, $\vec{j} = g\vec{E}$ (g is the conductivity of the medium), to write the source currents in terms of the fields themselves. Taking

$$\rho = 0, \quad \vec{j} = \sigma \vec{E} = -g \left(\vec{\nabla} \phi + \frac{\partial \vec{A}}{\partial t} \right),\tag{8.5.16}$$

starting from our original equations (8.2.3) and (8.2.5) and imposing a suitable gauge condition, one finds⁶ a simple equation for both ϕ and \vec{A} if the medium is also non-dispersive:

$$\left[\frac{1}{v^2} \frac{\partial^2}{\partial t^2} - \vec{\nabla}^2 + \mu g \frac{\partial}{\partial t} \right] \psi = 0\tag{8.5.17}$$

where the speed of light in a vacuum ($c = 1/\sqrt{\epsilon_0 \mu_0}$) is replaced by $v = 1/\sqrt{\epsilon \mu}$, the electric permittivity and magnetic susceptibility of the medium being ϵ and μ respectively, and where ψ is either ϕ or \vec{A} . The equation is again homogeneous, differing from the vacuum equation by the first derivative term. It is called the ‘‘Telegrapher’s equation’’. Again, we may seek a harmonic solution of the form

$$\psi(t, \vec{r}) = \psi_0 e^{-i(\omega t \pm \vec{\kappa} \cdot \vec{r})}\tag{8.5.18}$$

but now we find the (dispersion) relation

$$\frac{\omega^2}{v^2} - \vec{\kappa}^2 + i\omega g \mu = 0.\tag{8.5.19}$$

Given that the conductivity, electric permittivity and magnetic susceptibility of the medium are real, it is clear that ω and κ cannot both be real. This makes the solutions a great deal richer than the vacuum plane waves we have examined. They will be examined in some detail in Chapter 7.

⁵Problem: Obtain a solution of the wave equation in spherical coordinates (spherical symmetry) by making the following approximation: neglect the angular dependence, *i.e.*, let $\psi = \psi(r)$. Show that $\psi(r) = \frac{\psi_0}{r} e^{i(kr \pm \omega t)}$.

⁶See Chapter 7

Chapter 9

Space-time Symmetries

9.1 Introduction

Let us now consider the relationship that electromagnetism bears to the special theory of relativity. Recall the electromagnetic field equations

$$\begin{aligned}\frac{1}{c^2} \frac{\partial^2 \phi}{\partial t^2} - \vec{\nabla}^2 \phi &= \frac{\rho}{\epsilon_o} \\ \frac{1}{c^2} \frac{\partial^2 \vec{A}}{\partial t^2} - \vec{\nabla}^2 \vec{A} &= \mu_o \vec{j}\end{aligned}\tag{9.1.1}$$

in Lorentz gauge. We shall think of them as the action of an operator “ \square_x ”,

$$\square_x := \frac{1}{c^2} \frac{\partial^2}{\partial t^2} - \vec{\nabla}^2,\tag{9.1.2}$$

on a set of fields, (ϕ, \vec{A}) , which we collectively designate ψ . So we will write

$$\square_x \psi = f.\tag{9.1.3}$$

Now Newtonian mechanics is based on Galilean transformations: the “Galilean boosts”,

$$\begin{aligned}\vec{r} &\rightarrow \vec{r}' = \vec{r} - \vec{v}t \\ t &\rightarrow t' = t\end{aligned}\tag{9.1.4}$$

(provided that the origin of space and time is chosen appropriately) plus spatial rotations

$$\vec{r}' \rightarrow \vec{r}' = \hat{R} \cdot \vec{r}\tag{9.1.5}$$

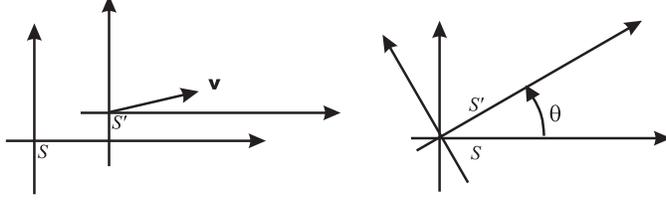


Figure 9.1: Boosts and rotations

where \hat{R} is a rotation matrix (see figure 2).

Consider a single particle within a collection of N particles with interactions between them. If we label the particles by integers, Newton's equations describing the evolution of a single particle, say particle n , may be written as,

$$m_n \frac{d^2 \vec{r}_n}{dt^2} = \vec{F}_n^{ext} + \vec{F}_n^{int} = \vec{F}_n^{ext} + \sum_{m \neq n} \vec{F}_{m \rightarrow n}^{int}, \quad (9.1.6)$$

where $\vec{F}_{m \rightarrow n}^{int}$ represents the (internal) force that particle m exerts over particle n . Assume that the internal forces are derivable from a potential and that the potential depends only on the spatial distance between the particles, *i.e.*,

$$\vec{F}_n^{int} = -\vec{\nabla}_n \Phi_n^{int} = -\sum_{m \neq n} \vec{\nabla}_n \Phi_{nm}(|\vec{r}_n - \vec{r}_m|). \quad (9.1.7)$$

This is compatible with the third law (of action and reaction) and it also makes the internal forces invariant w.r.t Galilean boosts. Thus, provided that the external force, F^{ext} , is invariant under Galilean boosts, the r.h.s. of Newton's equations are invariant. To see that this is so, specialize to just one space dimension and write the transformations in the following form (we are making this more complicated than it really is so as to introduce methods that will be useful in more complicated situations)

$$\begin{bmatrix} dt' \\ dx' \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ -v & 1 \end{bmatrix} \begin{bmatrix} dt \\ dx \end{bmatrix} \quad (9.1.8)$$

and the inverse transformations as

$$\begin{bmatrix} dt \\ dx \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ v & 1 \end{bmatrix} \begin{bmatrix} dt' \\ dx' \end{bmatrix}. \quad (9.1.9)$$

We can now read off

$$\frac{\partial}{\partial t'} = \frac{\partial t}{\partial t'} \frac{\partial}{\partial t} + \frac{\partial x}{\partial t'} \frac{\partial}{\partial x} = \frac{\partial}{\partial t} + v \frac{\partial}{\partial x} \quad (9.1.10)$$

and

$$\frac{\partial}{\partial x'} = \frac{\partial t}{\partial x'} \frac{\partial}{\partial t} + \frac{\partial x}{\partial x'} \frac{\partial}{\partial x} = \frac{\partial}{\partial x}. \quad (9.1.11)$$

Therefore

$$\frac{\partial}{\partial x'_n} \Phi_{nm}(|\vec{r}'_n - \vec{r}'_m|) = \frac{\partial}{\partial x_n} \Phi_{nm}(|\vec{r}_n - \vec{r}_m|), \quad (9.1.12)$$

so the r.h.s. is invariant (we assume that \vec{F}^{ext} is constructed to be invariant). Moreover $dt' = dt$ and the transformation is linear so that the l.h.s. of Newton's equations is also invariant under these transformations. Newtonian dynamics are therefore invariant under Galilean boosts.

9.2 Lorentz Transformations

This is not so for electrodynamics: let us see that the wave equation (even in the vacuum),

$$\frac{1}{c^2} \frac{\partial^2 \psi}{\partial t^2} - \vec{\nabla}^2 \psi = 0, \quad (9.2.1)$$

is not invariant under Galilean transformations. It is an experimental fact (the famous experiment of Michelson and Morley) that the speed of light in a vacuum is constant, independent of the observer. Using the transformations in (9.1.8) and (9.1.9) we have

$$\frac{\partial^2}{\partial t'^2} = \left(\frac{\partial}{\partial t} + \vec{v} \cdot \vec{\nabla} \right) \left(\frac{\partial}{\partial t} + \vec{v} \cdot \vec{\nabla} \right) \quad (9.2.2)$$

and

$$\vec{\nabla}'^2 = \vec{\nabla}^2. \quad (9.2.3)$$

Plugging this into the wave equation, we find

$$\frac{1}{c^2} \frac{\partial^2}{\partial t'^2} - \vec{\nabla}'^2 \rightarrow \frac{1}{c^2} \frac{\partial^2}{\partial t^2} - \vec{\nabla}^2 + \frac{2\vec{v}}{c^2} \cdot \vec{\nabla} \frac{\partial}{\partial t} + \frac{1}{c^2} (\vec{v} \cdot \vec{\nabla})(\vec{v} \cdot \vec{\nabla}). \quad (9.2.4)$$

Only the first two terms on the r.h.s. belong to the wave-equation. Moreover, there exists no known kinetic transformation of the wave-function that can lead to a return to the original form.¹ We conclude that the electromagnetic wave-equation is not invariant under Galilean transformations. This is either a disaster for electromagnetism or a disaster for

¹Problem: Show that, on the contrary, the Schroedinger equation is invariant under Galilean transformations if they are supplemented with the following kinetic transformation of the wave-function:

$$\psi \rightarrow \psi' = e^{-\frac{i}{\hbar}(\vec{p} \cdot \vec{r} - Et)} \psi$$

where $\vec{p} = m\vec{v}$ and $E = m\vec{v}^2/2$. What does this mean?

Newtonian mechanics. As we know, for very good reasons, Maxwell's theory was preferred over Newtonian mechanics and Galilean transformations. We must now ask: what are the transformations that keep Maxwell's equations *form* invariant (we will explain more precisely what *form* invariance means later).

To answer this question we will think of the wave-equation as made up of two distinct parts: the operator " \square_x " and the fields, which we collectively call ψ . Now " \square_x " is a second order operator and we will require it to be a scalar (invariant). Later we will see that for the equation itself to remain form invariant, the fields have to transform as scalars, vectors or tensors. This will imply certain transformation properties of the potentials (ϕ, \vec{A}) .

Let us work with Cartesian systems and consider some general transformations of the form

$$\begin{aligned} t &\rightarrow t' = t'(t, \vec{r}), \\ \vec{r} &\rightarrow \vec{r}' = \vec{r}'(t, \vec{r}). \end{aligned} \quad (9.2.5)$$

They must be

1. one-to-one: so that observers may be able to uniquely relate observations, and
2. invertible: so that the transformations can be made from any observer to the other – there is no preferred observer.

Our functions must therefore be bijective. Let us assume that they are linear,

$$\begin{aligned} t' &= -\frac{1}{c^2}(L_{00}t + \sum_i L_{0i}x_i), \\ x'_i &= L_{i0}t + \sum_j L_{ij}x_j. \end{aligned} \quad (9.2.6)$$

The reason for this peculiar definition will become clear later. For now let us only note that the L 's are some constants that we would like to evaluate. In matrix form the transformations could be written as

$$\begin{bmatrix} dt' \\ dx'_i \end{bmatrix} = \begin{bmatrix} -\frac{L_{00}}{c^2} & -\frac{L_{0j}}{c^2} \\ L_{i0} & L_{ij} \end{bmatrix} \begin{bmatrix} dt \\ dx_j \end{bmatrix}. \quad (9.2.7)$$

The matrix on the r.h.s. is really a 4×4 matrix and L_{ij} represents a 3×3 matrix (the spatial part). It must be invertible.

Take $L_{00} = -c^2$ and $L_{0i} = 0 = L_{i0}$. The resulting transformations are pure spatial rotations, transforming $x_i \rightarrow x'_i = \sum_j L_{ij}x_j$ and leaving $t \rightarrow t' = t$ unchanged. Clearly the wave-operator is a scalar under spatial rotations because in this case $\vec{\nabla}'^2 = \vec{\nabla}^2$ and $\partial_{t'}^2 = \partial_t^2$.

More interesting are the “boosts”. These involve inertial observers with relative velocities. Now $L_{i0} \neq 0 \neq L_{0i}$. Consider relative velocities along the x direction and the transformation

$$\begin{bmatrix} dt' \\ dx'_1 \\ dx'_2 \\ dx'_3 \end{bmatrix} = \begin{bmatrix} \alpha & \beta & 0 & 0 \\ \gamma & \delta & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} dt \\ dx_1 \\ dx_2 \\ dx_3 \end{bmatrix}. \quad (9.2.8)$$

Notice that $x'_2 = x_2$ and $x'_3 = x_3$. This is because we have assumed that space is homogeneous and isotropic so that a boost in the x_1 direction has no effect on the orthogonal coordinates x_2 and x_3 . We can consider then only the effective two dimensional matrix

$$\begin{bmatrix} dt' \\ dx' \end{bmatrix} = \begin{bmatrix} \alpha & \beta \\ \gamma & \delta \end{bmatrix} \begin{bmatrix} dt \\ dx \end{bmatrix} \quad (9.2.9)$$

(where $x_1 := x$). Thus we find the inverse transformation

$$\begin{bmatrix} dt \\ dx \end{bmatrix} = \frac{1}{\|\|} \begin{bmatrix} \delta & -\beta \\ -\gamma & \alpha \end{bmatrix} \begin{bmatrix} dt' \\ dx' \end{bmatrix}, \quad (9.2.10)$$

where $\|\|$ represents the determinant of the transformation, $\|\| = \alpha\delta - \beta\gamma$. Thus we have

$$\begin{aligned} \frac{\partial}{\partial t'} &= \frac{\partial t}{\partial t'} \frac{\partial}{\partial t} + \frac{\partial x}{\partial t'} \frac{\partial}{\partial x} = \frac{1}{\|\|} \left(+\delta \frac{\partial}{\partial t} - \gamma \frac{\partial}{\partial x} \right) \\ \frac{\partial}{\partial x'} &= \frac{\partial t}{\partial x'} \frac{\partial}{\partial t} + \frac{\partial x}{\partial x'} \frac{\partial}{\partial x} = \frac{1}{\|\|} \left(-\beta \frac{\partial}{\partial t} + \alpha \frac{\partial}{\partial x} \right), \end{aligned} \quad (9.2.11)$$

and our wave-operator becomes

$$\begin{aligned} \frac{1}{c^2} \frac{\partial^2}{\partial t'^2} - \vec{\nabla}'^2 &= \frac{1}{\|\|^2} \left(\frac{1}{c^2} \left(+\delta \frac{\partial}{\partial t} - \gamma \frac{\partial}{\partial x} \right)^2 - \left(-\beta \frac{\partial}{\partial t} + \alpha \frac{\partial}{\partial x} \right)^2 \right) \\ &= \frac{1}{\|\|^2} \left((\delta^2/c^2 - \beta^2) \frac{\partial^2}{\partial t^2} - (\alpha^2 - \gamma^2/c^2) \frac{\partial^2}{\partial x^2} \right. \\ &\quad \left. - 2(\alpha\beta - \gamma\delta/c^2) \frac{\partial^2}{\partial t \partial x} \right). \end{aligned} \quad (9.2.12)$$

To return to the wave-operator in the frame S we need to set

$$\begin{aligned} \frac{\delta^2}{c^2} - \beta^2 &= \frac{\|\|^2}{c^2}, \\ \alpha^2 - \frac{\gamma^2}{c^2} &= \|\|^2, \\ \alpha\beta - \frac{\gamma\delta}{c^2} &= 0. \end{aligned} \quad (9.2.13)$$

We have four unknowns and three constraints, so there is really just one parameter that determines all the unknowns. It is easy to find. Note that setting

$$\delta = \gamma \cosh \eta, \quad \beta = \frac{\gamma}{c} \sinh \eta \quad (9.2.14)$$

solves the first of these equations, as

$$\alpha = \gamma \cosh \omega, \quad \gamma = c \gamma \sinh \omega \quad (9.2.15)$$

solves the second. The last equation is then a relationship between η and ω . It implies that

$$\sinh \eta \cosh \omega - \sinh \omega \cosh \eta = \sinh(\eta - \omega) = 0 \rightarrow \eta = \omega. \quad (9.2.16)$$

Our boost in the x direction therefore looks like

$$\begin{bmatrix} dt' \\ dx' \end{bmatrix} = \gamma \begin{bmatrix} \cosh \eta & \frac{1}{c} \sinh \eta \\ c \sinh \eta & \cosh \eta \end{bmatrix} \begin{bmatrix} dt \\ dx \end{bmatrix}. \quad (9.2.17)$$

We notice that γ is not determined. We will henceforth take it to be unity.

What is the meaning of the parameter η ? Consider a test body, which has a velocity u observed in the S frame. Its velocity as measured in the S' frame would be (the velocity does not transform as a vector)

$$u' = \frac{dx'}{dt'} = \frac{(\cosh \eta)dx + c(\sinh \eta)dt}{(\cosh \eta)dt + \frac{1}{c}(\sinh \eta)dx}. \quad (9.2.18)$$

Dividing by $(\cosh \eta)dt$ we find

$$u' = \frac{u + c \tanh \eta}{1 + \frac{u}{c} \tanh \eta}. \quad (9.2.19)$$

Now suppose that the body is at rest in the frame S . This would mean that $u = 0$. But, because S' moves with velocity v relative to S , we can say that S should move with velocity $-v$ relative to S' . Therefore, because the test body is at rest in S , its velocity relative to S' should be $u' = -v$. Our formula gives

$$u' = -v = c \tanh \eta \rightarrow \tanh \eta = -\frac{v}{c}. \quad (9.2.20)$$

This in turn implies that

$$\cosh \eta = \frac{1}{\sqrt{1 - v^2/c^2}}, \quad \sinh \eta = -\frac{v/c}{\sqrt{1 - v^2/c^2}}, \quad (9.2.21)$$

so we can write the transformations in a recognizable form

$$t' = \frac{t - vx/c^2}{\sqrt{1 - v^2/c^2}},$$

$$\begin{aligned}
 x' &= \frac{x - vt}{\sqrt{1 - v^2/c^2}}, \\
 y' &= y, \\
 z' &= z.
 \end{aligned}
 \tag{9.2.22}$$

Notes:

- These are the Lorentz transformations of the special theory of relativity.² They reduce to Galilean transformations when $v/c \ll 1$.
- Because $\tanh \eta \in (-1, 1)$ it follows that the transformations are valid only for $v < c$. The velocity of light is the *limiting* velocity of material bodies and observers. There exists no transformation from the rest frame of light to the rest frame of a material body.
- In general the matrix \hat{L} is made up of boosts *and* rotations. Rotations do not, in general, commute with boosts and two boosts can lead to an overall rotation.

9.3 Tensors on the fly

One lesson that we learn is that we must work with the position vectors of *events* and these are “four-vectors”, *i.e.*, vectors having one time and three space components. It is no longer useful or even correct to think of space and time as separate entities because the Lorentz transformations mix the two. Continuing with a Cartesian system, label the coordinates as follows:

$$x^\mu = (x^0, x^i) : \quad \mu \in \{0, 1, 2, 3\}, \quad x^0 = t, \quad x^i = x_i. \tag{9.3.1}$$

Let us be particular about the position of the indices as superscripts, distinguishing between superscripts and subscripts (soon we will see that this is important) and consider a

²For the very curious: Lorentz transformations can be put in four categories:

- Proper orthochronous: L_+^\uparrow with $||| = +1, L^0_0 \geq +1$
- Proper non-orthochronous: L_+^\downarrow with $||| = +1, L^0_0 \leq -1$
- Improper orthochronous: L_-^\uparrow with $||| = -1, L^0_0 \geq +1$
- Improper non-orthochronous: L_-^\downarrow with $||| = -1, L^0_0 \leq -1$

What we have are therefore proper orthochronous transformations.

displacement, dx^μ , in frame S letting the corresponding displacement in frame S' be dx'^μ . By our transformations we know that

$$dx'^\mu \rightarrow dx^\mu = \sum_{\nu} L^{\mu}_{\nu} dx'^{\nu}, \quad (9.3.2)$$

where L^{μ}_{ν} is precisely the matrix we derived earlier for the special case of boosts in the x direction. In that case

$$\begin{aligned} L^0_0 &= -L_{00}/c^2 = \cosh \eta, \\ L^0_1 &= -L_{01}/c^2 = \sinh \eta/c, & L^0_i &= 0 \quad \forall i \in \{2, 3\}, \\ L^1_0 &= L_{10} = c \sinh \eta, & L^i_0 &= 0 \quad \forall i \in \{2, 3\}, \\ L^1_1 &= L_{11} = \cosh \eta, & L^i_j &= \delta^i_j \quad \forall i, j \in \{2, 3\}, \end{aligned} \quad (9.3.3)$$

where δ^i_j is the usual Kronecker $\hat{\delta}$ (unit matrix),

$$\delta^i_j = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases} \quad (9.3.4)$$

In spacetime, we may set up a vector space V by defining a set of four unit vectors, $\{\hat{u}_{(\mu)}\}$, called a **tetrad frame**, spanning V , so that an arbitrary proper displacement in space-time can be expressed as $d\vec{s} = \sum_{\mu} dx^\mu \hat{u}_{(\mu)}$. Since the displacement itself should not depend on the observer, it follows from (9.3.2) that under a Lorentz transformation

$$\hat{u}_{(\mu)} \rightarrow \hat{u}'_{(\mu)} = \sum_{\alpha} \hat{u}_{(\alpha)} (L^{-1})^{\alpha}_{\mu} \quad (9.3.5)$$

A vector is any object of the form $\vec{A} = \sum_{\mu} A^{\mu} \hat{u}_{(\mu)}$, with four “contravariant” components, A^{μ} , each of which transforms as dx^{μ} (so that \vec{A} is also observer independent), *i.e.*,

$$A^{\mu} \rightarrow A'^{\mu} = \sum_{\nu} L^{\mu}_{\nu} A^{\nu}. \quad (9.3.6)$$

It is both customary and useful to think of a vector in terms of its components, but it is somewhat inconvenient to explicitly write out the summation (Σ) every time we have sum over components. We notice, however, that only repeated indices get summed over; therefore we will use Einstein’s convention and drop the symbol Σ , but now with the understanding that repeated indices, occurring in pairs in which one member appears “up” (as a superscript) and the other “down” (as a subscript), automatically implies a sum. Thus, for example, we would write the above transformation of contravariant vectors as

$$A^{\mu} \rightarrow A'^{\mu} = L^{\mu}_{\nu} A^{\nu}. \quad (9.3.7)$$

Notice that the derivative operator does not transform as dx^μ , but according to the inverse transformation. In other words:

$$\frac{\partial}{\partial x^\mu} := \partial_\mu \rightarrow \frac{\partial}{\partial x'^\mu} := \partial'_\mu = \frac{\partial x^\alpha}{\partial x'^\mu} \partial_\alpha. \quad (9.3.8)$$

But since $\partial x'^\mu / \partial x^\alpha = L^\mu{}_\alpha$, and

$$\frac{\partial x^\alpha}{\partial x'^\mu} \frac{\partial x'^\mu}{\partial x^\beta} = \delta^\alpha{}_\beta = (L^{-1})^\alpha{}_\mu L^\mu{}_\beta, \quad (9.3.9)$$

it follows that

$$\partial'_\mu = \frac{\partial x^\alpha}{\partial x'^\mu} \partial_\alpha = \partial_\alpha (L^{-1})^\alpha{}_\mu, \quad (9.3.10)$$

so the derivatives of a scalar function, $\partial_\mu \phi(x)$, are not the components of a vector. Nevertheless, we see that

$$dx'^\mu \partial'_\mu \phi'(x') = (L^{-1})^\beta{}_\mu L^\mu{}_\alpha dx^\alpha \partial_\beta \phi(x) = \delta^\beta{}_\alpha dx^\alpha \partial_\beta \phi(x) = dx^\mu \partial_\mu \phi(x) \quad (9.3.11)$$

is invariant.

Given any vector space, V , one can consider the space of all linear maps from V to the real numbers, *i.e.*, maps of the form $\vec{\omega} : V \rightarrow \mathbb{R}$,

$$\vec{\omega}(\vec{A}) \stackrel{\text{def}}{=} \vec{\omega} \cdot \vec{A} \in \mathbb{R}$$

satisfying

$$\vec{\omega}(a\vec{A} + b\vec{C}) = a\vec{\omega}(\vec{A}) + b\vec{\omega}(\vec{C}) \quad (9.3.12)$$

where \vec{A} and \vec{C} are in V , and a and b are real numbers. One may now define the sum of two linear maps by

$$(a\vec{\omega} + b\vec{\eta})(\vec{A}) = a\vec{\omega}(\vec{A}) + b\vec{\eta}(\vec{A}) \quad (9.3.13)$$

then it is easy to see that these maps themselves form a vector space, called the **dual** vector space, $*V$. Given the tetrad $\{\hat{u}_{(\mu)}\}$, spanning V , we could introduce a basis for the dual vector space, $\{\hat{\theta}^{(\mu)}\}$, by requiring that

$$\hat{\theta}^{(\nu)}(\hat{u}_{(\mu)}) = \hat{\theta}^{(\nu)} \cdot \hat{u}_{(\mu)} = \delta^\nu{}_\mu \quad (9.3.14)$$

For the inner product to remain invariant, it must hold that, under a Lorentz transformation,

$$\hat{\theta}^{(\mu)} \rightarrow \hat{\theta}'^{(\mu)} = L^\mu{}_\alpha \hat{\theta}^{(\alpha)}. \quad (9.3.15)$$

Any member of the dual vector space, $\vec{\omega}$ can now be expressed as $\vec{\omega} = \omega_\mu \hat{\theta}^{(\mu)}$. ω_μ are called the “covariant” components of $\vec{\omega}$. They will transform as

$$\omega_\mu \rightarrow \omega'_\mu = \omega_\alpha (L^{-1})^\alpha{}_\mu, \quad (9.3.16)$$

so that, given any vector, \vec{A} , and any dual vector, $\vec{\omega}$, one forms a scalar

$$-\vec{\omega} \cdot \vec{A} = -\omega_\mu A^\mu. \quad (9.3.17)$$

This is the four dimensional dot product, the analogue of the three dimensional dot product we are familiar with. The simplest example of a dual vector is the gradient of a scalar function:

$$\nabla = \hat{\theta}^{(\mu)} \partial_\mu \phi(x), \quad (9.3.18)$$

as we saw earlier.

We could generalize the concept of vectors to tensors by simply defining a rank $(0, n)$ tensor to be an multilinear map from a tensor product (an ordered collection) of vectors to \mathbb{R} , *i.e.*, $\mathbb{T} : V \otimes V \dots \otimes V$ (n times) $\rightarrow \mathbb{R}$.³ A basis for \mathbb{T} will evidently be $\hat{\theta}^{(\mu_1)} \otimes \hat{\theta}^{(\mu_2)} \dots \otimes \hat{\theta}^{(\mu_n)}$ and we could express \mathbb{T} as

$$\mathbb{T} = T_{\mu_1 \mu_2 \dots \mu_n} \hat{\theta}^{(\mu_1)} \otimes \hat{\theta}^{(\mu_2)} \dots \otimes \hat{\theta}^{(\mu_n)}, \quad (9.3.19)$$

Its covariant components will transform as n copies of a dual vector,

$$T'_{\mu\nu\lambda\dots} = T_{\alpha\beta\gamma\dots} (L^{-1})^\alpha_\mu (L^{-1})^\beta_\nu (L^{-1})^\gamma_\lambda \dots \quad (9.3.20)$$

Similarly, we could define a rank $(m, 0)$ tensor to be an multilinear map from a tensor product of dual vectors to \mathbb{R} , *i.e.*, $\mathbb{T} : {}^*V \otimes {}^*V \dots \otimes {}^*V$ (m times) $\rightarrow \mathbb{R}$. Following the same reasoning as before, a basis for \mathbb{T} will be $\hat{u}_{(\mu_1)} \otimes \hat{u}_{(\mu_2)} \dots \otimes \hat{u}_{(\mu_m)}$ and we could express \mathbb{T} as

$$\mathbb{T} = T^{\mu_1 \mu_2 \dots \mu_m} \hat{u}_{(\mu_1)} \otimes \hat{u}_{(\mu_2)} \dots \otimes \hat{u}_{(\mu_m)} \quad (9.3.21)$$

so that its contravariant components will transform as m copies of a vector,

$$T^{\mu\nu\lambda\dots} = L^\mu_\alpha L^\nu_\beta L^\lambda_\gamma T^{\alpha\beta\gamma\dots} \quad (9.3.22)$$

More generally, we define “mixed” tensors as multilinear maps from a tensor product of vectors *and* dual vectors, $\mathbb{T} : {}^*V \otimes {}^*V \dots \otimes {}^*V$ (m times) $\otimes V \otimes V \dots \otimes V$ (n times) $\rightarrow \mathbb{R}$ and express it as

$$\mathbb{T} = T^{\mu\nu\dots\lambda\kappa\dots} \hat{u}_{(\mu)} \otimes \hat{u}_{(\nu)} \dots \otimes \hat{\theta}^{(\lambda)} \otimes \hat{\theta}^{(\kappa)} \dots, \quad (9.3.23)$$

with V and *V appearing in any order in the product (the above is simply one example). In this case, the tensor is said to have rank (m, n) . Thus vectors and dual vectors are but special cases of tensors: vectors are tensors of rank $(1, 0)$ and dual vectors are tensors of rank $(0, 1)$. Just as we think of vectors and dual vectors in terms of their components, we will also think of tensors in terms of their components. Thus we will speak of contravariant, covariant and mixed tensors according to their components.

³A multilinear map acts linearly on all its arguments.

There is a one to one relationship between the covariant and contravariant tensors: for every covariant tensor we can find a contravariant tensor and vice-versa. To see how this comes about, let us rewrite the proper distance (??) in a slightly different way, using matrix notation as follows:

$$ds^2 = -d\vec{s} \cdot d\vec{s} = -(\hat{u}_{(\mu)} \cdot \hat{u}_{(\nu)})dx^\mu dx^\nu = -\eta_{\mu\nu}dx^\mu dx^\nu, \quad (9.3.24)$$

or

$$\hat{u}_{(\mu)} \cdot \hat{u}_{(\nu)} = \eta_{\mu\nu} \quad (9.3.25)$$

where, according to (??), $\eta_{\mu\nu}$ is the matrix: $\text{diag}(-c^2, 1, 1, 1)$ *i.e.*,

$$\hat{\eta} = \eta_{\mu\nu} = \begin{bmatrix} -c^2 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}. \quad (9.3.26)$$

It is a covariant tensor of rank two as we see from its transformation properties and is called the Minkowski metric. Given that ds^2 is invariant, we must have

$$-ds^2 = \eta_{\mu\nu}dx^\mu dx^\nu \rightarrow \eta'_{\alpha\beta}dx'^\alpha dx'^\beta = \eta'_{\alpha\beta}L^\alpha{}_\mu L^\beta{}_\nu dx^\mu dx^\nu = \eta_{\mu\nu}dx^\mu dx^\nu, \quad (9.3.27)$$

which implies that

$$\eta_{\mu\nu} = L^\alpha{}_\mu L^\beta{}_\nu \eta'_{\alpha\beta}, \quad (9.3.28)$$

or, by taking inverses,

$$\eta'_{\alpha\beta} = (L^{-1})^\mu{}_\alpha (L^{-1})^\nu{}_\beta \eta_{\mu\nu}. \quad (9.3.29)$$

However, $\eta_{\mu\nu}$ is required to be an invariant tensor in Special Relativity, $\eta'_{\mu\nu} \equiv \eta_{\mu\nu}$, and this can be used in conjunction with (9.3.28) to derive expressions for the matrices \hat{L} . It is an alternative way of deriving the Lorentz transformations through the so-called generators of the transformation (see Appendix B). Now the metric is invertible ($\|\hat{\eta}\| \neq 0$), with inverse

$$\hat{\eta}^{-1} = \eta^{\mu\nu} = \begin{bmatrix} -\frac{1}{c^2} & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (9.3.30)$$

and, by construction, $\eta^{\mu\alpha}\eta_{\alpha\nu} = \delta^\mu{}_\nu$. (It may be easily shown that the inverse metric $\eta^{\mu\nu}$ transforms as

$$\eta'^{\alpha\beta} = L^\alpha{}_\mu L^\beta{}_\nu \eta^{\mu\nu} = \eta^{\alpha\beta}, \quad (9.3.31)$$

i.e., according the rule for a contravariant tensor of rank two.)

In (9.3.24), $\eta_{\mu\nu}$ acts upon two contravariant vectors (two factors of dx^μ) to create a scalar, the proper distance between two events. But we had seen that invariants are constructed from the product of contravariant tensors and covariant tensors. Thus we expect that $\eta_{\mu\nu}dx^\nu$ should transform as a *covariant* vector. In general, consider a contravariant vector A^μ and construct the quantity

$$A_\mu = \eta_{\mu\nu}A^\nu. \quad (9.3.32)$$

How does it transform? We see that

$$A_\mu \rightarrow A'_\mu = \eta_{\mu\nu}L^\nu{}_\alpha A^\alpha = \eta_{\mu\nu}L^\nu{}_\alpha \eta^{\alpha\gamma} \eta_{\gamma\lambda} A^\lambda = (\eta_{\mu\nu}L^\nu{}_\alpha \eta^{\alpha\gamma})(\eta_{\gamma\lambda} A^\lambda), \quad (9.3.33)$$

where we have used $\eta^{\alpha\gamma} \eta_{\gamma\lambda} = \delta^\alpha{}_\lambda$. But notice that (9.3.28) implies the identity

$$\begin{aligned} \eta_{\alpha\beta} &= \eta_{\mu\nu}L^\mu{}_\alpha L^\nu{}_\beta \rightarrow \eta^{\gamma\alpha} \eta_{\alpha\beta} = \eta_{\mu\nu} \eta^{\gamma\alpha} L^\mu{}_\alpha L^\nu{}_\beta \\ \rightarrow \delta^\gamma{}_\beta &= (\eta_{\nu\mu} L^\mu{}_\alpha \eta^{\alpha\gamma}) L^\nu{}_\beta = (L^{-1})^\gamma{}_\nu L^\nu{}_\beta \\ \rightarrow (L^{-1})^\gamma{}_\nu &= \eta_{\nu\mu} L^\mu{}_\alpha \eta^{\alpha\gamma}. \end{aligned} \quad (9.3.34)$$

Therefore (9.3.33) reads

$$A'_\mu = A_\gamma (L^{-1})^\gamma{}_\mu, \quad (9.3.35)$$

which is the transformation of a covariant vector. The Minkowski metric therefore maps contravariant vectors to covariant vectors. In the same way it maps contravariant tensors to covariant tensors:

$$T_{\alpha_1, \alpha_2, \dots, \alpha_n} = \eta_{\alpha_1 \beta_1} \eta_{\alpha_2 \beta_2} \dots \eta_{\alpha_n \beta_n} T^{\beta_1, \beta_2, \dots, \beta_n}. \quad (9.3.36)$$

Likewise, the inverse metric $\eta^{\mu\nu}$ maps covariant vectors to contravariant vectors, *i.e.*, the quantity A^μ defined by

$$A^\mu = \eta^{\mu\nu} A_\nu, \quad (9.3.37)$$

transforms as a contravariant vector.⁴ Therefore, it maps covariant tensors to contravariant tensors:

$$T^{\alpha_1, \alpha_2, \dots, \alpha_n} = \eta^{\alpha_1 \beta_1} \eta^{\alpha_2 \beta_2} \dots \eta^{\alpha_n \beta_n} T_{\beta_1, \beta_2, \dots, \beta_n}. \quad (9.3.38)$$

This relationship between covariant tensors and contravariant tensors is why we originally defined the boosts as in (9.2.7). Thus, $L^\mu{}_\nu = \eta^{\mu\alpha} L_{\alpha\nu}$ which gives

$$L^0{}_0 = \eta^{00} L_{00} = -L_{00}/c^2,$$

⁴Problem: Show this!

$$\begin{aligned}
L^0_i &= \eta^{00}L_{0i} = -L_{0i}/c^2, \\
L^i_0 &= \eta^{ij}L_{j0} = L_{i0}, \\
L^i_j &= \eta^{ik}L_{kj} = L_{ij}.
\end{aligned} \tag{9.3.39}$$

Moreover, there is a natural way to define the (invariant) magnitude of a four-vector, A^μ . It is simply

$$A^2 = -A^\mu A_\mu = -\eta_{\mu\nu}A^\mu A^\nu = -\eta^{\mu\nu}A_\mu A_\nu, \tag{9.3.40}$$

which is the equivalent of the familiar way of defining the magnitude of an ordinary three-vector.⁵ For example, the familiar operator \square_x can be written as

$$\square_x = -\eta^{\mu\nu}\partial_\mu\partial_\nu = -\partial^2, \tag{9.3.41}$$

in which form it is manifestly a scalar. We see once again that the basic difference between Newtonian space and Lorentzian space-time is that, in the case of the former, space and time do not mix and both are absolute. In this case it is sufficient to consider only spatial distances and Pythagoras' theorem ensures that the metric is just the Kronecker δ (with three positive eigenvalues), so there is no need to distinguish between covariant and contravariant indices. In the case of a Lorentzian space-time an observer's measurements of space and time are not independent, neither is absolute and so one is forced to consider the "distance" between *events* in space-time. The metric, $\eta_{\mu\nu}$, for space-time has signature $(-1, 3)$ *i.e.*, it has one negative eigenvalue and three positive eigenvalues.

For an arbitrary boost specified by a velocity $\vec{v} = (v_1, v_2, v_3) = (v^1, v^2, v^3)$, we find the following Lorentz transformations:

$$\begin{aligned}
L^0_0 &= \frac{1}{\sqrt{1 - \vec{v}^2/c^2}} = \gamma, \\
L^i_0 &= \gamma v^i, \\
L^0_i &= \frac{\gamma v_i}{c^2}, \\
L^i_j &= \delta^i_j + (\gamma - 1)\frac{v^i v_j}{\vec{v}^2}.
\end{aligned} \tag{9.3.42}$$

These are most easily derived using (9.3.28) and the fact that $\eta'_{\mu\nu} = \eta_{\mu\nu}$. They reduce to the transformations we had earlier for a boost in the x -direction, for then $\vec{v} = (v, 0, 0)$

⁵When $A^2 < 0$ the vector points within the light cone and is said to be "time-like". When $A^2 > 0$ it points outside the light cone and is called "space-like" and when $A^2 = 0$ the vector A is "light-like" or "null", pointing along the light cone.

and

$$\hat{L} = \begin{bmatrix} \gamma & -\frac{\gamma v}{c^2} & 0 & 0 \\ -\gamma v & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad (9.3.43)$$

which leads to precisely the transformations in (9.2.22). A more compact way to write (9.3.42) is

$$\begin{aligned} t' &= \gamma \left[t - \frac{(\vec{v} \cdot \vec{r})}{c^2} \right] \\ \vec{r}' &= \vec{r} - \gamma \vec{v} t + (\gamma - 1) \frac{\vec{v}}{v^2} (\vec{v} \cdot \vec{r}) \end{aligned} \quad (9.3.44)$$

for a general \vec{v} .

Spatial volume elements are not invariant under Lorentz transformations. We can make a rough argument for this as follows: suppose that the volume measured by the proper observer is dV then the observer moving relative to this observer with a velocity \vec{v} will observe the length dimension in the direction of motion contracted according to (??) and all perpendicular length dimensions will remain unchanged, so we expect $dV' = dV/\gamma$. A more precise treatment follows by mimicking the argument for length contraction. Consider the transformation from $(t, x) \rightarrow (t', x')$

$$dt' = L^0_0 dt + L^0_j dx^j, \quad dx'^i = L^i_0 dt + L^i_j dx^j \quad (9.3.45)$$

with $dt' = 0$ because length measurements must be made subject to a simultaneous measurement of the endpoints in every frame. Therefore $dt = -L^0_j dx^j / \gamma$ and

$$dx'^i = \left(-\frac{1}{\gamma} L^i_0 L^0_j + L^i_j \right) dx^j = \left(\delta^i_j + \frac{(1-\gamma)}{\gamma} \frac{v^i v_j}{v^2} \right) dx^j, \quad (9.3.46)$$

so taking the Jacobian of the transformation gives

$$d^3 \vec{r}' \rightarrow d^3 \vec{r} = d^3 \vec{r}' \left| \frac{\partial x'^i}{\partial x^j} \right| = d^3 \vec{r}' / \gamma. \quad (9.3.47)$$

The four dimensional volume element, $d^4 x$, is invariant for proper Lorentz transformations.

A consequence of the Lorentz transformation of volume is that the three dimensional δ function, $\delta^{(3)}(\vec{r} - \vec{r}_0)$, which is defined according to

$$\int d^3 \vec{r} \delta^{(3)}(\vec{r} - \vec{r}_0) = 1 \quad (9.3.48)$$

cannot be invariant either. If we require the defining integral to remain invariant then

$$\delta^{(3)}(\vec{r} - \vec{r}_0) \rightarrow \delta'^{(3)}(\vec{r}' - \vec{r}'_0) = \gamma \delta^{(3)}(\vec{r} - \vec{r}_0). \quad (9.3.49)$$

The four dimensional delta function, $\delta^{(4)}(x - x_0)$, will, however, be invariant.

9.4 Covariance of Maxwell's equations

A dynamical equation is said to be form invariant, or *covariant*, if both sides of the equation have the same transformation properties. It is a general principle (Einstein's third principle of relativity, based on the notion that all inertial observers are equivalent) that the fundamental dynamical equations of physics must be the same for all inertial observers. We have constructed the wave operator " \square_x " to be a scalar under Lorentz transformations. Therefore, if we want the equation

$$\left[\frac{1}{c^2} \frac{\partial^2}{\partial t^2} - \vec{\nabla}^2 \right] \psi = f \quad (9.4.1)$$

to be covariant both ψ and f must transform as scalars, or vectors, or tensors under the same transformations. If both ψ and f are scalars then covariance is obvious. If both sides transform as, for example, covectors, *i.e.*, ψ is some covector A_μ and f is some covector j_μ then

$$\square'_x A'_\mu = j'_\mu \rightarrow \square_x (L^{-1})^\alpha{}_\mu A_\alpha = (L^{-1})^\alpha{}_\mu j_\alpha \quad (9.4.2)$$

(applying the transformation properties of covectors and the fact that \square_x is constructed to be a scalar) and, therefore,

$$(L^{-1})^\alpha{}_\mu \square_x A_\alpha = (L^{-1})^\alpha{}_\mu j_\alpha, \quad (9.4.3)$$

because L is a constant (global transformation). This is possible if and only if

$$\square_x A_\mu = j_\mu, \quad (9.4.4)$$

so we see that both the observer S' and the observer S use the same equations of motion. The same argument can be applied to any tensor field.

Now there is a natural way to construct a vector field from the potentials (ϕ, \vec{A}) . Define the contravariant 4-vector potential A^μ as follows:

$$A^\mu = \left(\frac{\phi}{c^2}, \vec{A} \right); \quad A^0 = \frac{\phi}{c^2}, \quad A^i = A_i. \quad (9.4.5)$$

Obviously, the corresponding covariant 4-vector potential will be

$$A_\mu = \eta_{\mu\nu} A^\nu = (-\phi, \vec{A}). \quad (9.4.6)$$

Define the second rank antisymmetric tensor (called the Maxwell field strength tensor) as

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu. \quad (9.4.7)$$

It is obviously a covariant tensor because both the partial derivatives and A_μ transform according to the inverse Lorentz transformation, (L^{-1}) .⁶ It is antisymmetric in its indices and, importantly, made up of field gradients. We'll see that its components are just the electric and magnetic fields. Calculate (noting that $F_{00} = 0 = F_{ii}$)

$$F_{0i} = \partial_0 A_i - \partial_i A_0 = \frac{\partial A_i}{\partial t} + \vec{\nabla}_i \phi = -E_i, \quad (9.4.8)$$

and

$$F_{ij} = \partial_i A_j - \partial_j A_i = \epsilon_{ijk} B^k, \quad (9.4.9)$$

where ϵ_{ijk} is the Levi-Civita tensor.⁷ We'll therefore write the matrix $F_{\mu\nu}$ as

$$F_{\mu\nu} = \begin{bmatrix} 0 & -E_1 & -E_2 & -E_3 \\ E_1 & 0 & B_3 & -B_2 \\ E_2 & -B_3 & 0 & B_1 \\ E_3 & B_2 & -B_1 & 0 \end{bmatrix}. \quad (9.4.10)$$

What does the corresponding contravariant tensor look like? Convince yourselves that

$$F^{\mu\nu} = \begin{bmatrix} 0 & E_1/c^2 & E_2/c^2 & E_3/c^2 \\ -E_1/c^2 & 0 & B_3 & -B_2 \\ -E_2/c^2 & -B_3 & 0 & B_1 \\ -E_3/c^2 & B_2 & -B_1 & 0 \end{bmatrix}, \quad (9.4.11)$$

and verify that the quantity $-F^{\mu\nu} F_{\mu\nu}/2 = \vec{E}^2/c^2 - \vec{B}^2$ is a *Lorentz scalar*, *i.e.*, its value is the same for all inertial observers. This is just one example of a field invariant that is constructed from the Maxwell field strength tensor. Likewise the product $*F^{\mu\nu} F_{\mu\nu} \sim \vec{E} \cdot \vec{B}$ is also a Lorentz scalar. Other invariants can be constructed.⁸

It is extremely important to know the invariants of any theory. The fact that they do not change from inertial observer to inertial observer means that they can act as powerful constraints. For example, consider the following question: does there exist an electromagnetic field that is purely electric in one inertial frame and purely magnetic in another? Without knowing the explicit form of the transformations of \vec{E} and \vec{B} we can

⁶Problem: Show explicitly that $F'_{\mu\nu} = (L^{-1})^\alpha{}_\mu (L^{-1})^\beta{}_\nu F_{\alpha\beta}$. This transformation can now be used to derive the transformation properties of the electric and magnetic fields – see Homework (2) problem (2).

⁷Recall that the three dimensional Levi-Civita symbol is defined by

$$\epsilon_{ijk} = \begin{cases} +1, & (i, j, k) \text{ is an even permutation of } (1, 2, 3) \\ -1, & (i, j, k) \text{ is an odd permutation of } (1, 2, 3) \\ 0, & \text{otherwise} \end{cases}$$

Problem: Show therefore that $B^i = \frac{1}{2} \epsilon^{ijk} F_{jk}$.

⁸see Homework (2) problem (1).

deduce that the answer is “no”. Suppose that the field is purely electric (non-vanishing) in frame S and purely magnetic in frame S' , if S' exists. Then we should have

$$\frac{\vec{E}^2}{c^2} = -(\vec{B}')^2 \quad (9.4.12)$$

but the left hand side is ≥ 0 and the right hand side is ≤ 0 . Therefore both must be zero which contradicts our starting assumption that $\vec{E} \neq 0$. To generalize this example a bit, consider a frame S in which there is both an electric field, \vec{E} and a magnetic field, \vec{B} . Is there a frame, S' , in which the electric field would disappear? Again, let us see what the invariants say: we should have

$$\begin{aligned} \vec{E} \cdot \vec{B} &= 0 \\ \frac{\vec{E}^2}{c^2} - \vec{B}^2 &= -(\vec{B}')^2 \end{aligned} \quad (9.4.13)$$

The first equation tells us that this would be possible only if the electric and magnetic fields were perpendicular in the original frame, S . We will return to this point shortly.

Consider the transformations of the electric and magnetic fields. They they do not transform as vectors but as the components of a tensor (the Maxwell field strength tensor, of course). Thus, for example, we could use

$$\begin{aligned} \frac{E'^i}{c^2} &= F'^{0i} = L^0{}_\alpha L^i{}_\beta F^{\alpha\beta} \\ \epsilon^{ijk} B'_k &= F'^{ij} = L^i{}_\alpha L^j{}_\beta F^{\alpha\beta} \end{aligned} \quad (9.4.14)$$

to determine⁹

$$\begin{aligned} \vec{E}' &= \gamma(\vec{E} + \vec{v} \times \vec{B}) - \frac{\gamma^2}{(\gamma + 1)} \frac{\vec{v}}{c^2} (\vec{v} \cdot \vec{E}) \\ \vec{B}' &= \gamma(\vec{B} - \frac{\vec{v}}{c^2} \times \vec{E}) - \frac{\gamma^2}{(\gamma + 1)} \frac{\vec{v}}{c^2} (\vec{v} \cdot \vec{B}) \end{aligned} \quad (9.4.15)$$

They do not transform as the spatial part of the displacement vector x^μ . To develop on the question in the previous paragraph: we saw that it might be possible to eliminate the electric field by a boost if the two fields were perpendicular. For example, consider the plane electromagnetic waves we obtained in section (5). They certainly satisfy the condition $\vec{E} \cdot \vec{B} = 0$. However, they also satisfy $E^2/c^2 = B^2$, so $(\vec{B}')^2$ is also vanishing. Is it possible then to eliminate both the electric and the magnetic field merely by choosing

⁹see homework 2 Problem (2)

a suitable frame? If this were possible, the photon would simply disappear! Suppose that $\vec{E}' = 0 = \vec{B}'$ (in frame S'), then according to (9.4.15),

$$\begin{aligned}\vec{E} + \vec{v} \times \vec{B} &= \frac{\gamma}{(\gamma + 1)} \frac{\vec{v}}{c^2} (\vec{v} \cdot \vec{E}) \\ \vec{B} - \frac{\vec{v}}{c^2} \times \vec{E} &= \frac{\gamma}{(\gamma + 1)} \frac{\vec{v}}{c^2} (\vec{v} \cdot \vec{B})\end{aligned}\quad (9.4.16)$$

But, because $\vec{E} \cdot \vec{B} = 0$, we have $(\vec{v} \cdot \vec{E})(\vec{v} \cdot \vec{B}) = 0$, so the velocity should be perpendicular either to \vec{E} or to \vec{B} . Suppose it is perpendicular to \vec{E} . Then it follows that $\vec{E} = -\vec{v} \times \vec{B}$. But, for electromagnetic waves in the vacuum, $\vec{E} = c\vec{B} \times \hat{k}$, so $\vec{v} = c\hat{k}$. S' is a frame that travels in the direction of propagation at the speed of light. The result is nonsense, of course, because the transformations (9.4.15) are not valid for $v = c$. Had we chosen \vec{v} perpendicular to \vec{B} instead, we would find $c^2\vec{B} = \vec{v} \times \vec{E}$. But we already know that $\hat{k} \times \vec{E} = c\vec{B}$ allowing us to once again identify $\vec{v} = c\hat{k}$. The conclusions are the same: you cannot get rid of a photon by a boost. Of course, this is also obvious from the fact that the equation $\square_x A_\mu = 0$ is Lorentz covariant.

Again, note that gauge transformations of the 4-vector potential can be written as

$$A_\mu \rightarrow A'_\mu = A_\mu + \partial_\mu \Lambda, \quad (9.4.17)$$

which is another nice way to see how the formalism unifies the two potentials and their properties. $F_{\mu\nu}$ remains unchanged under these transformations, *i.e.*, $F_{\mu\nu}$ is invariant under gauge transformations.

Maxwell's equations are first order differential equations in the electric and magnetic fields (or second order in the potentials). Let us therefore evaluate the quantity $\partial_\mu F^{\mu\nu}$. In the first place we notice that this is a contravariant *vector*, for consider its transformation properties:

$$\begin{aligned}\partial_\mu F^{\mu\nu} \rightarrow \partial'_\mu F'^{\mu\nu} &= (L^{-1})^\alpha{}_\mu L^\mu{}_\beta L^\nu{}_\gamma \partial_\alpha F^{\beta\gamma} \\ &= \delta^\alpha{}_\beta L^\nu{}_\gamma \partial_\alpha F^{\beta\gamma} = L^\nu{}_\gamma \partial_\alpha F^{\alpha\gamma}.\end{aligned}\quad (9.4.18)$$

They are of a contravariant vector. Then let us evaluate ($F^{00} = 0$):

$$\partial_\mu F^{\mu 0} = \partial_i F^{i0} = -\frac{1}{c^2} \vec{\nabla} \cdot \vec{E} = -\frac{\rho}{\epsilon_0 c^2} = -\mu_0 \rho \quad (9.4.19)$$

(using $c^2 = 1/\epsilon_0 \mu_0$) and

$$\partial_\mu F^{\mu i} = \partial_0 F^{0i} + \partial_j F^{ji} = \frac{\partial E_i}{\partial t} + \epsilon^{jik} \partial_j B_k$$

$$= - \left[(\vec{\nabla} \times \vec{B})_i - \frac{\partial E_i}{\partial t} \right] = -\mu_0 j_i. \quad (9.4.20)$$

We conclude that

$$\partial_\mu F^{\mu\nu} = -j^\nu, \quad (9.4.21)$$

where j^μ is defined by

$$j^\mu = (j^0, j^i); \quad j^0 = \mu_0 \rho, \quad j^i = \mu_0 j_i. \quad (9.4.22)$$

Moreover, because the l.h.s. of (9.4.21) transforms as a contravariant vector, therefore j^μ should transform as a contravariant vector. Thus we know how the matter fields transform as well. The continuity equation now has the particularly simple form

$$\partial_\mu j^\mu = 0 \quad (9.4.23)$$

and the entire content of Maxwell's dynamics is in the simple equation of (9.4.21). As we have seen, this equation must be supplemented by a gauge condition. The Lorentz gauge condition reads

$$\vec{\nabla} \cdot \vec{A} + \frac{1}{c^2} \frac{\partial \phi}{\partial t} = \partial_\mu A^\mu = 0, \quad (9.4.24)$$

in which form it is manifestly a scalar condition, remaining the same for all inertial observers.

What about the Bianchi identities? Define the “dual” tensor

$$*F^{\mu\nu} = \frac{1}{2} \epsilon^{\mu\nu\alpha\beta} F_{\alpha\beta}, \quad (9.4.25)$$

where $\epsilon^{\mu\nu\alpha\beta}$ is the 4-dimensional Levi-Civita tensor defined in the usual way by

$$\epsilon^{\mu\nu\alpha\beta} = \begin{cases} +1, & (\mu, \nu, \alpha, \beta) \text{ is an even permutation of } (0, 1, 2, 3) \\ -1, & (\mu, \nu, \alpha, \beta) \text{ is an odd permutation of } (0, 1, 2, 3) \\ 0, & \text{otherwise} \end{cases} \quad (9.4.26)$$

The four dimensional Levi-Civita symbol is really a tensor *density*, but in Minkowski space there is no noteworthy difference between a tensor density and a tensor. For our purposes, therefore, the Levi-Civita symbol defined above is a contravariant tensor of rank four.¹⁰ It should be clear from the properties of the Levi-Civita tensor that $*F^{\mu\nu}$ is antisymmetric. Its components are

$$*F^{0i} = \frac{1}{2} \epsilon^{0ijk} F_{jk} = \frac{1}{2} \epsilon^{ijk} F_{jk} = B^i,$$

¹⁰Problem: Prove this! Use the fact that the Lorentz transformation matrices have unit determinant.

$${}^*F^{ij} = \frac{1}{2}[\epsilon^{ij0k}F_{0k} + \epsilon^{ijk0}F_{k0}] = -\epsilon^{ijk}F_{k0} = -\epsilon^{ijk}E_k. \quad (9.4.27)$$

We could write the tensor in matrix form as

$${}^*F^{\mu\nu} = \begin{bmatrix} 0 & B_1 & B_2 & B_3 \\ -B_1 & 0 & -E_3 & E_2 \\ -B_2 & E_3 & 0 & -E_1 \\ -B_3 & -E_2 & E_1 & 0 \end{bmatrix} \quad (9.4.28)$$

and, likewise, the corresponding covariant form, ${}^*F_{\mu\nu} = \eta_{\mu\alpha}\eta_{\nu\beta}{}^*F^{\alpha\beta}$ as

$${}^*F_{\mu\nu} = \begin{bmatrix} 0 & -c^2B_1 & -c^2B_2 & -c^2B_3 \\ c^2B_1 & 0 & -E_3 & E_2 \\ c^2B_2 & E_3 & 0 & -E_1 \\ c^2B_3 & -E_2 & E_1 & 0 \end{bmatrix}. \quad (9.4.29)$$

If we compare the expressions for $F^{\mu\nu}$ and ${}^*F^{\mu\nu}$ we will see that the effect of taking the dual is to transform $\vec{E} \rightarrow c^2\vec{B}$ and $\vec{B} \rightarrow -\vec{E}$. This kind of a transformation is called a duality transformation. Notice now that

$$\partial_\mu {}^*F^{\mu\nu} = \epsilon^{\mu\nu\alpha\beta}\partial_\mu F_{\alpha\beta} = \epsilon^{\mu\nu\alpha\beta}\partial_\mu(\partial_\alpha A_\beta - \partial_\beta A_\alpha) \equiv 0 \quad (9.4.30)$$

(because whereas $\epsilon^{\mu\nu\alpha\beta}$ is antisymmetric w.r.t. interchanges of its indices, the operators $\partial_\mu\partial_\alpha$ and $\partial_\mu\partial_\beta$ are both symmetric). Let us now calculate $\partial_\mu {}^*F^{\mu\nu}$:

$$\begin{aligned} \partial_\mu {}^*F^{\mu 0} &= \partial_i {}^*F^{i0} = -\vec{\nabla} \cdot \vec{B} = 0, \\ \partial_\mu {}^*F^{\mu i} &= \partial_0 {}^*F^{0i} + \partial_j {}^*F^{ji} = \frac{\partial B_i}{\partial t} - \epsilon^{jik}\partial_j E_k \\ &= (\vec{\nabla} \times \vec{E})_i + \frac{\partial B_i}{\partial t} = 0. \end{aligned} \quad (9.4.31)$$

These will be recognized as the Bianchi identities. Thus the Bianchi identities may be written in a single equation as

$$\partial_\mu {}^*F^{\mu\nu} = 0. \quad (9.4.32)$$

Maxwell's equations (all four of them) are therefore¹¹

$$\partial_\mu {}^*F^{\mu\nu} = 0, \quad \text{Bianchi identities,}$$

¹¹For those who want to flex those muscles developed in Mechanics a bit, the following exercises are instructive and highly recommended:

- Show that the l.h.s. of the dynamical equations can be obtained from varying the action $\int d^4x F^{\mu\nu} F_{\mu\nu}$

$$\partial_\mu F^{\mu\nu} = -j^\nu, \quad \text{dynamics.} \quad (9.4.33)$$

Notice that in the vacuum ($j^\mu = 0$), Maxwell's equations are *invariant* under a duality transformation.

9.5 Discrete Symmetries

The Lorentz transformations are continuous because they depend on the continuous parameters v_i , which are simply the components of the velocity of one frame relative to the other. There are also discrete transformations which may be of interest. We will consider, in particular, three discrete transformations, two of them act directly on space-time, the third acts on charge.

9.5.1 Parity

Parity, or space inversion, is the transformation that reflects the spatial coordinates, as if in a mirror, so that

$$\vec{r} \rightarrow -\vec{r}, \quad t \rightarrow t \quad (9.5.1)$$

We can represent this transformation as a matrix operation on space-time, just as we did the Lorentz transformation. Thus

$$x^\mu = P^\mu_\nu x^\nu \quad (9.5.2)$$

-
- Show that the Bianchi identities are obtained from the action $\int d^4x^* F^{\mu\nu} F_{\mu\nu}$. This integral can be written as a total derivative:

$$\int d^4x F^{\mu\nu*} F_{\mu\nu} = 4 \int d^4x \partial_\mu W^\mu = 4 \int_\Sigma d\sigma^\mu W_\mu$$

where

$$W^\mu = \epsilon^{\mu\alpha\beta\gamma} A_\alpha \partial_\beta A_\gamma$$

and the last integration is carried out on the bounding surface, Σ , therefore it depends on the topology. In other words, $\int d^4x F^* F$ is a topological term that can be added to the Lagrangian $\int d^4x F^2$. It does not contribute to the (local) equations of motion but does contribute to the action if non-trivial boundaries are present. The quantity $F^* F$ is an invariant called the Euler invariant.

- Perform a Hamiltonian decomposition of the action and show that ϕ is a Lagrange multiplier field. How many real dynamical fields (functions) does the electromagnetic field involve?
- Use Noether's theorem to obtain the stress-energy tensor of the electromagnetic field, $T^{\mu\nu}$. Is it symmetric? If not, symmetrize it.
- Define the angular momentum density $M^{ij} = x^i T^{0j} - x^j T^{0i}$
- Compute the angular momentum density of the plane wave we had obtained earlier.

These are only for the especially brave! The less courageous will find the answers in Chapter 5.

where

$$P^\mu_\nu = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix} \quad (9.5.3)$$

The spatial velocity and the spatial acceleration clearly get reversed under a parity transformation. We can also determine how the metric tensor transforms under a parity transformation by looking at the line element,

$$ds^2 = \eta_{\mu\nu} dx^\mu dx^\nu \rightarrow \eta_{\mu\nu} P^\mu_\alpha dx^\alpha P^\nu_\beta dx^\beta = (P^T_\alpha{}^\mu \eta_{\mu\nu} P^\nu_\beta) dx^\alpha dx^\beta = \eta'_{\alpha\beta} dx^\alpha dx^\beta \quad (9.5.4)$$

where P^T is the transpose of P . But one verifies by direct computation that

$$\eta'_{\alpha\beta} = (P^T_\alpha{}^\mu \eta_{\mu\nu} P^\nu_\beta) = \eta_{\alpha\beta} \quad (9.5.5)$$

so it is invariant. What about the electromagnetic field? The transformation of the electric and magnetic fields can be inferred directly from Lorentz force on a single charged particle, if we assume that the basic laws of mechanics are invariant under parity transformations. Then

$$\vec{F} = e[\vec{E} + \vec{v} \times \vec{B}] \quad (9.5.6)$$

should stay the same and since, under a parity transformation, $\vec{F} \rightarrow -\vec{F}$ and $\vec{v} \rightarrow -\vec{v}$, all else being unaffected, it follows that the equation remains the same if and only if

$$\begin{aligned} \vec{E} &\rightarrow -\vec{E} \\ \vec{B} &\rightarrow \vec{B} \end{aligned} \quad (9.5.7)$$

One could say that one is interested only in the electric and magnetic fields because that is what we measure. Nevertheless, what about the electromagnetic potential A_μ ? We see that

$$\begin{aligned} \vec{B} &= \vec{\nabla} \times \vec{A} \\ \vec{E} &= -\vec{\nabla}\phi - \frac{\partial \vec{A}}{\partial t} \end{aligned} \quad (9.5.8)$$

are consistent only if $\vec{A} \rightarrow -\vec{A}$ and $\phi \rightarrow \phi$. In other words, A^μ transforms as the coordinate displacement dx^μ under a parity transformation. A^μ is said to be a “vector” field under parity transformations. If it had **not** transformed as dx^μ under parity, it would be called a “pseudovector” field because it would transform as a vector under Lorentz transformations and as a scalar under parity.

9.5.2 Time reversal

Time reversal, as the name suggests, reflects only the time coordinate,

$$\vec{r} \rightarrow \vec{r}, \quad t \rightarrow -t \quad (9.5.9)$$

Again, we can represent this transformation as a matrix operation on the space-time coordinates,

$$x^\mu \rightarrow T^\mu_\nu x^\nu$$

$$T^\mu_\nu = \begin{bmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (9.5.10)$$

Following precisely the same argument as before, we see that the metric tensor is invariant under time reversal and that $\vec{F} \rightarrow \vec{F}$ and $\vec{v} \rightarrow -\vec{v}$ together imply that $\vec{E} \rightarrow \vec{E}$ and $\vec{B} \rightarrow -\vec{B}$. From the definition of these fields in terms of the electromagnetic potential, we see that $\vec{A} \rightarrow -\vec{A}$ and $\phi \rightarrow \phi$, which is precisely the same as it was under parity.

9.5.3 Charge Conjugation

Charge conjugation is a transformation that changes the sign of the charge, leaving the space-time coordinates unchanged

$$x^\mu \rightarrow x^\mu, \quad e \rightarrow -e. \quad (9.5.11)$$

Naturally, the metric is unaffected by charge conjugation. However, the Lorentz force law tells us that $\vec{E} \rightarrow -\vec{E}$ and $\vec{B} \rightarrow -\vec{B}$, which is only possible if $\vec{A} \rightarrow -\vec{A}$ and $\phi \rightarrow -\phi$ or $A^\mu \rightarrow -A^\mu$.

9.6 Electromagnetic Duality

“Electromagnetic Duality” is a notable and curious property of the electromagnetic field in the absence of sources. The vacuum Maxwell equations remain invariant under the continuous transformation¹²

$$\vec{E} - ic\vec{B} \rightarrow e^{i\theta}(\vec{E} - ic\vec{B}), \quad \theta \in [0, 2\pi) \quad (9.6.1)$$

¹²Problem: Check this claim.

which, for $\theta = \frac{\pi}{2}$ in particular, interchanges the electric and magnetic fields according to

$$\vec{E} \rightarrow c\vec{B}, \quad \vec{B} \rightarrow -\frac{\vec{E}}{c} \quad (9.6.2)$$

(recall that $c = 1/\sqrt{\epsilon_o\mu_o}$). This starting symmetry is broken in the presence of sources. However, we can attempt (as did Dirac) to “complete” Maxwell’s equations by introducing sources for the magnetic field as well, so that (8.1.9) turns into

$$\begin{aligned} \vec{\nabla} \cdot \vec{E} &= \mu_o c^2 \rho^e \\ \vec{\nabla} \times \vec{E} + \frac{\partial \vec{B}}{\partial t} &= -\mu_o \vec{j}^m \\ \vec{\nabla} \cdot \vec{B} &= \mu_o \rho^m \\ \vec{\nabla} \times \vec{B} - \epsilon_o \mu_o \frac{\partial \vec{E}}{\partial t} &= \mu_o \vec{j}^e \end{aligned} \quad (9.6.3)$$

where we have used the superscripts “e” and “m” respectively to distinguish between the electric and magnetic charge and current densities. The vacuum duality is restored by transforming the charges and currents as well,

$$\begin{aligned} \left(\rho^e - i \frac{\rho^m}{c} \right) &\rightarrow e^{i\theta} \left(\rho^e - i \frac{\rho^m}{c} \right) \\ \left(\vec{j}^e - i \frac{\vec{j}^m}{c} \right) &\rightarrow e^{i\theta} \left(\vec{j}^e - i \frac{\vec{j}^m}{c} \right) \end{aligned} \quad (9.6.4)$$

A single magnetic charge is called a Dirac monopole. It is easy to see that the magnetic field for a single, static monopole is

$$\vec{B} = \frac{\mu_o g}{4\pi r^2} \hat{r}, \quad (9.6.5)$$

where g is the magnetic charge. Away from the monopole the magnetic field is derivable from a vector potential (since $\vec{\nabla} \cdot \vec{B} = 0$), but the vector potential describing this particular magnetic field,

$$\vec{A} = \frac{\mu_o g}{4\pi r} \tan \frac{\theta}{2} \hat{\phi} \quad (9.6.6)$$

is singular on the half line $\theta = \pi$. This singularity cannot be removed by any gauge transformation. Physically, it corresponds to an infinite solenoid confining a magnetic flux entering the monopole from infinity that equals the outgoing magnetic flux from the monopole. It is called a Dirac “string”. The presence of a Dirac string (the infinite solenoid) would pose a severe threat to the consistency of electromagnetism as a quantum theory due to the fact that wave functions of particles would pick up a phase when surrounding a region

in which a magnetic flux tube is present. This is the “Aharonov-Bohm’ effect. One can show that the problem is circumvented if and only if the electric and magnetic charges satisfy the following condition:

$$eq = \frac{2\pi n}{c}, \quad n \in \mathbb{N}. \quad (9.6.7)$$

Therefore the presence of a single magnetic monopole in the universe would imply that electric charge is quantized in units of $2\pi/gc$. No magnetic monopoles have yet been discovered.

Chapter 10

More general coordinate systems*

10.1 Introduction

As we have seen, very often the symmetries of a given physical system make Cartesian coordinates cumbersome because it may turn out to be difficult to implement the boundary conditions suited to the system in these coordinates. In that case, as we know well, we turn to coordinate systems that are better adapted to the given symmetries. The new coordinates are not usually Cartesian (for example, think of the problem of determining geodesics on a sphere). A generic feature of such systems – and one that is exploited in the physical problem – is that the coordinate surfaces are curved. They are therefore called “curvilinear” systems. In this chapter we will develop some machinery to work with such systems.

Suppose that we perform a coordinate transformation from a set of Cartesian coordinates to a set of curvilinear coordinates that are given by ξ^μ . Imagine that the Cartesian coordinates extend over the entire spacetime and let them be given by x^a . We will assume that the new coordinates are invertible functions of the Cartesian coordinates, *i.e.*, $\xi^\mu = \xi^\mu(x)$ and $x^a = x^a(\xi)$. In the Cartesian system, the distance between two spatial points is given by the 4-dimensional equivalent of Pythagoras’ theorem:

$$ds^2 = -\eta_{ab}dx^a dx^b \quad (10.1.1)$$

But the distance could just as well be expressed in terms of the new coordinates, $\xi^\mu(x)$, so, exploiting Pythagoras’ theorem, which is given in a Cartesian frame, we write (Einstein’s summation convention used throughout)

$$ds^2 = -\eta_{ab}dx^a dx^b = -\eta_{ab} \frac{\partial x^a}{\partial \xi^\mu} \frac{\partial x^b}{\partial \xi^\nu} d\xi^\mu d\xi^\nu = -g_{\mu\nu}(\xi) d\xi^\mu d\xi^\nu \quad (10.1.2)$$

where we have made use of the chain rule:

$$dx^a = \frac{\partial x^a}{\partial \xi^\mu} d\xi^\mu \quad (10.1.3)$$

and called

$$g_{\mu\nu}(\xi) = \eta_{ab} \frac{\partial x^a}{\partial \xi^\mu} \frac{\partial x^b}{\partial \xi^\nu} \quad (10.1.4)$$

This is the metric (it gives the distance between infinitesimally separated points) and it is, in general, a function of the new coordinates. Henceforth throughout this section, we'll use the following notation: indices from the beginning of the alphabet, a, b, c, \dots , will represent a Cartesian basis and greek indices μ, ν, \dots will represent a general (curvilinear) basis. This will serve to distinguish indices that originate in the Cartesian system from those whose origin is in the curvilinear system.

Let us define the matrix

$$e_\mu^a(\xi) = \frac{\partial x^a}{\partial \xi^\mu} \quad (10.1.5)$$

It is a function of position and is called the “vielbein”. A useful way is to think of it as a collection of four vectors, $\{\vec{e}_\mu\} = (\vec{e}_0, \vec{e}_1, \vec{e}_2, \vec{e}_3)$, whose Cartesian components are given by $[\vec{e}_\mu]^a = e_\mu^a$. These are the basis vectors of the new coordinate system $\xi^\mu(x)$. They are (generally) functions of the position because the new basis is not (generally) rigid. If we consider the identity transformations, for example, *i.e.*, $\xi^\mu = (t, x, y, z)$ then it's easy to see that $\vec{e}_0 = (1, 0, 0, 0)$, $\vec{e}_1 = (0, 1, 0, 0)$, $\vec{e}_2 = (0, 0, 1, 0)$ and $\vec{e}_3 = (0, 0, 0, 1)$ or, said in another way, $e_\mu^a = \delta_\mu^a$.

The metric in (10.1.4) can be thought of as the matrix whose components are the (Cartesian) inner products of the \vec{e}_μ , *i.e.*,¹

$$g_{\mu\nu}(\xi) = \eta_{ab} e_\mu^a e_\nu^b = \vec{e}_\mu \cdot \vec{e}_\nu \quad (10.1.6)$$

and it is manifestly a scalar under Lorentz transformations. Also, g_{ij} is invertible, if the transformation $x \rightarrow \xi$ is invertible, in which case we can define the inverse metric by

$$g^{\mu\nu} g_{\nu\kappa} = \delta_\kappa^\mu \quad (10.1.7)$$

and it's easy to see that

$$g^{\mu\nu} = \eta^{ab} \frac{\partial \xi^\mu}{\partial x^a} \frac{\partial \xi^\nu}{\partial x^b} = \vec{E}^\mu \cdot \vec{E}^\nu \quad (10.1.8)$$

where \vec{E}^μ is the vector with (covariant) Cartesian components

$$[\vec{E}^\mu]_a = E_a^\mu = \frac{\partial \xi^\mu}{\partial x^a} \quad (10.1.9)$$

¹Problem: check that the identity transformation leads to $g_{\mu\nu} = \eta_{\mu\nu}$

It is called the *inverse* vielbein because, clearly,

$$e_{\mu}^a E_b^{\mu} = \frac{\partial x^a}{\partial \xi^{\mu}} \frac{\partial \xi^{\mu}}{\partial x^b} = \delta_b^a, \quad e_{\mu}^a E_a^{\nu} = \frac{\partial x^a}{\partial \xi^{\mu}} \frac{\partial \xi^{\nu}}{\partial x^a} = \delta_{\mu}^{\nu} \quad (10.1.10)$$

where we have repeatedly used the chain rule from elementary calculus.

10.2 Vectors and Tensors

Now the set $\{\vec{e}_{\mu}\}$ forms a complete basis in which any four vector can be expanded, *i.e.*,

$$\mathbb{A} = A^{\mu} \vec{e}_{\mu} \quad (10.2.1)$$

and A^{μ} are the contravariant components of the vector in the directions given by the \vec{e}_{μ} . We can think of the r.h.s. as giving the contravariant *Cartesian* components of the vector according to

$$A^a = A^{\mu} e_{\mu}^a \quad (10.2.2)$$

in the original frame, while the contravariant components, A^{μ} , in the basis \vec{e}_{μ} , are given in terms of its Cartesian components by the inverse relation

$$A^a E_a^{\mu} = A^{\nu} e_{\nu}^a E_a^{\mu} = A^{\nu} \delta_{\nu}^{\mu} = A^{\mu} \quad (10.2.3)$$

Any vector can be specified by specifying either the components A^a or the components A^{μ} . Naturally, there is also a description in terms of the covariant components of the vector using the inverse vielbeins:

$$\mathbb{A} = A_{\mu} \vec{E}^{\mu} \quad (10.2.4)$$

and the r.h.s. may be thought of as specifying the covariant Cartesian components of \mathbb{A} ,

$$A_a = A_{\mu} E_a^{\mu}, \quad (10.2.5)$$

its covariant components, A_{μ} , in the basis \vec{E}^{μ} being given by

$$A_{\mu} = e_{\mu}^a A_a \quad (10.2.6)$$

in keeping with the relations for contravariant vectors.

We must next consider the transformation properties of the components, Cartesian or curvilinear. Note that there are two kinds of transformations to think about, *viz.*, Lorentz transformations, which concern the original Cartesian basis and general coordinate transformations, which concern the new (curvilinear) basis.

- Lorentz transformations act on the Cartesian coordinates, x^a , so, when $x^a \rightarrow x'^a$,

$$e'^a_\mu = \frac{\partial x'^a}{\partial \xi^\mu} = L^a_b \frac{\partial x^b}{\partial \xi^\mu} = L^a_b e^a_\mu \quad (10.2.7)$$

showing that e^a_μ transforms as a contravariant vector w.r.t. Lorentz transformations. On the other hand, E^μ_a will transform as a covariant vector

$$E'^\mu_a = \frac{\partial \xi^\mu}{\partial x'^a} = (L^{-1})^b_a \frac{\partial \xi^\mu}{\partial x^b} = (L^{-1})^b_a E^\mu_b \quad (10.2.8)$$

w.r.t. the same transformations

- Under general coordinate transformations, $\xi^\mu \rightarrow \xi'^\mu$,

$$e'^a_\mu = \frac{\partial x^a}{\partial \xi'^\mu} = \frac{\partial \xi^\nu}{\partial \xi'^\mu} \frac{\partial x^a}{\partial \xi^\nu} = \frac{\partial \xi^\nu}{\partial \xi'^\mu} e^a_\nu = (\Lambda^{-1})^\nu_\mu e^a_\nu \quad (10.2.9)$$

or

$$\vec{e}'_\mu = \frac{\partial \xi^\nu}{\partial \xi'^\mu} \vec{e}_\nu = (\Lambda^{-1})^\nu_\mu \vec{e}_\nu \quad (10.2.10)$$

and

$$E'^\mu_a = \frac{\partial \xi'^\mu}{\partial x^a} = \frac{\partial \xi'^\mu}{\partial \xi^\nu} \frac{\partial \xi^\nu}{\partial x^a} = \frac{\partial \xi'^\mu}{\partial \xi^\nu} E^\nu_a = \Lambda^\mu_\nu E^\nu_a \quad (10.2.11)$$

or

$$\vec{E}'^\mu = \frac{\partial \xi'^\mu}{\partial \xi^\nu} \vec{E}^\nu = \Lambda^\mu_\nu \vec{E}^\nu \quad (10.2.12)$$

Note that $\hat{\Lambda}$, unlike \hat{L} , is **not** necessarily a constant matrix. These transformation properties imply that the metric in (10.1.4) transforms (under coordinate transformations) as

$$g'_{\mu\nu} = (\Lambda^{-1})^\alpha_\mu (\Lambda^{-1})^\beta_\nu g_{\alpha\beta} \quad (10.2.13)$$

and

$$g'^{\mu\nu} = \Lambda^\mu_\alpha \Lambda^\nu_\beta g^{\alpha\beta} \quad (10.2.14)$$

but is a Lorentz *scalar*.

It should be clear that the Cartesian components, A^a (A_a) transform only under Lorentz transformations (they are “coordinate scalars”) and the components A^μ (A_μ) transform only under general coordinate transformations (they are “Lorentz scalars”). How do they transform? They follow the same rules as the Cartesian components, but transform under $\hat{\Lambda}$ instead of \hat{L} . As a vector does not depend on the basis in which it is expanded,

$$\mathbb{A} = A'^\mu \vec{e}'_\mu = A'^\mu (\Lambda^{-1})^\nu_\mu \vec{e}_\nu = A^\nu \vec{e}_\nu \quad (10.2.15)$$

implying obviously that

$$A'^{\mu} = \Lambda^{\mu}_{\nu} A^{\nu} \quad (10.2.16)$$

and a completely analogous argument shows that

$$A'_{\mu} = (\Lambda^{-1})^{\nu}_{\mu} A_{\nu} \quad (10.2.17)$$

is the transformation property of the covariant components. The contravariant components and the covariant components transform inversely to one another, so they must be related by the metric

$$\begin{aligned} A_{\mu} &= g_{\mu\nu} A^{\nu} \\ A^{\mu} &= g^{\mu\nu} A_{\nu} \end{aligned} \quad (10.2.18)$$

because the metric and its inverse have precisely the transformation properties required. Moreover, it's easy to see now how we might construct scalars (under both Lorentz and general coordinate transformations):

$$\mathbb{A}^2 = \mathbb{A} \cdot \mathbb{A} = (A^{\mu} \vec{e}_{\mu}) \cdot (A^{\nu} \vec{e}_{\nu}) = A^{\mu} A^{\nu} \vec{e}_{\mu} \cdot \vec{e}_{\nu} = g_{\mu\nu} A^{\mu} A^{\nu} = A^{\mu} A_{\mu} \quad (10.2.19)$$

That $\mathbb{A} \cdot \mathbb{A}$ is really a scalar follows from

$$g_{\mu\nu} A^{\mu} A^{\nu} = \eta_{ab} e_{\mu}^a e_{\nu}^b A^{\mu} A^{\nu} = \eta_{ab} A^a A^b = A^a A_a \quad (10.2.20)$$

or directly from

$$A'^{\mu} A'_{\mu} = \Lambda^{\mu}_{\alpha} (\Lambda^{-1})^{\beta}_{\mu} A^{\alpha} A_{\beta} = \delta^{\beta}_{\alpha} A^{\alpha} A_{\beta} = A^{\alpha} A_{\alpha} \quad (10.2.21)$$

As usual we will define tensors as copies of vectors, their components in any basis being given by

$$\mathbb{T} = T^{\mu\nu\lambda\dots} \vec{e}_{\mu} \otimes \vec{e}_{\nu} \otimes \vec{e}_{\lambda} \dots = T_{\mu\nu\lambda\dots} \vec{E}^{\mu} \otimes \vec{E}^{\nu} \otimes \vec{E}^{\lambda} \dots \quad (10.2.22)$$

where $T^{\mu\nu\lambda\dots}$ and $T_{\mu\nu\lambda\dots}$ are the contravariant and covariant components of \mathbb{T} respectively. Then their transformation properties are given by

$$T'^{\mu\nu\lambda\dots} = \Lambda^{\mu}_{\alpha} \Lambda^{\nu}_{\beta} \Lambda^{\lambda}_{\gamma} T^{\alpha\beta\gamma\dots} \quad (10.2.23)$$

and

$$T'_{\mu\nu\lambda\dots} = (\Lambda^{-1})^{\alpha}_{\mu} (\Lambda^{-1})^{\beta}_{\nu} (\Lambda^{-1})^{\gamma}_{\lambda} T_{\alpha\beta\gamma\dots} \quad (10.2.24)$$

respectively. Just as for vectors, the covariant and contravariant components of a tensor are related by the metric (tensor):

$$T^{\mu\nu\lambda\dots} = g^{\mu\alpha} g^{\nu\beta} g^{\lambda\gamma} \dots T_{\alpha\beta\gamma\dots} \quad (10.2.25)$$

and

$$T_{\mu\nu\lambda\dots} = g_{\mu\alpha}g_{\nu\beta}g_{\lambda\gamma}\dots T^{\alpha\beta\gamma\dots} \quad (10.2.26)$$

and one can interpolate between components in the original Cartesian basis and in the curvilinear basis by simply applying the vielbein and its inverse, just as we did for vectors

$$\begin{aligned} T^{ab\dots} &= e_{\mu}^a e_{\nu}^b \dots T^{\mu\nu\dots}, & T^{\mu\nu\dots} &= E_a^{\mu} E_b^{\nu} \dots T^{ab\dots} \\ T_{ab\dots} &= E_a^{\mu} E_b^{\nu} \dots T_{\mu\nu\dots}, & T_{\mu\nu\dots} &= e_{\mu}^a e_{\nu}^b \dots T^{ab\dots} \end{aligned} \quad (10.2.27)$$

10.3 Differentiation

In differentiating a tensor, we are generally interested in measuring the rates of change of the tensor as we move from point to point on the manifold. To do so we measure the difference between the values of the tensor at infinitesimally close points, finally taking the limit as the points approach each other. However, depending on what differences we measure, the resulting rate of change may not have definite transformation properties under general coordinate transformations. Below we will consider two ways to define the “derivative” of a tensor so that the derivative is itself a tensor.

10.3.1 Lie Derivative

Often we may be interested in how the components of a given vector or tensor or even just a function change(s) as we move along some curve, $\xi^{\mu}(\lambda)$, parametrized by λ , from a point p to another point p' . To define the Lie derivative, we consider a special set of curves, which are constructed from coordinate transformations. Consider a one parameter family of coordinate transformations $\xi'^{\mu}(\lambda, \xi)$ so that the $\lambda = 0$ transformation is just the identity transformation, $\xi'(0, \xi) = \xi$. Let the coordinates of point p be ξ_p^{μ} . Holding ξ_p^{μ} fixed, $\xi'^{\mu}(\lambda, \xi_p)$ represents the a curve passing through p at $\lambda = 0$. Suppose that we have chosen our one parameter family of transformations so that the curve $\xi'^{\mu}(\lambda, \xi_p)$ passes through p' at $\delta\lambda$. Let $U^{\mu}(\lambda, \xi_p)$ be tangent to the curve. The point p' is therefore represented by

$$\xi'^{\mu} = \xi'^{\mu}(\delta\lambda, \xi_p) = \xi_p^{\mu} + \delta\lambda U^{\mu}(\xi_p) \quad (10.3.1)$$

This is the “active” view of coordinate transformations, where they are used to actually “push points around”. If \mathbb{T} is a tensor, it transforms as

$$T'^{\mu_1\mu_2\dots\nu_1\nu_2\dots}(\xi') = \frac{\partial \xi'^{\mu_1}}{\partial \xi^{\alpha_1}} \dots \frac{\partial \xi'^{\beta_1}}{\partial \xi'^{\nu_1}} \dots T^{\alpha_1\alpha_2\dots\beta_1\beta_2\dots}(\xi) \quad (10.3.2)$$

and therefore the left hand side of the above should give the value of \mathbb{T} at p' from its value at p . The Lie derivative of \mathbb{T} is then defined as

$$[\mathcal{L}_U \mathbb{T}]^{\mu_1\mu_2\dots\nu_1\nu_2\dots} = \lim_{\delta\lambda \rightarrow 0} \frac{1}{\delta\lambda} [T^{\mu_1\mu_2\dots\nu_1\nu_2\dots}(\xi) - T'^{\mu_1\mu_2\dots\nu_1\nu_2\dots}(\xi)] \quad (10.3.3)$$

It measures the rate of change of the *functional form of the components* of a tensor field by a coordinate transformation, in the direction of U .

For scalar functions, we see immediately that this is just the directional derivative, for if \mathbb{T} is a scalar function, $f(\xi)$, then $f'(\xi') = f(\xi) \Rightarrow f(\xi) - f'(\xi) = \delta\lambda U^\mu \partial_\mu f$ (to order λ), therefore

$$\mathcal{L}_U f(x) = \lim_{\delta\lambda \rightarrow 0} \frac{1}{\delta\lambda} [f(\xi) - f'(\xi)] = U^\mu \partial_\mu f(\xi) \quad (10.3.4)$$

If \mathbb{T} is a vector field, $V^\mu(\xi)$, then

$$\begin{aligned} [\mathcal{L}_U V]^\mu &= \lim_{\delta\lambda \rightarrow 0} \frac{1}{\delta\lambda} [V^\mu(\xi) - V'^\mu(\xi)] \\ &= \lim_{\delta\lambda \rightarrow 0} \frac{1}{\delta\lambda} \left[\frac{\partial \xi^\mu}{\partial \xi'^\alpha} V'^\alpha(\xi') - V'^\mu(\xi) \right] \\ &= \lim_{\delta\lambda \rightarrow 0} \frac{1}{\delta\lambda} \left[\frac{\partial \xi^\mu}{\partial \xi'^\alpha} (V'^\alpha(\xi) + \delta\lambda U^\kappa \partial_\kappa V^\alpha + \dots) - V'^\mu(\xi) \right] \\ &= \lim_{\delta\lambda \rightarrow 0} \frac{1}{\delta\lambda} [(\delta_\alpha^\mu - \delta\lambda \partial_\alpha U^\mu + \dots)(V'^\alpha(\xi) + \delta\lambda U^\kappa \partial_\kappa V^\alpha + \dots) - V'^\mu(\xi)] \\ &= U^\kappa \partial_\kappa V^\mu - V^\kappa \partial_\kappa U^\mu \end{aligned} \quad (10.3.5)$$

Finally, for a co-vector field, $W_\mu(\xi)$

$$\begin{aligned} \mathcal{L}_U W &= \lim_{\delta\lambda \rightarrow 0} \frac{1}{\delta\lambda} [W_\mu(\xi) - W'_\mu(\xi)] \\ &= \lim_{\delta\lambda \rightarrow 0} \frac{1}{\delta\lambda} \left[\frac{\partial \xi'^\alpha}{\partial \xi^\mu} W'_\alpha(\xi') - W'_\mu(\xi) \right] \\ &= \lim_{\delta\lambda \rightarrow 0} \frac{1}{\delta\lambda} [(\delta_\mu^\alpha + \delta\lambda \partial_\mu U^\alpha + \dots)(W'_\alpha(\xi) + \delta\lambda U^\kappa \partial_\kappa W_\alpha + \dots) - W'_\mu(\xi)] \\ &= U^\kappa \partial_\kappa W_\mu + W_\alpha \partial_\mu U^\alpha \end{aligned} \quad (10.3.6)$$

and so on for tensors of higher rank.² If $\mathcal{L}_U \mathbb{T} = 0$, then \mathbb{T} does not change as we move along the integral curve of U . In this case, the vector U is called a “symmetry” of \mathbb{T} . Note that the Lie derivative of a tensor field \mathbb{T} is of the same rank as \mathbb{T} itself.

²Obtain the Lie derivative of second rank contravariant, covariant and mixed tensors. In general, the Lie derivative of a mixed tensor takes the form

$$\begin{aligned} [\mathcal{L}_U \mathbb{T}]^{\mu_1 \mu_2 \dots}_{\nu_1 \nu_2 \dots} &= U^\sigma \partial_\sigma T^{\mu_1 \mu_2 \dots}_{\nu_1 \nu_2 \dots} - T^{\sigma \mu_2 \dots}_{\nu_1 \nu_2 \dots} \partial_\sigma U^{\mu_1} - \dots \\ &\quad + T^{\mu_1 \mu_2 \dots}_{\sigma \nu_2 \dots} \partial_{\nu_1} U^\sigma + \dots \end{aligned}$$

where the ellipsis means that we repeat the terms of each index of the same type.

10.3.2 Covariant Derivative: the Connection

The Lie derivative can be thought of as an operator that acts upon a tensor to yield another tensor of the same rank. However, when we think of a derivative, we think of the operator ∂_a (say), which has the effect of increasing the rank of the tensor. Thus, for example, if \mathbb{T} is a rank (m, n) tensor (m contravariant indices and n covariant indices) then $\partial\mathbb{T}$ is a tensor of rank $(m, n + 1)$, if the partial derivative is applied in a Cartesian coordinate system. However, $\partial\mathbb{T}$ is not a tensor in a general coordinate system, as we will see below. We would like to obtain a derivative operator, ∇ , in general curvilinear coordinates that plays the role of ∂ in Cartesian coordinates, *i.e.*, an operator which acts on an (m, n) tensor to give an $(m, n + 1)$ tensor. So let us begin with vectors. Imagine transporting a vector from some point p to some other point p' . The basis vectors (the vielbeins) and their inverses are not necessarily constant during this transport – they would be constant only if the coordinate system is not curvilinear. Instead of asking about changes in the components of a vector \mathbb{A} , let's ask instead how the vector as a whole changes as we move from p to p' . We find

$$\delta\mathbb{A} = (\delta A^\mu)\vec{e}_\mu + A^\mu(\delta\vec{e}_\mu) \quad (10.3.7)$$

Assume that the change in \vec{e}_μ is a linear combination of the \vec{e}_ν themselves. This is a reasonable assumption because the basis at p is complete. Then

$$\delta\vec{e}_\mu = (\delta\Gamma_\mu^\nu)\vec{e}_\nu \quad (10.3.8)$$

and $\Gamma_\mu^\nu \equiv \Gamma_\mu^\nu(x)$ will be in general a function of the position. In fact we can obtain $\delta\Gamma_\mu^\nu$ in terms of \vec{e}_μ and \vec{E}^μ as follows: begin with

$$\vec{e}_\mu \cdot \vec{E}^\sigma = \delta_\mu^\sigma \Rightarrow (\delta\vec{e}_\mu) \cdot \vec{E}^\sigma = -\vec{e}_\mu \cdot (\delta\vec{E}^\sigma) \quad (10.3.9)$$

but, using (10.3.8), we see that

$$\delta\Gamma_\mu^\nu(\vec{e}_\nu \cdot \vec{E}^\sigma) = \delta\Gamma_\mu^\sigma = -\vec{e}_\mu \cdot (\delta\vec{E}^\sigma) = (\delta\vec{e}_\mu) \cdot \vec{E}^\sigma \quad (10.3.10)$$

We can write

$$\delta\mathbb{A} = (\delta A^\mu)\vec{e}_\mu + A^\mu(\delta\Gamma_\mu^\nu)\vec{e}_\nu \quad (10.3.11)$$

and we see that the change in \mathbb{A} is made up of two parts: (i) the change in its components and (ii) the changing the basis, as we move from one point to another. The term $\delta\Gamma$ takes into account the change in basis. The difference, $\delta\mathbb{A}$, is also a vector and is expandable in the basis \vec{e}_μ . If we consequently write it as $\delta\mathbb{A} = (DA^\mu)\vec{e}_\mu$, we find

$$DA^\mu = \delta A^\mu + (\delta\Gamma_\mu^\nu)A^\nu \quad (10.3.12)$$

but what does DA^μ represent? Notice that if the basis is rigid then $\delta\vec{e}_\mu = 0 = \delta\Gamma_\mu^\nu$ and there is no difference between the variations DA^μ and δA^μ . This equality fails in a general

coordinate system, however, and the second term is important. The derivative corresponding to the infinitesimal change given in (10.3.12) is called the “covariant derivative” of A^μ ,

$$D_\nu A^\mu \equiv \nabla_\nu A^\mu = \partial_\nu A^\mu + (\partial_\nu \Gamma_\lambda^\mu) A^\lambda = \partial_\nu A^\mu + \Gamma_{\nu\lambda}^\mu A^\lambda \quad (10.3.13)$$

and the 3-index object $\Gamma_{\nu\lambda}^\mu = \partial_\nu \Gamma_\lambda^\mu$ is called a “connection”. Using (10.3.10), it can be written as

$$\Gamma_{\nu\lambda}^\mu = (\partial_\nu \vec{e}_\lambda) \cdot \vec{E}^\mu. \quad (10.3.14)$$

It is interesting to see that the ordinary derivative, $\partial_\nu A^\mu$ of a contravariant vector does not transform as a (mixed) tensor, but the covariant derivative, $\nabla_\nu A^\mu$, does. The fact that the covariant derivative transforms as a tensor is of great importance. As we have mentioned, the laws of physics should not depend on one’s choice of coordinates. This means that they should “look the same” in any system, which is possible only if the two sides of any dynamical equation transform in the same manner, *i.e.*, either as scalars, vectors or tensors under transformations between coordinate systems. Thus, covariant derivatives and not ordinary derivatives are more meaningful in physics.

First let’s see that $\partial_\nu A^\mu$ is not a tensor:

$$\frac{\partial A'^\mu}{\partial \xi'^\nu} = \frac{\partial \xi^\lambda}{\partial \xi'^\nu} \frac{\partial}{\partial \xi^\lambda} \left(\frac{\partial \xi'^\mu}{\partial \xi^\kappa} A^\kappa \right) = \frac{\partial \xi^\lambda}{\partial \xi'^\nu} \frac{\partial \xi'^\mu}{\partial \xi^\kappa} \frac{\partial A^\kappa}{\partial \xi^\lambda} + \frac{\partial \xi^\lambda}{\partial \xi'^\nu} \frac{\partial^2 \xi'^\mu}{\partial \xi^\lambda \xi^\kappa} A^\kappa \quad (10.3.15)$$

The first term on the r.h.s. corresponds to the tensor transformation, but the second term spoils the transformation properties of $\partial_\nu A^\mu$. Let us then examine the transformation properties of $\nabla_\nu A^\mu$:

$$\begin{aligned} \nabla'_\nu A'^\mu &= \partial'_\nu A'^\mu + \Gamma'_{\nu\kappa}{}^\mu A'^\kappa \\ &= \frac{\partial \xi^\lambda}{\partial \xi'^\nu} \frac{\partial \xi'^\mu}{\partial \xi^\gamma} \frac{\partial A^\gamma}{\partial \xi^\lambda} + \frac{\partial \xi^\lambda}{\partial \xi'^\nu} \frac{\partial^2 \xi'^\mu}{\partial \xi^\lambda \xi^\gamma} A^\gamma + \frac{\partial \xi'^\kappa}{\partial \xi^\gamma} \Gamma'_{\nu\kappa}{}^\mu A^\gamma \end{aligned} \quad (10.3.16)$$

If we can show that

$$\frac{\partial \xi'^\kappa}{\partial \xi^\gamma} \Gamma'_{\nu\kappa}{}^\mu = \frac{\partial \xi^\lambda}{\partial \xi'^\nu} \frac{\partial \xi'^\mu}{\partial \xi^\sigma} \Gamma_{\lambda\gamma}^\sigma - \frac{\partial \xi^\lambda}{\partial \xi'^\nu} \frac{\partial^2 \xi'^\mu}{\partial \xi^\lambda \xi^\gamma} \quad (10.3.17)$$

then we will have

$$\nabla'_\nu A'^\mu = \frac{\partial \xi^\lambda}{\partial \xi'^\nu} \frac{\partial \xi'^\mu}{\partial \xi^\sigma} \left[\frac{\partial A^\sigma}{\partial \xi^\lambda} + \Gamma_{\lambda\gamma}^\sigma A^\gamma \right] = \frac{\partial \xi^\lambda}{\partial \xi'^\nu} \frac{\partial \xi'^\mu}{\partial \xi^\sigma} \nabla_\lambda A^\sigma = (\Lambda^{-1})^\lambda{}_\nu \Lambda^\mu{}_\sigma \nabla_\lambda A^\sigma \quad (10.3.18)$$

and we will have accomplished the task of showing that $\nabla_\nu A^\mu$ is a tensor. It is not so difficult to show (10.3.17). First put it in the form

$$\Gamma'_{\nu\kappa}{}^\mu = \frac{\partial \xi^\lambda}{\partial \xi'^\nu} \frac{\partial \xi'^\mu}{\partial \xi^\sigma} \frac{\partial \xi^\gamma}{\partial \xi'^\kappa} \Gamma_{\lambda\gamma}^\sigma - \frac{\partial \xi^\lambda}{\partial \xi'^\nu} \frac{\partial^2 \xi'^\mu}{\partial \xi^\lambda \xi^\gamma} \frac{\partial \xi^\gamma}{\partial \xi'^\kappa} \quad (10.3.19)$$

To show (10.3.19), write

$$\Gamma'_{\nu\kappa}{}^\mu = \partial'_\nu \Gamma'_{\kappa}{}^\mu = (\partial'_\nu \vec{e}'_{\kappa}) \cdot \vec{E}'^\mu = -\vec{e}'_{\kappa} \cdot (\partial'_\nu \vec{E}'^\mu) \quad (10.3.20)$$

where we have used the fact that $\Gamma'_{\nu\kappa}{}^\mu = (\partial'_\nu \vec{e}'_{\kappa}) \cdot \vec{E}'^\mu$, which follows from the definition of $\delta\Gamma$ in (10.3.8), and $\vec{e}'_{\kappa} \cdot \vec{E}'^\mu = \delta_{\kappa}^\mu$. Then

$$\begin{aligned} \Gamma'_{\nu\kappa}{}^\mu &= -\frac{\partial\xi^\gamma}{\partial\xi'^{\nu\kappa}} \frac{\partial\xi^\lambda}{\partial\xi'^{\nu\kappa}} \vec{e}'_{\gamma} \cdot \frac{\partial}{\partial\xi^\lambda} \left(\frac{\partial\xi'^{\mu}}{\partial\xi^\sigma} \vec{E}^\sigma \right) \\ &= -\frac{\partial\xi^\gamma}{\partial\xi'^{\nu\kappa}} \frac{\partial\xi^\lambda}{\partial\xi'^{\nu\kappa}} \frac{\partial\xi'^{\mu}}{\partial\xi^\sigma} \vec{e}'_{\gamma} \cdot (\partial_\lambda \vec{E}^\sigma) - \frac{\partial\xi^\gamma}{\partial\xi'^{\nu\kappa}} \frac{\partial\xi^\lambda}{\partial\xi'^{\nu\kappa}} \frac{\partial^2\xi'^{\mu}}{\partial\xi^\lambda\partial\xi^\sigma} \vec{e}'_{\gamma} \cdot \vec{E}^\sigma \\ &= \frac{\partial\xi^\gamma}{\partial\xi'^{\nu\kappa}} \frac{\partial\xi^\lambda}{\partial\xi'^{\nu\kappa}} \frac{\partial\xi'^{\mu}}{\partial\xi^\sigma} \Gamma_{\lambda\gamma}^\sigma - \frac{\partial\xi^\lambda}{\partial\xi'^{\nu\kappa}} \frac{\partial^2\xi'^{\mu}}{\partial\xi^\lambda\partial\xi^\sigma} \frac{\partial\xi^\gamma}{\partial\xi'^{\nu\kappa}} \end{aligned} \quad (10.3.21)$$

which is the desired result. Again, notice that without the second term the above would correspond to a tensor transformation, but the second term spoils the transformation properties. In fact it is precisely because of the presence of the second term that $\nabla_\nu A^\mu$ transforms as a tensor. Note also that if the unprimed coordinates were Cartesian, $(\sigma, \lambda, \gamma) \equiv (a, b, c)$, then $\Gamma_{bc}^a \equiv 0$ and

$$\Gamma_{\nu\kappa}^\mu = -\frac{\partial x^b}{\partial\xi^\nu} \frac{\partial^2\xi^\mu}{\partial x^b\partial x^c} \frac{\partial x^c}{\partial\xi^\kappa} \quad (10.3.22)$$

which shows that $\Gamma_{\nu\kappa}^\mu$ is symmetric in (ν, κ) .

In the Cartesian basis, the derivative of a vector is just $\partial_a A^b$. If we now transform to the curvilinear coordinates,

$$\partial_a A^b = \frac{\partial\xi^\mu}{\partial x^a} \frac{\partial}{\partial\xi^\mu} (A^\nu e_\nu^b) = E_a^\mu (\partial_\mu A^\nu) e_\nu^b + E_a^\mu A^\nu (\partial_\mu e_\nu^b) \quad (10.3.23)$$

so that

$$\begin{aligned} e_\sigma^a E_b^\lambda \partial_a A^b &= e_\sigma^a E_b^\lambda E_a^\mu (\partial_\mu A^\nu) e_\nu^b + e_\sigma^a E_b^\lambda E_a^\mu A^\nu (\partial_\mu e_\nu^b) \\ &= \partial_\sigma A^\lambda + E_b^\lambda A^\nu (\partial_\sigma e_\nu^b) = \partial_\sigma A^\lambda + \Gamma_{\sigma\nu}^\lambda A^\nu \end{aligned} \quad (10.3.24)$$

If we think of $\partial_a A^b$ as the components of a (mixed) tensor in the Cartesian system then, in a general coordinate system, its components should be given by $e_\sigma^a E_b^\lambda \partial_a A^b$. The above equation shows that its components in the general coordinate basis are given by the components of the covariant derivative. In other words, derivatives of vectors in the Cartesian coordinate system must be replaced by covariant derivatives in general coordinate systems.

The connection measures the rate of change of the basis as we move from point to point and is computed from the metric $g_{\mu\nu}$ as we will now see. Consider the change in metric as we move from x to $x + dx$,

$$\begin{aligned}\delta g_{\mu\nu} &= \delta \vec{e}_\mu \cdot \vec{e}_\nu + \vec{e}_\mu \cdot \delta \vec{e}_\nu = (\delta \Gamma_\mu^\kappa) \vec{e}_\kappa \cdot \vec{e}_\nu + \vec{e}_\mu \cdot \vec{e}_\kappa (\delta \Gamma_\nu^\kappa) \\ \rightarrow \quad \partial_\gamma g_{\mu\nu} &= \Gamma_{\gamma\mu}^\kappa g_{\kappa\nu} + \Gamma_{\gamma\nu}^\kappa g_{\mu\kappa}\end{aligned}\quad (10.3.25)$$

If we take the combination

$$\begin{aligned}\partial_\gamma g_{\mu\nu} + \partial_\nu g_{\gamma\mu} - \partial_\mu g_{\nu\gamma} &= \Gamma_{\gamma\mu}^\kappa g_{\kappa\nu} + \Gamma_{\gamma\nu}^\kappa g_{\mu\kappa} + \Gamma_{\nu\gamma}^\kappa g_{\kappa\mu} + \Gamma_{\nu\mu}^\kappa g_{\gamma\kappa} - \Gamma_{\mu\nu}^\kappa g_{\kappa\gamma} - \Gamma_{\mu\gamma}^\kappa g_{\nu\kappa} \\ &= \Gamma_{\gamma\nu}^\kappa g_{\mu\kappa} + \Gamma_{\nu\gamma}^\kappa g_{\mu\kappa}\end{aligned}\quad (10.3.26)$$

and use the fact we just saw that $\Gamma_{\gamma\nu}^\kappa = \Gamma_{\nu\gamma}^\kappa$, *i.e.*, $\partial_{[\gamma} \Gamma_{\nu]}^\kappa \equiv 0$, then

$$\partial_\gamma g_{\mu\nu} + \partial_\nu g_{\gamma\mu} - \partial_\mu g_{\nu\gamma} = 2\Gamma_{\gamma\nu}^\kappa g_{\mu\kappa}\quad (10.3.27)$$

and

$$\Gamma_{\gamma\nu}^\kappa g_{\mu\kappa} g^{\mu\rho} = \Gamma_{\gamma\nu}^\kappa \delta_\kappa^\rho = \Gamma_{\gamma\nu}^\rho = \frac{1}{2} g^{\rho\mu} [\partial_\nu g_{\mu\gamma} + \partial_\gamma g_{\mu\nu} - \partial_\mu g_{\gamma\nu}]\quad (10.3.28)$$

where we have freely used the fact that $g_{\mu\nu}$ is symmetric.

It should be clear that the covariant derivative of a tensor copies the covariant derivative of the vector. Setting,

$$\mathbb{T} = T^{\mu\nu\dots} \vec{e}_\mu \vec{e}_\nu \dots\quad (10.3.29)$$

we get

$$\delta \mathbb{T} = \delta T^{\mu\nu\dots} \vec{e}_\mu \vec{e}_\nu \dots + T^{\mu\nu\dots} (\delta \vec{e}_\mu) \vec{e}_\nu \dots + T^{\mu\nu\dots} \vec{e}_\mu (\delta \vec{e}_\nu) \dots + \dots\quad (10.3.30)$$

from which it follows that

$$\nabla_\gamma T^{\mu\nu\dots} = \partial_\gamma T^{\mu\nu\dots} + \Gamma_{\gamma\lambda}^\mu T^{\lambda\nu\dots} + \Gamma_{\gamma\lambda}^\nu T^{\mu\lambda\dots} + \dots\quad (10.3.31)$$

We have defined the covariant derivatives of a contravariant vector. How about the covariant derivative of a covector? We should find

$$\delta \mathbb{A} = (\delta A_\mu) \vec{E}^\mu + A_\mu (\delta \vec{E}^\mu)\quad (10.3.32)$$

and we want to know what $\delta \vec{E}^\mu$ is. Use the fact that

$$e_\mu^a E_b^\mu = \delta_b^a \rightarrow (\delta e_\mu^a) E_b^\mu + e_\mu^a (\delta E_b^\mu) = 0\quad (10.3.33)$$

Therefore

$$e_\mu^a (\delta E_b^\mu) = -(\delta e_\mu^a) E_b^\mu = -(\delta \Gamma_\mu^\kappa) e_\kappa^a E_b^\mu\quad (10.3.34)$$

and, multiplying the l.h.s. by E_a^ν gives

$$E_a^\nu e_\mu^a (\delta E_b^\mu) = \delta_\mu^\nu (\delta E_b^\mu) = \delta E_b^\nu = -(\delta \Gamma_\mu^\kappa) E_a^\nu e_\kappa^a E_b^\mu = -(\delta \Gamma_\mu^\kappa) \delta_k^\nu E_b^\mu = -(\delta \Gamma_\mu^\nu) E_b^\mu \quad (10.3.35)$$

Thus we have found simply that

$$\delta \vec{E}^\mu = -(\delta \Gamma_\mu^\nu) \vec{E}^\nu \quad (10.3.36)$$

which should be compared with (10.3.8) for the variation of \vec{e}_μ . Therefore

$$\delta \mathbb{A} = (DA_\mu) \vec{E}^\mu = (\delta A_\mu) \vec{E}^\mu - (\delta \Gamma_\mu^\kappa) A_\mu \vec{E}^\kappa \quad (10.3.37)$$

and we could write the covariant derivative of the covector, A_μ

$$\nabla_\nu A_\mu = \partial_\nu A_\mu - \Gamma_{\nu\mu}^\kappa A_\kappa \quad (10.3.38)$$

and of a co-tensor, $T_{\mu\nu\kappa\dots}$

$$\nabla_\gamma T_{\mu\nu\dots} = \partial_\gamma T_{\mu\nu\dots} - \Gamma_{\gamma\mu}^\lambda T_{\lambda\nu\dots} - \Gamma_{\gamma\nu}^\lambda T_{\mu\lambda\dots} + \dots \quad (10.3.39)$$

in complete analogy with the covariant derivative of contravectors and tensors. In particular we see that

$$\nabla_\gamma g_{\mu\nu} = \partial_\gamma g_{\mu\nu} - \Gamma_{\gamma\mu}^\kappa g_{\kappa\nu} - \Gamma_{\gamma\nu}^\kappa g_{\mu\kappa} \equiv 0 \equiv \nabla_\gamma g^{\mu\nu} \quad (10.3.40)$$

by (10.3.25). This is called the “metricity” property.³

10.3.3 Absolute Derivative: parallel transport

Having defined the covariant derivative operator, in a general coordinate system, we may now define the *absolute derivative* of a tensor, \mathbb{T} , along some curve with tangent vector U as the projection of the covariant derivative on the tangent, *i.e.*, if the curve is specified by $\xi^\mu(\lambda)$,

$$\frac{D\mathbb{T}}{D\lambda} = U \cdot \nabla \mathbb{T} \quad (10.3.41)$$

The absolute derivative measures the total rate of change of the vector along the curve $\xi^\mu(\lambda)$ and is a tensor of the same rank as \mathbb{T} itself. It is also called the directional derivative of \mathbb{T} .

³Using the Lie derivative of the metric (a rank two co-tensor) show that if U is a symmetry of the metric then it must satisfy

$$\nabla_{(\mu} U_{\nu)} = \nabla_\mu U_\nu + \nabla_\nu U_\mu = 0$$

The symmetry vectors of the metric are called Killing vectors. In Minkowski space there are 10 of them and they generate the Poincaré group: translations, spatial rotations and boosts.

A tensor \mathbb{T} will be said to be “parallel transported” along a curve $\xi^\mu(\lambda)$ if and only if

$$\frac{D\mathbb{T}}{D\lambda} = f(\lambda)\mathbb{T} \quad (10.3.42)$$

where $f(\lambda)$ is an arbitrary function of the curve’s parameter. Let A^μ be parallelly transported along the curve, then

$$U \cdot \nabla A^\mu = U^\sigma (\partial_\sigma A^\mu + \Gamma_{\sigma\kappa}^\mu A^\kappa) = \frac{dA^\mu}{d\lambda} + \Gamma_{\sigma\kappa}^\mu U^\sigma A^\kappa = f(\lambda)A^\mu \quad (10.3.43)$$

This is the condition for parallel transport. In particular, if A^μ is the tangent (velocity) vector of the curve itself, we see that

$$\frac{dU^\mu}{d\lambda^2} + \Gamma_{\sigma\kappa}^\mu U^\sigma U^\kappa = \frac{d^2\xi^\mu}{d\lambda^2} + \Gamma_{\sigma\kappa}^\mu \frac{d\xi^\sigma}{d\lambda} \frac{d\xi^\kappa}{d\lambda} = f(\lambda) \frac{d\xi^\mu}{d\lambda} \quad (10.3.44)$$

This is a second order equation for $\xi^\mu(\lambda)$. We notice that if the coordinates were Cartesian, the connections would vanish and with $f(\lambda) = 0$ we would find simply

$$\frac{d^2x^a}{d\lambda^2} = 0 \quad (10.3.45)$$

which we recognize to be the equation of a straight line, the shortest distance between two points (the geodesic). The equation (10.3.44) generalizes this equation for geodesics to arbitrary curved coordinate systems and is called the geodesic equation.

10.3.4 The Laplacian

A very important operator in physics is the Laplacian. It is an invariant under coordinate transformations being defined, in an arbitrary system of coordinates, as $\square_x = \nabla_\mu \nabla^\mu$. Because it involves the covariant derivative its action will depend on whether it operates on a scalar, a vector or a tensor. Consider its operation on a scalar function, ϕ (remember that $\nabla^\mu \phi = \partial^\mu \phi$ is a vector)

$$\square_x \phi = \nabla_\mu \nabla^\mu \phi = \partial_\mu \nabla^\mu \phi + \Gamma_{\mu\kappa}^\mu \nabla^\kappa \phi = \partial_\mu g^{\mu\nu} \partial_\nu \phi + \Gamma_{\mu\kappa}^\mu g^{\kappa\nu} \partial_\nu \phi \quad (10.3.46)$$

where we have used the fact that the covariant derivative operating on a scalar function is just the partial derivative. But

$$\Gamma_{\mu\kappa}^\mu = \frac{1}{2} g^{\mu\rho} [\partial_\kappa g_{\rho\mu} + \partial_\mu g_{\rho\kappa} - \partial_\rho g_{\mu\kappa}] \quad (10.3.47)$$

Interchanging $(\mu\rho)$ in the middle term shows that it cancels the last, so

$$\Gamma_{\mu\kappa}^\mu = \frac{1}{2} g^{\mu\rho} \partial_\kappa g_{\rho\mu} \quad (10.3.48)$$

This expression may be further simplified: let g be the determinant of $g_{\mu\nu}$, then

$$\ln g = \text{tr} \ln \hat{g} \rightarrow \delta \ln g = \frac{\delta g}{g} = \text{tr} \hat{g}^{-1} \delta \hat{g} = g^{\mu\rho} \delta g_{\mu\rho} \quad (10.3.49)$$

and therefore

$$\frac{1}{g} \partial_\kappa g = g^{\mu\rho} \partial_\kappa g_{\mu\rho}, \quad \Gamma_{\mu\kappa}^\mu = \partial_\kappa \ln \sqrt{g} \quad (10.3.50)$$

which means that

$$\square_x \phi = \partial_\mu (g^{\mu\nu} \partial_\nu \phi) + g^{\mu\nu} (\partial_\mu \ln \sqrt{g}) \partial_\nu \phi = \frac{1}{\sqrt{g}} \partial_\mu \sqrt{g} g^{\mu\nu} \partial_\nu \phi \quad (10.3.51)$$

This is a very compact formula. Life is not so easy if the Laplacian, \square_x , operates on a vector (worse, on a tensor), instead of a scalar. Then we have

$$\begin{aligned} \square_x A^\mu &= \nabla_\nu \nabla^\nu A^\mu = g^{\nu\kappa} \nabla_\nu \nabla_\kappa A^\mu = g^{\nu\kappa} [\partial_\nu \nabla_\kappa A^\mu - \Gamma_{\nu\kappa}^\lambda \nabla_\lambda A^\mu + \Gamma_{\nu\lambda}^\mu \nabla_\kappa A^\lambda] \\ &= g^{\nu\kappa} [\partial_\nu (\partial_\kappa A^\mu + \Gamma_{\kappa\lambda}^\mu A^\lambda) - \Gamma_{\nu\kappa}^\lambda (\partial_\lambda A^\mu + \Gamma_{\lambda\gamma}^\mu A^\gamma) \\ &\quad + \Gamma_{\nu\lambda}^\mu (\partial_\kappa A^\lambda + \Gamma_{\kappa\gamma}^\lambda A^\gamma)] \end{aligned} \quad (10.3.52)$$

which is certainly more complicated. Let's see how this works through some common examples. Only the results will be given, the details are left to the reader.

10.4 Examples

Spherical Coordinates

Take the following coordinate functions: $\xi^\mu = (t, r, \theta, \phi)$ where

$$\begin{aligned} t &= t \\ r &= \sqrt{x^2 + y^2 + z^2} \\ \theta &= \cos^{-1} \left(\frac{z}{\sqrt{x^2 + y^2 + z^2}} \right) \\ \varphi &= \tan^{-1} \left(\frac{y}{x} \right) \end{aligned} \quad (10.4.1)$$

and the inverse transformations: $x^a = x^a(\xi)$

$$\begin{aligned} t &= t \\ x &= r \sin \theta \cos \varphi \\ y &= r \sin \theta \sin \varphi \end{aligned}$$

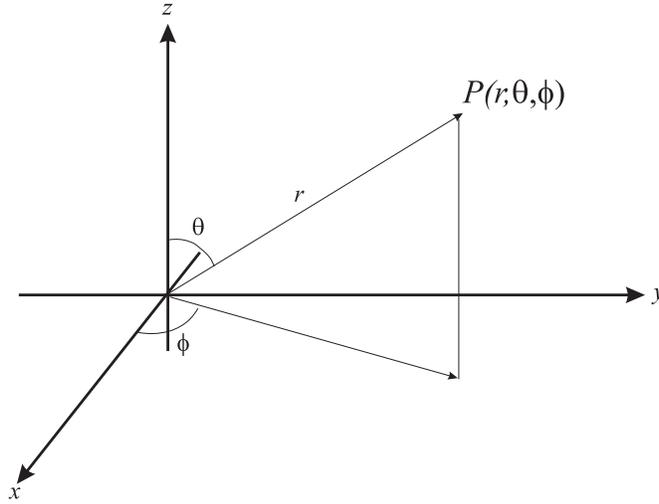


Figure 10.1: Spherical coordinates

$$z = r \cos \theta \quad (10.4.2)$$

Let's compute the vielbein

$$\begin{aligned} \vec{e}_t &= \left(\frac{\partial t}{\partial t}, \frac{\partial x}{\partial t}, \frac{\partial y}{\partial t}, \frac{\partial z}{\partial t} \right) = (1, 0, 0, 0) \\ \vec{e}_r &= \left(\frac{\partial t}{\partial r}, \frac{\partial x}{\partial r}, \frac{\partial y}{\partial r}, \frac{\partial z}{\partial r} \right) = (0, \sin \theta \cos \varphi, \sin \theta \sin \varphi, \cos \theta) \\ \vec{e}_\theta &= \left(\frac{\partial t}{\partial \theta}, \frac{\partial x}{\partial \theta}, \frac{\partial y}{\partial \theta}, \frac{\partial z}{\partial \theta} \right) = r(0, \cos \theta \cos \varphi, \cos \theta \sin \varphi, -\sin \theta) \\ \vec{e}_\varphi &= \left(\frac{\partial t}{\partial \varphi}, \frac{\partial x}{\partial \varphi}, \frac{\partial y}{\partial \varphi}, \frac{\partial z}{\partial \varphi} \right) = r(0, -\sin \theta \sin \varphi, \sin \theta \cos \varphi, 0) \end{aligned} \quad (10.4.3)$$

and its inverse

$$\begin{aligned} \vec{E}^t &= \left(\frac{\partial t}{\partial t}, \frac{\partial t}{\partial x}, \frac{\partial t}{\partial y}, \frac{\partial t}{\partial z} \right) = (1, 0, 0, 0) \\ \vec{E}^r &= \left(\frac{\partial r}{\partial t}, \frac{\partial r}{\partial x}, \frac{\partial r}{\partial y}, \frac{\partial r}{\partial z} \right) = (0, \sin \theta \cos \varphi, \sin \theta \sin \varphi, \cos \theta) \\ \vec{E}^\theta &= \left(\frac{\partial \theta}{\partial t}, \frac{\partial \theta}{\partial x}, \frac{\partial \theta}{\partial y}, \frac{\partial \theta}{\partial z} \right) = \frac{1}{r}(0, \cos \theta \cos \varphi, \cos \theta \sin \varphi, -\sin \theta) \end{aligned}$$

$$\vec{E}^\varphi = \left(\frac{\partial \varphi}{\partial t}, \frac{\partial \varphi}{\partial x}, \frac{\partial \varphi}{\partial y}, \frac{\partial \varphi}{\partial z} \right) = \frac{1}{r} (0, -\sin \theta \sin \varphi, \sin \theta \cos \varphi, 0) \quad (10.4.4)$$

It is easy to check that $\vec{e}_m \cdot \vec{E}^n = \delta_m^n$ and that $e_m^a E_b^m = \delta_b^a$. Now compute the inner products to get the metric function: $g_{tt} = -1$, $g_{rr} = 1$, $g_{\theta\theta} = r^2$ and $g_{\varphi\varphi} = r^2 \sin^2 \theta$ (all other components vanish). In matrix notation,

$$g_{\mu\nu} = \begin{bmatrix} -c^2 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & r^2 & 0 \\ 0 & 0 & 0 & r^2 \sin^2 \theta \end{bmatrix} \quad (10.4.5)$$

and the distance function is given explicitly by,

$$ds^2 = c^2 dt^2 - (dr^2 + r^2 d\theta^2 + r^2 \sin^2 \theta d\varphi^2) \quad (10.4.6)$$

Next compute the connections using either $\Gamma_{\nu\kappa}^\mu = (\partial_\nu \vec{e}_\kappa) \cdot \vec{E}^\mu$ or (10.3.28) to get the non-vanishing components

$$\begin{aligned} \Gamma_{\theta\theta}^r &= -r, & \Gamma_{\varphi\varphi}^r &= -r \sin^2 \theta \\ \Gamma_{r\theta}^\theta &= \Gamma_{\theta r}^\theta = \frac{1}{r}, & \Gamma_{r\varphi}^\varphi &= \Gamma_{\varphi r}^\varphi = \frac{1}{r} \\ \Gamma_{\varphi\varphi}^\theta &= -\sin \theta \cos \theta, & \Gamma_{\varphi\theta}^\varphi &= \Gamma_{\theta\varphi}^\varphi = \cot \theta \end{aligned} \quad (10.4.7)$$

(all others vanish). What is the action of the Laplacian, \square_x , on a scalar function?

$$\square_x \phi = \frac{1}{\sqrt{g}} \partial_\mu (\sqrt{g} g^{\mu\nu} \partial_\nu \phi) = -\frac{1}{c^2} \partial_t^2 \phi + \frac{1}{r^2} \partial_r (r^2 \partial_r \phi) + \frac{1}{r^2 \sin \theta} \partial_\theta (\sin \theta \partial_\theta \phi) + \frac{1}{r^2 \sin^2 \theta} \partial_\varphi^2 \phi \quad (10.4.8)$$

the spatial part of which will be recognized as the standard result from ordinary vector analysis. Its action on vectors is quite a bit more complicated but can be written out,

$$\begin{aligned} \square_x A^0 &= \left[-\frac{1}{c^2} \partial_t^2 A^0 + \frac{1}{r^2} \partial_r (r^2 \partial_r A^0) + \frac{1}{r^2 \sin \theta} \partial_\theta (\sin \theta \partial_\theta A^0) + \frac{1}{r^2 \sin^2 \theta} \partial_\varphi^2 A^0 \right] \\ \square_x A^r &= \left[-\frac{1}{c^2} \partial_t^2 A^r + \frac{1}{r^2} \partial_r (r^2 \partial_r A^r) + \frac{1}{r^2 \sin \theta} \partial_\theta (\sin \theta \partial_\theta A^r) + \frac{1}{r^2 \sin^2 \theta} \partial_\varphi^2 A^r \right. \\ &\quad \left. - \frac{2}{r^2} (A^r + r \cot \theta A^\theta + r \partial_\theta A^\theta + r \partial_\varphi A^\varphi) \right] \\ \square_x A^\theta &= \left[-\frac{1}{c^2} \partial_t^2 A^\theta + \frac{1}{r^4} \partial_r (r^4 \partial_r A^\theta) + \frac{1}{r^2 \sin \theta} \partial_\theta (\sin \theta \partial_\theta A^\theta) + \frac{1}{r^2 \sin^2 \theta} \partial_\varphi^2 A^\theta \right. \\ &\quad \left. + \frac{2}{r^3} (\partial_\theta A^r - \frac{1}{2} r \cos 2\theta A^\theta - r \cot \theta \partial_\varphi A^\varphi) \right] \end{aligned}$$

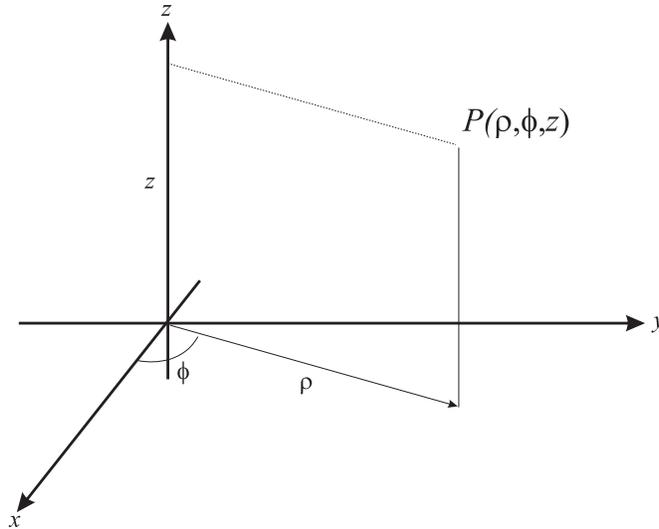


Figure 10.2: Cylindrical coordinates

$$\square_x A^\varphi = \left[-\frac{1}{c^2} \partial_t^2 A^\varphi + \frac{1}{r^4} \partial_r (r^4 \partial_r A^\varphi) + \frac{1}{r^2 \sin^3 \theta} \partial_\theta (\sin^3 \theta \partial_\theta A^\varphi) + \frac{1}{r^2 \sin^2 \theta} \partial_\varphi^2 A^\varphi + \frac{2}{r^3 \sin^3 \theta} (\sin \theta \partial_\varphi A^r + r \cos \theta \partial_\varphi A^\theta) \right] \quad (10.4.9)$$

We see that the Laplacian acts on the time component, A^0 , of A^μ , just exactly as it does on a scalar. This is because the coordinate transformation was purely spatial. On the other hand, its action on the space components mixes them.

Cylindrical coordinates

Take the following coordinate functions: $\xi^\mu = (t, \rho, \varphi, z)$ where

$$\begin{aligned} t &= t \\ \rho &= \sqrt{x^2 + y^2} \\ \varphi &= \tan^{-1} \left(\frac{y}{x} \right) \\ z &= z \end{aligned} \quad (10.4.10)$$

and the inverse transformations: $x^a = x^a(\xi)$

$$t = t$$

$$\begin{aligned}
x &= \rho \cos \varphi \\
y &= \rho \sin \varphi \\
z &= z
\end{aligned} \tag{10.4.11}$$

Let's compute the vielbein

$$\begin{aligned}
\vec{e}_t &= \left(\frac{\partial t}{\partial t}, \frac{\partial x}{\partial t}, \frac{\partial y}{\partial t}, \frac{\partial z}{\partial t} \right) = (1, 0, 0, 0) \\
\vec{e}_\rho &= \left(\frac{\partial t}{\partial \rho}, \frac{\partial x}{\partial \rho}, \frac{\partial y}{\partial \rho}, \frac{\partial z}{\partial \rho} \right) = (0, \cos \varphi, \sin \varphi, 0) \\
\vec{e}_\varphi &= \left(\frac{\partial t}{\partial \varphi}, \frac{\partial x}{\partial \varphi}, \frac{\partial y}{\partial \varphi}, \frac{\partial z}{\partial \varphi} \right) = \rho(0, -\sin \varphi, \cos \varphi, 0) \\
\vec{e}_z &= \left(\frac{\partial t}{\partial z}, \frac{\partial x}{\partial z}, \frac{\partial y}{\partial z}, \frac{\partial z}{\partial z} \right) = (0, 0, 0, 1)
\end{aligned} \tag{10.4.12}$$

and its inverse

$$\begin{aligned}
\vec{E}^t &= \left(\frac{\partial t}{\partial t}, \frac{\partial t}{\partial x}, \frac{\partial t}{\partial y}, \frac{\partial t}{\partial z} \right) = (1, 0, 0, 0) \\
\vec{E}^\rho &= \left(\frac{\partial \rho}{\partial t}, \frac{\partial \rho}{\partial x}, \frac{\partial \rho}{\partial y}, \frac{\partial \rho}{\partial z} \right) = (0, \cos \varphi, \sin \varphi, 0) \\
\vec{E}^\varphi &= \left(\frac{\partial \varphi}{\partial t}, \frac{\partial \varphi}{\partial x}, \frac{\partial \varphi}{\partial y}, \frac{\partial \varphi}{\partial z} \right) = \frac{1}{\rho}(0, -\sin \varphi, \cos \varphi, 0) \\
\vec{E}^z &= \left(\frac{\partial z}{\partial t}, \frac{\partial z}{\partial x}, \frac{\partial z}{\partial y}, \frac{\partial z}{\partial z} \right) = (0, 0, 0, 1)
\end{aligned} \tag{10.4.13}$$

Again, it's easy to check that $\vec{e}_m \cdot \vec{E}^n = \delta_m^n$ and that $e_m^a E_b^m = \delta_b^a$. Now compute the inner products to get the metric function: $g_{tt} = -1$, $g_{\rho\rho} = 1$, $g_{\varphi\varphi} = \rho^2$ and $g_{zz} = 1$ (all other components vanish). In matrix notation,

$$g_{\mu\nu} = \begin{bmatrix} -c^2 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \rho^2 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{10.4.14}$$

and the distance function is given explicitly by,

$$ds^2 = c^2 dt^2 - (d\rho^2 + \rho^2 d\varphi^2 + dz^2) \tag{10.4.15}$$

The non-vanishing components of the connections, obtained by using either $\Gamma_{\nu\kappa}^\mu = (\partial_\nu \vec{e}_\kappa) \cdot \vec{E}^\mu$ or (10.3.28) are just

$$\Gamma_{\varphi\varphi}^\rho = -\rho,$$

$$\Gamma_{\rho\varphi}^{\varphi} = \Gamma_{\varphi\rho}^{\varphi} = \frac{1}{\rho} \quad (10.4.16)$$

(all others vanish), while the action of the Laplacian, \square_x , on a scalar function is

$$\square_x \phi = \frac{1}{\sqrt{g}} \partial_{\mu} (\sqrt{g} g^{\mu\nu} \partial_{\nu} \phi) = -\frac{1}{c^2} \partial_t^2 \phi + \frac{1}{\rho} \partial_{\rho} (\rho \partial_{\rho} \phi) + \frac{1}{\rho^2} \partial_{\varphi}^2 \phi + \partial_z^2 \phi \quad (10.4.17)$$

the spatial part of which being, as before, the standard result from ordinary vector analysis. Its action on vectors can be written out and we leave this as a straightforward exercise.^{4,5}

10.5 Integration: The Volume Element

When passing from Cartesian coordinates to general coordinates one must also take care to account for the change in the integration measure, which follows the usual rule,

$$\int d^4x \rightarrow \int d^4\xi \left\| \frac{\partial x}{\partial \xi} \right\| \quad (10.5.1)$$

where $\|\partial\xi/\partial x\|$ represents the Jacobian of the transformation. Now notice that under the coordinate transformation that took $x^a \rightarrow \xi^{\mu}$, the metric also underwent a transformation

$$\eta_{ab} \rightarrow g_{\mu\nu} = \eta_{ab} \frac{\partial x^a}{\partial \xi^{\mu}} \frac{\partial x^a}{\partial \xi^{\nu}} = \eta_{ab} e_{\mu}^a e_{\nu}^b \quad (10.5.2)$$

It follows, upon taking determinants, that

$$\|\hat{g}\| = \|\hat{\eta}\| \left\| \frac{\partial x}{\partial \xi} \right\|^2 = -c^2 \left\| \frac{\partial x}{\partial \xi} \right\|^2 \quad (10.5.3)$$

where we have used $\|\hat{\eta}\| = -c^2$. Therefore,

$$\left\| \frac{\partial x}{\partial \xi} \right\| = \frac{1}{c} \sqrt{-\|\hat{g}\|} \quad (10.5.4)$$

We have previously used the notation g for the determinant of the metric, \hat{g} . Continuing with this notation we notice that (10.5.1) can be written as

$$\int d^4x \rightarrow \frac{1}{c} \int d^4\xi \sqrt{-g} \quad (10.5.5)$$

⁴Problem: Write out $\square_x A^{\mu}$ for each component of A^{μ} in cylindrical coordinates.

⁵Problem: Work out the details of the Rindler spacetime.

Again, we see that if the transformation is just a Lorentz transformation,

$$\int d^4x \rightarrow \int d^4x' \quad (10.5.6)$$

because the metric $g_{\mu\nu}$ then continues to be just the Lorentz metric. Thus we are led to define

$$\frac{1}{c} \int d^4\xi \sqrt{-g} \quad (10.5.7)$$

as the correct volume integration measure in *any* system of coordinates. Simple examples are spherical coordinates:

$$\frac{1}{c} \int d^4\xi \sqrt{-g} = \int dt \int dr r^2 \int_0^\pi d\theta \sin\theta \int_0^{2\pi} d\varphi \quad (10.5.8)$$

and cylindrical coordinates:

$$\frac{1}{c} \int d^4\xi \sqrt{-g} = \int dt \int dz \int d\rho \rho^2 \int_0^{2\pi} d\varphi \quad (10.5.9)$$

but (10.5.7) has general applicability.

Chapter 11

Solutions of the Wave Equation

11.1 Green's functions

Recall that the equations for the electromagnetic field in the Coulomb gauge read

$$\vec{\nabla}^2 \phi = -\frac{\rho(\vec{r}, t)}{\epsilon_0}, \quad \square_x \vec{A} = \mu_0 \vec{j}(\vec{r}, t) \quad (11.1.1)$$

and in the Lorentz gauge they read

$$\square_x A_\mu = j_\mu(x). \quad (11.1.2)$$

The equations are linear and they can be solved exactly. In this chapter, we ask for the general solutions and consider what boundary conditions must be applied when two or more media are present in a given physical problem. We'll begin examining solutions of the equation for ϕ in the Coulomb gauge. This is because it is quite a bit simpler to do than examining solutions to the others, but it contains many of the computational details anyway. As a bonus we'll prove the identity

$$\vec{\nabla}^2 \frac{1}{|\vec{r} - \vec{r}'|} = -4\pi \delta(\vec{r} - \vec{r}'). \quad (11.1.3)$$

11.1.1 Spatial Green's function

Assume that the charge distribution we are interested in vanishes in all but a finite region of space and that the only boundary condition on $\phi(\vec{r}, t)$ is the "trivial" condition at infinity, *i.e.*, that $\phi(\vec{r}, t)$ falls off to zero at least as fast as $1/r$ there. Now ask for a solution of the form

$$\phi(\vec{r}, t) = \int d^3 \vec{r}' G(\vec{r}, \vec{r}') \rho(\vec{r}', t). \quad (11.1.4)$$

The three dimensional Laplacian is translation invariant outside the charge distribution and, under these boundary conditions, we expect $G(\vec{r}, \vec{r}')$ to be translation invariant also. This means that $G(\vec{r}, \vec{r}') = G(\vec{r} - \vec{r}')$. It is easy to see that

$$\vec{\nabla}^2 \phi = \int d^3 \vec{r}' \vec{\nabla}^2 G(\vec{r}, \vec{r}') \rho(\vec{r}', t) = -\frac{\rho(\vec{r}, t)}{\epsilon_o}, \quad (11.1.5)$$

implies that

$$\vec{\nabla}^2 G(\vec{r}, \vec{r}') = -\frac{1}{\epsilon_o} \delta^3(\vec{r} - \vec{r}'). \quad (11.1.6)$$

Any function that satisfies an equation of this type (where the r.h.s. is a δ -function) is called a Green's function. Let us write

$$G(\vec{r}, \vec{r}') = \int \frac{d^3 \vec{k}}{(2\pi)^3} \mathcal{G}(\vec{k}) e^{i\vec{k} \cdot (\vec{r} - \vec{r}')}, \quad (11.1.7)$$

then,

$$\vec{\nabla}^2 G(\vec{r}, \vec{r}') = \int \frac{d^3 \vec{k}}{(2\pi)^3} (-k^2 \mathcal{G}(\vec{k})) e^{i\vec{k} \cdot (\vec{r} - \vec{r}')}. \quad (11.1.8)$$

But, keeping in mind the Fourier representation of the δ -function

$$\delta^3(\vec{r} - \vec{r}') = \int \frac{d^3 \vec{k}}{(2\pi)^3} e^{i\vec{k} \cdot (\vec{r} - \vec{r}')}, \quad (11.1.9)$$

we find that $\mathcal{G}(\vec{k})$ must be given as

$$\mathcal{G}(\vec{k}) = \frac{1}{\epsilon_o k^2} \quad (11.1.10)$$

and therefore

$$G(\vec{r}, \vec{r}') = \int \frac{d^3 \vec{k}}{(2\pi)^3 \epsilon_o} \frac{e^{i\vec{k} \cdot (\vec{r} - \vec{r}')}}{k^2} = -\frac{1}{4\pi^2 \epsilon_o} \int_0^\infty dk \int_0^\pi d(\cos \theta) e^{ik\Delta r \cos \theta}, \quad (11.1.11)$$

where we have used spherical coordinates in \vec{k} -space, $k = |\vec{k}|$ and $\Delta r = |\vec{r} - \vec{r}'|$. Integrating over the polar angle, we get

$$G(\vec{r}, \vec{r}') = \frac{1}{4\pi^2 \epsilon_o} \int_0^\infty \frac{dk}{ik\Delta r} \left[e^{ik\Delta r} - e^{-ik\Delta r} \right] = -\frac{i}{4\pi^2 \epsilon_o \Delta r} \int_{-\infty}^\infty \frac{dk}{k} e^{ik\Delta r}, \quad (11.1.12)$$

where we have adjoined the second integral to the first by extending the limits of integration to run from $(-\infty, \infty)$. Our central problem is then to evaluate this integral, but it has a pole at $k = 0$. We therefore *define* it to be the principal part of an integral in

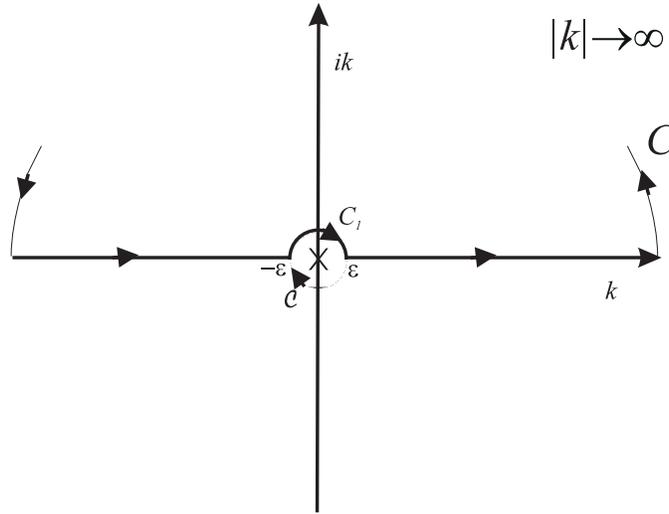


Figure 11.1: Contour of integration

the complex k -plane along the contour shown in figure 3. Note that the contour *must* be closed in the upper half plane so that the integral along the semi-circle at infinity vanishes, contributing nothing to the integration (by Jordan's lemma). If one defines the integral by its principal value,

$$P \int_{-\infty}^{\infty} \frac{dk}{k} e^{ik\Delta r} = \lim_{\epsilon \rightarrow 0} \left[\int_{-\infty}^{-\epsilon} \frac{dk}{k} e^{ik\Delta r} + \int_{\epsilon}^{\infty} \frac{dk}{k} e^{ik\Delta r} \right], \quad (11.1.13)$$

then it is easy to see that

$$P \int_{-\infty}^{\infty} \frac{dk}{k} e^{ik\Delta r} = \oint_C \frac{dk}{k} e^{ik\Delta r} - \int_{C_1} \frac{dk}{k} e^{ik\Delta r} \quad (11.1.14)$$

(the contribution of the semicircle at infinity vanishes by Jordan's lemma). But the first integral on the r.h.s. is evidently zero because the contour contains no poles. It follows that

$$\begin{aligned} P \int_{-\infty}^{\infty} \frac{dk}{k} e^{ik\Delta r} &= - \int_{C_1} \frac{dk}{k} e^{ik\Delta r} = -\frac{1}{2} \times \oint_C \frac{dk}{k} e^{ik\Delta r} \\ &= -\frac{1}{2} \times 2\pi i \sum \text{Res}_C = \pi i, \end{aligned} \quad (11.1.15)$$

where we have used the fact that C is a *clockwise* loop, which implies an overall minus sign, and that the residue at $k = 0$ is just 1. We substitute this result in the expression

(11.1.11) to obtain

$$G(\vec{r}, \vec{r}') = \frac{1}{4\pi\epsilon_0|\vec{r} - \vec{r}'|}. \quad (11.1.16)$$

Therefore we have found the general solution for ϕ in the Coulomb gauge, it is

$$\phi(\vec{r}) = \int \frac{d^3r'}{4\pi\epsilon_0|\vec{r} - \vec{r}'|} \rho(\vec{r}', t) \quad (11.1.17)$$

and, because of (11.1.6), we have also shown (11.1.3).¹

11.1.2 Invariant Green's function

We expect to be able to use the same techniques to solve (11.1.2) as we did to solve for $\phi(\vec{r}, t)$ above. Let us therefore look for a solution of the form

$$A_\mu(x) = \int d^4x' G(x, x') j_\mu(x'). \quad (11.1.18)$$

Again, the four dimensional Laplacian is translation invariant so we expect that $G(x, x')$ is also translation invariant, *i.e.*, $G(x, x') = G(x - x')$. Moreover, $G(x, x')$ is manifestly a scalar and

$$\square_x G(x, x') = \delta^4(x - x'). \quad (11.1.19)$$

Now express $G(x, x')$ as a four dimensional Fourier transform,

$$G(x, x') = \int \frac{d^4k}{(2\pi)^4} G(k) e^{ik(x-x')} \quad (11.1.20)$$

and recall that

$$\delta^4(x - x') = \int \frac{d^4k}{(2\pi)^4} e^{ik(x-x')}. \quad (11.1.21)$$

Now, $\square_x \equiv -\eta^{\mu\nu} \partial_\mu \partial_\nu$, so because

$$\begin{aligned} \square_x G(x, x') &= \int \frac{d^4k}{(2\pi)^4} G(k) (-\eta^{\mu\nu} \partial_\mu \partial_\nu) e^{ik(x-x')} = \int \frac{d^4k}{(2\pi)^4} G(k) (\eta^{\mu\nu} k_\mu k_\nu) e^{ik(x-x')} \\ &= \int \frac{d^4k}{(2\pi)^4} (-k^2 G(k)) e^{ik(x-x')}, \end{aligned} \quad (11.1.22)$$

¹Problem: The choice of contour is not at all unique. Show that one could take the contour C to *include* the pole at $k = 0$ (C_1 counter-clockwise in the lower half plane) with the same final result. Use the integral we have just solved to obtain

$$\int_0^\infty dx \frac{\sin x}{x} = \frac{\pi}{2}$$

where $k^2 = -\eta^{\mu\nu}k_\mu k_\nu = -\eta_{\mu\nu}k^\mu k^\nu$ it follows that

$$G(k) = -\frac{1}{k^2}, \quad (11.1.23)$$

or

$$G(x, x') = -\int \frac{d^4k}{(2\pi)^4} \frac{e^{ik(x-x')}}{k^2}. \quad (11.1.24)$$

Now $k_\mu = (-\omega, \vec{k})$ and $x - x' = (t - t', \vec{r} - \vec{r}')$ (this is defined so that $k_\mu x^\mu$ is dimensionless) are four vectors. So $k^2 = -\eta_{\mu\nu}k^\mu k^\nu = -\vec{k}^2 + \omega^2/c^2$ and $k \cdot (x - x') = \vec{k} \cdot (\vec{r} - \vec{r}') - \omega(t - t')$. Therefore the integral is more complicated than it seems in our condensed notation:

$$G(x, x') = \int \frac{d^3\vec{k}d\omega}{(2\pi)^4} \frac{e^{i[\vec{k} \cdot (\vec{r} - \vec{r}') - \omega(t - t')]}{k^2 - \omega^2/c^2}. \quad (11.1.25)$$

Let us first perform the integral over the solid angle.

$$G(x, x') = \frac{1}{(2\pi)^3} \int_0^\infty k^2 dk \int_{-\infty}^\infty d\omega \frac{e^{-i\omega\Delta t}}{k^2 - \omega^2/c^2} \int_{-1}^1 d\mu e^{ik\Delta r\mu}, \quad (11.1.26)$$

where $\mu = \cos\theta$, $\Delta r = |\vec{r} - \vec{r}'|$ and $\Delta t = t - t'$, and we are now using $k = |\vec{k}|$. Integrating over μ gives

$$\begin{aligned} G(x, x') &= \frac{-i}{(2\pi)^3 \Delta r} \int_0^\infty k dk \left[e^{ik\Delta r} - e^{-ik\Delta r} \right] \int_{-\infty}^\infty d\omega \frac{e^{-i\omega\Delta t}}{k^2 - \omega^2/c^2} \\ &= \frac{-i}{(2\pi)^3 \Delta r} \int_{-\infty}^\infty k dk e^{ik\Delta r} \int_{-\infty}^\infty d\omega \frac{e^{-i\omega\Delta t}}{k^2 - \omega^2/c^2}. \end{aligned} \quad (11.1.27)$$

Now let's evaluate the ω -integral. Notice that the integral has two poles, at $\omega = \pm ck$. Furthermore, Δt can be either positive or negative (unlike $\Delta r = |\vec{r} - \vec{r}'| \geq 0$) so one must close the contour carefully (see figures 4 and 5).

- For $\Delta t > 0$, we should close the contour in the *lower* half plane.
- For $\Delta t < 0$, we should close the contour in the *upper* half plane.

This rule ensures that the integral over the semicircles at infinity will not contribute to the contour integral (Jordan's lemma). Even if we fix the way the contour will be closed at infinity, we are still faced with two possible ways to go around each of the singular points. We will not define the integral as its principal part this time (it will shortly become clear why not) but instead by the contour chosen. Suppose we choose contour R .

- If $\Delta t < 0$ then we must close the contour in the upper half plane. Then the contour C contains no poles and the integral vanishes.

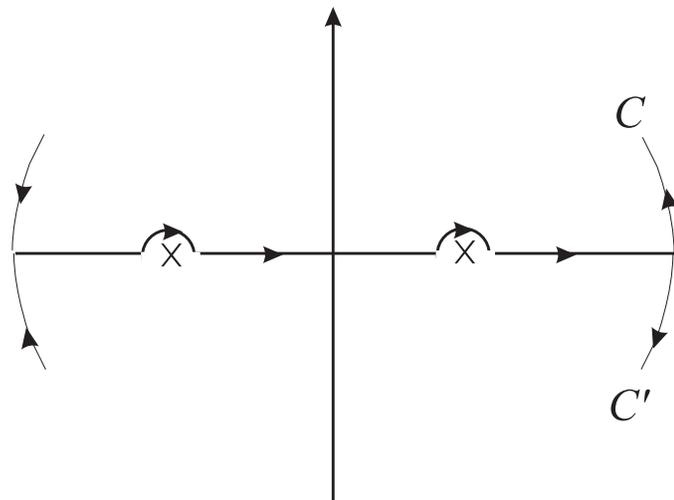


Figure 11.2: Contour R

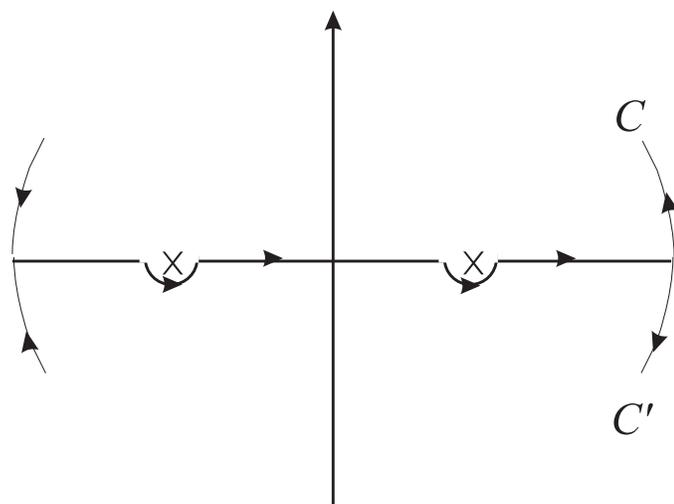


Figure 11.3: Contour A

- If $\Delta t > 0$ then we close the contour in the lower half plane and the contour C' contains two poles, so the integral does not vanish. We calculate its value

$$\oint_{C'} d\omega \frac{e^{-i\omega\Delta t}}{k^2 - \omega^2/c^2} = \frac{-i\pi}{k} [e^{ick\Delta t} - e^{-ick\Delta t}] \quad (11.1.28)$$

(the minus arises because this contour is clockwise). Then

$$\begin{aligned} G(x, x') &= \frac{1}{8\pi^2\Delta r} \int_{-\infty}^{\infty} dk \left[e^{ik(\Delta r - c\Delta t)} - e^{ik(\Delta r + c\Delta t)} \right] \\ &= \frac{1}{4\pi\Delta r} [\delta(\Delta r - c\Delta t) - \delta(\Delta r + c\Delta t)], \end{aligned} \quad (11.1.29)$$

where we have used the fact that

$$\int_{-\infty}^{\infty} \frac{dk}{2\pi} e^{ikZ} = \delta(Z). \quad (11.1.30)$$

Clearly, because $\Delta r > 0$ and $\Delta t > 0$ it follows that the argument of the second δ -function above is never zero and so we may write (keeping in mind that $\Delta t > 0$)

$$G_R(x, x') = \frac{\Theta(t - t')}{4\pi|\vec{r} - \vec{r}'|} \delta(|\vec{r} - \vec{r}'| - c(t - t')), \quad (11.1.31)$$

where $\Theta(t - t')$ is the Heaviside Θ -function, defined by²

$$\Theta(x) = \begin{cases} 1 & \text{if } x > 0 \\ 0 & \text{if } x < 0 \end{cases} \quad (11.1.32)$$

Let us examine this function carefully. It is non-vanishing only when $t > t'$. Now (t', \vec{r}') are coordinates associated with the sources (j_μ) and (t, \vec{r}) are associated with the observation. This Green's function is vanishing for observation times that are *less than* the source time. We understand this as saying that the source time must always be *earlier* than the matter time. This is therefore called the *retarded* or *causal* Green's function. Moreover, the δ -function is not supported unless $\Delta r = c\Delta t$, so information of the source distribution travels precisely at the speed of light and its support is precisely on the forward light cone.

Now let us consider the contour A . The treatment is similar and we'll go through it briefly:

- If $\Delta t > 0$ we must close the contour in the lower half plane, but a look at figure 5 shows that the contour closed in the lower half plane contains no poles, so the Green's function is precisely zero for $\Delta t > 0$.

²Note: $\Theta'(x) = \delta(x)$.

- If, on the contrary, $\Delta t < 0$ the contour is closed in the upper half plane and the Green's function is not vanishing. So, for $\Delta t < 0$ we have

$$\oint_{C'} d\omega \frac{e^{-i\omega\Delta t}}{k^2 - \omega^2/c^2} = \frac{i\pi}{k} [e^{ick\Delta t} - e^{-ick\Delta t}] \quad (11.1.33)$$

(no extra minus sign here!), which gives

$$\begin{aligned} G(x, x') &= \frac{1}{8\pi^2\Delta r} \int_{-\infty}^{\infty} dk \left[e^{ik(\Delta r + c\Delta t)} - e^{ik(\Delta r - c\Delta t)} \right] \\ &= \frac{1}{4\pi\Delta r} [\delta(\Delta r + c\Delta t) - \delta(\Delta r - c\Delta t)]. \end{aligned} \quad (11.1.34)$$

Again, because $\Delta r > 0$ and $\Delta t < 0$ the second δ -function is never supported so

$$G_A(x, x') = \frac{\Theta(t' - t)}{4\pi|\vec{r} - \vec{r}'|} \delta(|\vec{r} - \vec{r}'| + c(t - t')). \quad (11.1.35)$$

Notice that it is non-vanishing only when $t' > t$. This Green's function is vanishing for observation times that are *greater than* the source time. We understand this as saying that the source time must always be *later* than the matter time. This is therefore called the *advanced* or *acausal* Green's function. Moreover, the δ -function is not supported unless $\Delta r = -c\Delta t$, so information of the source distribution travels precisely at the speed of light, but on the backward light cone.³

We are now ready to write down our solutions to the four dimensional wave equation in the presence of sources. Because the equations are linear we may add to the solution in (11.1.18) an arbitrary solution of the homogeneous equation $\square_x A_\mu = 0$. Then we have

$$A_\mu(x) = A_\mu^{\text{in}}(x) + \int d^4x' G_R(x, x') j_\mu(x'), \quad (11.1.36)$$

or

$$A_\mu(x) = A_\mu^{\text{out}}(x) + \int d^4x' G_A(x, x') j_\mu(x'). \quad (11.1.37)$$

The reason for the superscripts “in” and “out” are the following: as $t \rightarrow -\infty$ the second term (integral) in (11.1.36) approaches zero and the only contribution at $t \rightarrow -\infty$ comes from the term $A_\mu^{\text{in}}(x)$, so these functions clearly have the interpretation of being an “incoming” field. On the contrary, as $t \rightarrow +\infty$ the integral in (11.1.37) approaches zero and the only contribution is from the term $A_\mu^{\text{out}}(x)$, which is therefore interpreted as an “outgoing” field.

³Problem: Explain in terms of causality why we chose *not* to define the Green's function by its principal part.

There are several other ways of going around the singularities at $\omega = \pm ck$, but most of them are of interest only in the quantum theory. In this course we will be interested only in $G_{R,A}$. Later on we will use them to compute the electromagnetic fields of moving charges, but for now we will spend a few weeks describing some simple phenomena which can be understood using only the plane wave solution to the homogeneous equation ($j^\mu = 0$) we obtained earlier.

11.2 Mode Expansions

11.3 Boundary conditions

We have seen that the equations for the electromagnetic field (A_μ) are linear second order differential equations. Of course, they must be solved subject to boundary conditions and in the previous sections we obtained solutions subject to the trivial conditions that the field falls off at least as fast as $1/r$ at infinity. In general, fields that obey second order differential equations are required in physics to be $C^{(1)}$ except, possibly, at the location of sources, where they are required to be continuous but not necessarily differentiable. Maxwell's equations, given in terms of \vec{E} and \vec{B} , are the first integrals of the field equations for A_μ . As the electric and magnetic fields are first derivatives of the potentials A_μ , we expect them to be continuous everywhere except, possibly, at the location of sources. What does this mean for \vec{E} and \vec{B} at the boundary between two media? The discontinuities encountered by the electric and magnetic fields at the interface between two media can be deduced from the integral form of Maxwell's equations themselves. Let us begin by determining the implications of the Bianchi identities ((a) no magnetic monopoles and (b) Faraday's law of induction),

$$\vec{\nabla} \cdot \vec{B} = 0 = \vec{\nabla} \times \vec{E} + \frac{\partial \vec{B}}{\partial t}. \quad (11.3.1)$$

- Consider the first equation. We have

$$\vec{\nabla} \cdot \vec{B} = 0 \rightarrow \oint_S \vec{B} \cdot d\vec{S} = 0, \quad (11.3.2)$$

where S is a closed surface. Consider a cylindrical “pill-box”, as shown in figure 9, that contains a patch of the interface between the two media. But, as we are interested only in the boundary between the media “1” and “2” we can make the “pill-box” infinitesimally thin so that the sides of the cylinder do not contribute to the surface integral. Then, if B_{1n} is the normal component of \vec{B} in medium “1” (the projection of \vec{B} on the normal \hat{n} , *i.e.*, $B_{1n} = \vec{B}^{(1)} \cdot \hat{n}_1 = \vec{B}^{(1)} \cdot \hat{n}$) and if B_{2n} is *minus*

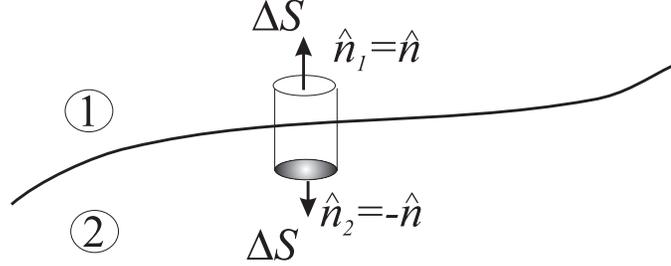


Figure 11.4: Boundary conditions for the normal component

the normal component of \vec{B} in medium “2”, *i.e.*, $B_{2n} = -\vec{B}^{(2)} \cdot \hat{n}_2 = B^{(2)} \cdot n$) we have

$$(B_{1n} - B_{2n})\Delta S = 0 \rightarrow B_{1n} = B_{2n}, \quad (11.3.3)$$

which means that *the normal component of the magnetic field must be continuous across the boundary between the media.*

- Consider second equation, *i.e.*, Faraday’s law (see figure 10),

$$\vec{\nabla} \times \vec{E} + \frac{\partial \vec{B}}{\partial t} = 0 \rightarrow \int_S \vec{\nabla} \times \vec{E} \cdot d\vec{S} + \frac{\partial}{\partial t} \int_S \vec{B} \cdot d\vec{S} = 0, \quad (11.3.4)$$

or

$$\int_S \vec{\nabla} \times \vec{E} \cdot d\vec{S} = -\frac{\partial}{\partial t} \int_S \vec{B} \cdot d\vec{S}. \quad (11.3.5)$$

Again, by shrinking the sides of the curve C (see figure 10) the area within the curve can be made as small as we wish. Therefore, assuming that \vec{B} is well behaved (does not behave as a δ -function, for example) the integral on the r.h.s. approaches zero. But applying Stokes theorem we find

$$\int_S \vec{\nabla} \times \vec{E} \cdot d\vec{S} = \oint_C \vec{E} \cdot d\vec{r} = 0. \quad (11.3.6)$$

Let \hat{t}_1 and \hat{t}_2 be the tangents to the curve in the media “1” and “2” respectively. Clearly, $\hat{t}_2 = -\hat{t}_1 = -\hat{t}$ (say). Then, if, as before, we define $E_{1t} = \vec{E}^{(1)} \cdot \hat{t}_1 = \vec{E}^{(1)} \cdot \hat{t}$ and $E_{2t} = -\vec{E}^{(2)} \cdot \hat{t}_2 = \vec{E}^{(2)} \cdot \hat{t}$, the integral yields

$$\oint_C \vec{E} \cdot d\vec{r} = (E_{1t} - E_{2t})\Delta l = 0, \quad (11.3.7)$$

whence it follows that *the tangential component of the electric field must be continuous across the boundary between the media.*

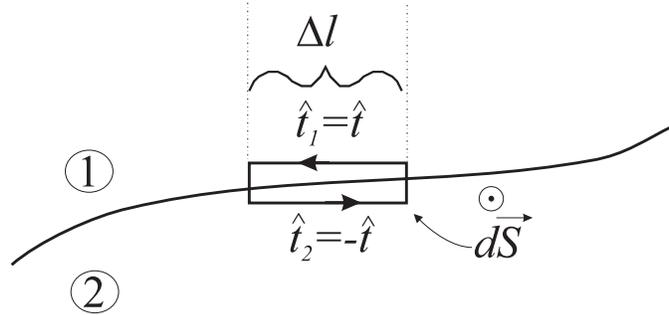


Figure 11.5: Boundary conditions for the tangential component

These two conditions were derived using the Bianchi identities. Let us now see what the dynamical equations have to say. These are

$$\begin{aligned}\vec{\nabla} \cdot \vec{D} &= \rho_f \\ \vec{\nabla} \times \vec{H} - \frac{\partial \vec{D}}{\partial t} &= \vec{j}_f\end{aligned}\quad (11.3.8)$$

and involve the sources. We expect, therefore, that these equations will imply, under certain conditions, the discontinuity of some of the components of the electric and magnetic field

- Beginning with the first equation, consider its integral form (see figure 9)

$$\oint_S \vec{D} \cdot d\vec{S} = \int_V \rho_f dV = Q_{\text{cont}}^f, \quad (11.3.9)$$

where V is the volume of the pill-box, S is its bounding surface and Q_{cont}^f is the free charge contained in the pill-box. If, as before, we shorten the box so that it becomes infinitesimally thick then the charge contained will simply be the free charge on the surface between the media and contained within the pill-box. Using the same notation as we had before, let σ_f be the surface density of the free charge. Then

$$(D_{1n} - D_{2n})\Delta S = \sigma_f \Delta S \rightarrow D_{1n} - D_{2n} = \sigma_f. \quad (11.3.10)$$

We conclude that *the normal component of the electric displacement vector is discontinuous across a boundary between two media if free surface charges are present, the discontinuity being proportional to the surface density of the free charge.*

- Finally, integrate the second equation (see figure 10) to get

$$\oint_C \vec{H} \cdot d\vec{r} = \frac{\partial}{\partial t} \int_S \vec{D} \cdot d\vec{S} + \int_S \vec{j}_f \cdot d\vec{S} \quad (11.3.11)$$

and assume that the electric displacement vector is well behaved, then making the thickness of the curve C infinitesimal we find that the first integral on the r.h.s. vanishes and we are left with

$$\oint_C \vec{H} \cdot d\vec{r} = \int_S \vec{j}_f \cdot d\vec{S}. \quad (11.3.12)$$

Now *if the free current density is well behaved* then the right hand side is zero because the area contained within the curve is vanishing and we conclude that *the tangential component of the magnetic intensity vector is continuous across the boundary*. However, if the current density is not well behaved (this occurs, for example in an ideal conductor where the conductivity goes to infinity) then the last integral need not vanish. Call

$$j_{\perp} \Delta l = \int_S \vec{j}_f \cdot d\vec{S}. \quad (11.3.13)$$

Then we obviously find a discontinuity in the tangential component of \vec{H}

$$H_{1t} - H_{2t} = j_{\perp}. \quad (11.3.14)$$

These four conditions summarize the boundary conditions appropriate to the electric and magnetic fields at the boundary between two media. We shall have occasion to use them shortly.

11.4 Poynting's theorem: energy and momentum density

We know that the electrostatic potential energy of a system of charges producing an electric field is given by

$$U_E = \frac{1}{2} \int_{\text{Vol}} d^3\vec{r} \vec{E} \cdot \vec{D}, \quad (11.4.1)$$

where "Vol" is the electric field external to all conductors. Therefore

$$u_E = \frac{1}{2} \vec{E} \cdot \vec{D} \quad (11.4.2)$$

can be thought of as the electric energy density. Similarly,

$$U_M = \frac{1}{2} \int_{\text{Vol}} d^3\vec{r} \vec{H} \cdot \vec{B} \quad (11.4.3)$$

(where “Vol” is now *all* of space) is the energy of the magnetic field with density

$$u_M = \frac{1}{2} \vec{H} \cdot \vec{B}. \quad (11.4.4)$$

Consider the equations

$$\begin{aligned} \vec{\nabla} \times \vec{E} + \frac{\partial \vec{B}}{\partial t} &= 0 \\ \vec{\nabla} \times \vec{H} - \frac{\partial \vec{D}}{\partial t} &= \vec{j}_f \end{aligned} \quad (11.4.5)$$

and take the scalar product of the first with \vec{H} ,⁴

$$\vec{H} \cdot (\vec{\nabla} \times \vec{E}) + \vec{H} \cdot \frac{\partial \vec{B}}{\partial t} = \vec{\nabla} \cdot (\vec{E} \times \vec{H}) + \vec{E} \cdot (\vec{\nabla} \times \vec{H}) + \vec{H} \cdot \frac{\partial \vec{B}}{\partial t} = 0, \quad (11.4.6)$$

so that

$$\vec{\nabla} \cdot (\vec{E} \times \vec{H}) = -\vec{H} \cdot \frac{\partial \vec{B}}{\partial t} - \vec{E} \cdot (\vec{\nabla} \times \vec{H}) = -\vec{H} \cdot \frac{\partial \vec{B}}{\partial t} - \vec{E} \cdot \frac{\partial \vec{D}}{\partial t} - \vec{j}_f \cdot \vec{E}, \quad (11.4.7)$$

where we have used

$$\vec{\nabla} \times \vec{H} = \frac{\partial \vec{D}}{\partial t} + \vec{j}_f \quad (11.4.8)$$

in the last step above. Now, if the media are isotropic, linear and non-dispersive then $\vec{D} = \epsilon \vec{E}$ and $\vec{B} = \mu \vec{H}$ so that

$$\vec{E} \cdot \frac{\partial \vec{D}}{\partial t} = \epsilon \vec{E} \cdot \frac{\partial \vec{E}}{\partial t} = \frac{1}{2} \epsilon \frac{\partial \vec{E}^2}{\partial t} = \frac{1}{2} \frac{\partial}{\partial t} (\vec{E} \cdot \vec{D}) \quad (11.4.9)$$

and similarly

$$\vec{H} \cdot \frac{\partial \vec{B}}{\partial t} = \frac{1}{2} \frac{\partial}{\partial t} (\vec{H} \cdot \vec{B}). \quad (11.4.10)$$

It follows that

$$\vec{\nabla} \cdot (\vec{E} \times \vec{H}) = -\frac{1}{2} \frac{\partial}{\partial t} (\vec{E} \cdot \vec{D} + \vec{H} \cdot \vec{B}) - \vec{j}_f \cdot \vec{E}. \quad (11.4.11)$$

⁴The first term is

$$\vec{H} \cdot (\vec{\nabla} \times \vec{E}) = \epsilon^{ijk} H_i (\partial_j E_k) = \partial_j (\epsilon^{ijk} H_i E_k) - \epsilon^{ijk} E_k \partial_j H_i$$

so

$$\vec{H} \cdot (\vec{\nabla} \times \vec{E}) = \vec{\nabla} \cdot (\vec{E} \times \vec{H}) + \vec{E} \cdot (\vec{\nabla} \times \vec{H})$$

Given the expressions in (11.4.2) and (11.4.4) for the electric and magnetic energy densities, it certainly makes sense to define an “energy density” of the electromagnetic field by

$$u = \frac{1}{2}(\vec{E} \cdot \vec{D} + \vec{H} \cdot \vec{B}) = u_E + u_M. \quad (11.4.12)$$

Let’s call the vector $\vec{S} = \vec{E} \times \vec{H}$, then (11.4.11) reads

$$\frac{\partial u}{\partial t} + \vec{\nabla} \cdot \vec{S} = -\vec{j}_f \cdot \vec{E} \quad (11.4.13)$$

In the absence of charges, $\vec{j}_f = 0$ and this equation looks strikingly like a continuity equation. The r.h.s. may therefore be thought of as a “source” term. We can understand it’s significance by integrating over a volume V bounded by some surface Σ ; then

$$\frac{d}{dt} \int_V d^3\vec{r} u + \oint_{\Sigma} d\vec{\sigma} \cdot \vec{S} = - \int_V d^3\vec{r} \vec{j}_f \cdot \vec{E}, \quad (11.4.14)$$

and

- the first term represents the rate of change of electromagnetic energy stored in the volume V ,
- the second term represents the rate of flow of energy across the bounding surface, Σ , and
- the source term should represent the rate at which the energy is transferred either *into* or *out of* the volume V , depending upon its sign.

To understand the meaning of the source term better, consider the work done per unit time by the electromagnetic (Lorentz) force \vec{F} acting on a single charged particle:

$$\vec{F} \cdot \vec{v} = q(\vec{E} + \vec{v} \times \vec{B}) \cdot \vec{v} = q\vec{v} \cdot \vec{E}. \quad (11.4.15)$$

If there were N particles, with charges q_n and velocities \vec{v}_n , then the work done per unit time would be

$$\sum_n q_n \vec{v}_n \cdot \vec{E}, \quad (11.4.16)$$

but, by definition, the current density is

$$\vec{j}_f(\vec{r}, t) = \sum_n q_n \vec{v}_n \delta^3(\vec{r} - \vec{r}_n(t)) \quad (11.4.17)$$

giving, for the volume integral of interest,

$$\int_V d^3\vec{r} \vec{j}_f \cdot \vec{E} = \sum_n \int_V d^3\vec{r} q_n \vec{v}_n \cdot \vec{E} \delta^3(\vec{r} - \vec{r}_n(t)) = \sum_n q_n \vec{v}_n \cdot \vec{E}(\vec{r}_n(t)), \quad (11.4.18)$$

which, as we have seen, is precisely the work performed by the electromagnetic field per second. The r.h.s. of (11.4.14) is the negative of this, so it is the work performed *upon* (or transferred to) the electromagnetic field per second by the transport currents.

The vector $\vec{E} \times \vec{H}$ represents a momentum density, *i.e.*, momentum per unit area, for the electromagnetic field. It is called the Poynting vector.

Chapter 12

Lagrangians and Hamiltonians*

In a modern approach, it is insufficient to formulate a theory exclusively in terms of the equations of motion. One must always attempt to formulate it in terms of an “action” principle. While the action itself does not have direct physical content in classical physics (it does in the quantum theory, as a probability amplitude for classical configurations), it possesses all the local symmetries of the theory. Symmetries lead to conservation laws, and these are derived most elegantly from the action. Moreover, the quantum theory is so heavily dependent on the canonical formulation that the use of an action is indispensable.

12.1 Lagrangian description of mechanics

Although our goal is to arrive at a fully relativistic formulation of electrodynamics in terms of an action principle, it is worth remembering first how the mechanics of point particles is formulated.¹ Consider a system having f degrees of freedom and described by a set of coordinates, q_i , ($i \in \{1, 2, 3 \dots N\}$) and velocities, \dot{q}_i . The Lagrangian is a function of the coordinates and velocities,

$$\mathcal{L} = \mathcal{L}(q_i, \dot{q}_i, t) \tag{12.1.1}$$

and the action is the functional

$$\mathcal{S}[q] = \int_1^2 dt \mathcal{L}(q_i, \dot{q}_i, t) \tag{12.1.2}$$

where “1” and “2” are fixed points in the configuration space and only trajectories, $q_i(t)$, that begin at “1” and terminate at “2” will be considered. Hamilton’s principle declares

¹Our treatment will only emphasize those aspects that are useful to achieve our goal. For a complete treatment the reader should consult one of the many excellent texts in classical mechanics.

that the classical trajectories are obtained by requiring the action to be stationary (extremized) if

$$\mathcal{L} = T(q_i, \dot{q}_i, t) - V(q_i, \dot{q}_i, t) \quad (12.1.3)$$

where T is the total kinetic energy of the system and V is its total potential energy. Suppose that all the coordinates are independent, so that $f = N$. To find the stationary points of the action functional one considers infinitesimal variations of the coordinates. Now variations may be of two types: (a) variation of the functional form

$$q_i(t) \rightarrow q'(t) = q_i(t) + \delta_0 q_i(t) = q_i(t) + \alpha \eta(t) \quad (12.1.4)$$

where α is some small parameter and $\eta(t)$ is an arbitrary function, and (b) variation of the parameter t :

$$t \rightarrow t' = t + \epsilon(t) \quad (12.1.5)$$

where $\epsilon(t)$ is an arbitrary but infinitesimal change in t . A variation of the parameter can be expected to induce a variation of the configuration space variables according to

$$\delta_1 q_i(t) = q_i(t + \epsilon(t)) - q_i(t) = \epsilon(t) \dot{q}_i(t) \quad (12.1.6)$$

In general, therefore, an arbitrary variation is made up of two parts: (i) a part that is purely functional, which we represent by δ_0 , and (ii) a part that arises from a change in the (time) parameter, which we denote by δ_1 ,

$$q_i(t) \rightarrow q'_i(t') = q'_i(t) + \epsilon \dot{q}_i(t) = q_i(t) + \delta_0 q_i(t) + \delta_1 \dot{q}_i(t) \quad (12.1.7)$$

giving the “total” variation,

$$\delta q(t) = q'_i(t') - q_i(t) = \delta_0 q_i(t) + \delta_1 \dot{q}_i(t) \quad (12.1.8)$$

up to first order in the variations and which we will henceforth represent by δ .

Our first goal is to determine the configurations that lead to a stationary action functional. This is achieved by considering a special sub-set of the possible variations, *viz.* functional variations of the coordinates that vanish at the end points (at t_1 and t_2 , *i.e.*, $\eta(t_1) = \eta(t_2) = 0$). Under such a variation, the action suffers the change

$$\begin{aligned} \delta \mathcal{S}[q] &= \int_1^2 dt \left[\frac{\partial \mathcal{L}}{\partial q_i} \delta_0 q_i + \frac{\partial \mathcal{L}}{\partial \dot{q}_i} \delta_0 \dot{q}_i \right] \\ &= \int_1^2 dt \left[\frac{\partial \mathcal{L}}{\partial q_i} \delta_0 q_i + \frac{\partial \mathcal{L}}{\partial \dot{q}_i} \delta_0 \frac{dq_i}{dt} \right] \\ &= \int_1^2 dt \left[\frac{\partial \mathcal{L}}{\partial q_i} \delta_0 q_i + \frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{q}_i} \delta_0 q_i \right) - \frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{q}_i} \right) \delta_0 q_i \right] \end{aligned}$$

$$= \left. \frac{\partial \mathcal{L}}{\partial \dot{q}_i} \delta_0 q_i \right|_1^2 + \int_1^2 dt \left[\frac{\partial \mathcal{L}}{\partial q_i} \delta_0 q_i - \frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{q}_i} \right) \delta_0 q_i \right] \quad (12.1.9)$$

Now the first term vanishes because the functional variations, $\delta_0 q_i$, are chosen to vanish on the boundaries. Thus

$$\delta \mathcal{S}[q] = \int_1^2 dt \left[\frac{\partial \mathcal{L}}{\partial q_i} - \frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{q}_i} \right) \right] \delta_0 q_i \quad (12.1.10)$$

According to Hamilton's principle, the action is required to be stationary ($\delta \mathcal{S} = 0$), then the integral on the r.h.s. is vanishing, which is possible only if the integrand vanishes. However, because $\delta_0 q_i$ is otherwise arbitrary, this means that

$$\frac{\partial \mathcal{L}}{\partial q_i} - \frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{q}_i} \right) = 0 \quad (12.1.11)$$

which are the so-called Euler equations.² We have assumed, of course, that the Lagrangian contains at most the first derivatives of the generalized coordinates. If it contains higher derivatives, appropriate modifications to the above derivation of the Euler equations must be made.

If the system is subject to external constraints, then the coordinates, q_i , are not all independent and there are relations between them. Suppose that there are M such relations (constraints) so that the number of degrees of freedom is $f = N - M$. We will consider only the case of holonomic constraints, or constraints that can be reduced by integration to holonomic constraints, in which case they can be given by M relations of the form $g_a(q_i, t) = 0$, between the coordinates only (and not the velocities). Then the Lagrangian must be augmented with Lagrange multipliers, λ_a , one for each relation, in defining the action:

$$\mathcal{S}[q] = \int_1^2 dt \left[\mathcal{L} + \sum_a \lambda_a g_a(q_i, t) \right] \quad (12.1.12)$$

and variations must be performed treating the coordinates q_i as independent. Variations w.r.t. λ_a must also be performed and they simply return the constraints,

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial q_i} - \frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{q}_i} \right) + \sum_a \lambda_a \frac{\partial g_a(q_i, t)}{\partial q_i} &= 0 \\ g_a(q_i, t) &= 0 \end{aligned} \quad (12.1.13)$$

These are the Euler-Lagrange equations of motion.

²Problem: Check that the Euler equations are Newton's equations for a single particle with kinetic energy $\frac{1}{2}m\vec{v}^2$ and potential energy $V(\vec{r})$.

12.2 Noether's theorem in Mechanics

As mentioned in the introduction to this section, a powerful way to view the conservation laws of mechanics is to think of them as resulting from the symmetries of the Lagrangian. Let us see how this works for a system with a finite number of degrees of freedom. Consider the change in the action under an arbitrary variation of the parameter t , *i.e.*, $t \rightarrow t' = t + \epsilon(t)$ as well as a functional variation of the coordinates, $q_i(t)$. We find

$$\delta\mathcal{S} = \int dt[\dot{\epsilon}(t)\mathcal{L} + \delta\mathcal{L}] = \int dt[\dot{\epsilon}(t)\mathcal{L} + \epsilon(t)\dot{\mathcal{L}} + \delta_0\mathcal{L}] = \int dt \left[\frac{d}{dt}(\epsilon\mathcal{L}) + \delta_0\mathcal{L} \right] \quad (12.2.1)$$

where we have used $dt \rightarrow dt' = dt(1 + \dot{\epsilon})$, or $\delta(dt) = \delta_1(dt) = dt\dot{\epsilon}$ and $\delta_1\mathcal{L} = \epsilon(t)\dot{\mathcal{L}}$. But we know how to write $\delta_0\mathcal{L}$, so

$$\delta\mathcal{S} = \int dt \left[\frac{d}{dt}(\epsilon\mathcal{L}) + \frac{\partial\mathcal{L}}{\partial q_i}\delta_0q_i + \frac{\partial\mathcal{L}}{\partial \dot{q}_i}\delta_0\dot{q}_i \right] \quad (12.2.2)$$

It is easy to see, from the definition of the functional derivative, that the operation of taking a time derivative commutes with the operation of making a functional variation, *i.e.*, $[\delta_0, \frac{d}{dt}]q_i(t) = 0$, so

$$\begin{aligned} \delta\mathcal{S} &= \int_1^2 dt \left[\frac{d}{dt}(\epsilon\mathcal{L}) + \frac{\partial\mathcal{L}}{\partial q_i}\delta_0q_i + \frac{d}{dt} \left(\frac{\partial\mathcal{L}}{\partial \dot{q}_i}\delta_0q_i \right) - \frac{d}{dt} \left(\frac{\partial\mathcal{L}}{\partial \dot{q}_i} \right) \delta_0q_i \right] \\ &= \int_1^2 dt \left[\frac{d}{dt}(\epsilon\mathcal{L}) + \frac{d}{dt} \left(\frac{\partial\mathcal{L}}{\partial \dot{q}_i}\delta_0q_i \right) \right] = \int_1^2 dt \frac{d}{dt} \left[(\epsilon\mathcal{L}) + \left(\frac{\partial\mathcal{L}}{\partial \dot{q}_i}\delta_0q_i \right) \right] \end{aligned} \quad (12.2.3)$$

after using Euler's equations. Note that we have not set the total derivative term

$$\int_1^2 dt \frac{d}{dt} \left(\frac{\partial\mathcal{L}}{\partial \dot{q}_i}\delta_0q_i \right)$$

directly to zero because we do not assume that the functional variation vanishes at the boundary. It is convenient at this point to exchange the functional variation of $q_i(t)$ for a total variation according to

$$\delta_0q_i(t) = \delta q_i(t) - \delta_1q_i(t) = \delta q_i(t) - \epsilon(t)\dot{q}_i(t) \quad (12.2.4)$$

and write

$$\delta\mathcal{S} = \int_1^2 dt \frac{d}{dt} \left[\left(\mathcal{L} - \frac{\partial\mathcal{L}}{\partial \dot{q}_i}\dot{q}_i \right) \delta t + \left(\frac{\partial\mathcal{L}}{\partial \dot{q}_i} \right) \delta q_i \right] \quad (12.2.5)$$

where we have called $\epsilon(t) = t'(t) - t = \delta t$. This is done only to maintain the same notation in both terms. Imagine that the variations result from some global transformations and

depend on a certain set of parameters, ω^a , $a \in \{1, 2, \dots\}$. These global transformations could be, for example, translations, rotations, boosts, etc. We can write

$$\delta\mathcal{S} = \int_1^2 dt \frac{d}{dt} \left[\left(\mathcal{L} - \frac{\partial\mathcal{L}}{\partial\dot{q}_i} \dot{q}_i \right) \frac{\delta t}{\delta\omega^a} + \left(\frac{\partial\mathcal{L}}{\partial\dot{q}_i} \right) \frac{\delta q_i}{\delta\omega^a} \right] \delta\omega^a \quad (12.2.6)$$

If $\mathcal{S}[q]$ is invariant under the global transformation, we have $\delta\mathcal{S}[q] = 0$ and therefore (ω^a is arbitrary)

$$\left(\mathcal{L} - \frac{\partial\mathcal{L}}{\partial\dot{q}_i} \dot{q}_i \right) \frac{\delta t}{\delta\omega^a} + \left(\frac{\partial\mathcal{L}}{\partial\dot{q}_i} \right) \frac{\delta q_i}{\delta\omega^a} \quad (12.2.7)$$

is a constant of the motion (it is conserved). This is “Noether’s theorem” applied to systems with a finite number of degrees of freedom. Let us examine some special consequences:

- Time translations: Take $t \rightarrow t' = t + \epsilon$ ($\epsilon = \delta\omega^a$ constant) and $q_i(t) \rightarrow q'_i(t') = q_i(t)$ (trajectories do not depend on the choice of the time origin) so $\delta q = 0$. If \mathcal{S} is invariant under time translations, $\delta\mathcal{S} = 0$ and

$$\mathcal{H} = \frac{\partial\mathcal{L}}{\partial\dot{q}_i} \dot{q}_i - \mathcal{L} \quad (12.2.8)$$

is a constant of the motion. The quantity

$$p_i = \frac{\partial\mathcal{L}}{\partial\dot{q}_i} \quad (12.2.9)$$

is the “momentum” conjugate to the coordinate q_i . \mathcal{H} is, of course, the “Hamiltonian”

$$\mathcal{H} = p_i \dot{q}_i - \mathcal{L} \quad (12.2.10)$$

of the system which, in this case, is interpreted as its energy. Thus, energy conservation is a consequence of time translation invariance.

- Spatial translations: Take $q_i \rightarrow q_i + a_i$ and $t \rightarrow t'(t) = t$, so that $\delta\omega^a = a_i = \delta q_i$ and $\delta t = 0$. If the action is invariant under spatial translations of some or all of the N coordinates then the momenta,

$$p_i = \frac{\partial\mathcal{L}}{\partial\dot{q}_i} \quad (12.2.11)$$

conjugate to those coordinates is conserved. Thus momentum conservation in any direction is a consequence of space translation invariance of the action. A coordinate whose conjugate momentum is conserved is called cyclic.

- Spatial Rotations: Consider now a single particle with coordinates $q_i = x_i$ (the usual Cartesian coordinates). General rotations of the coordinate system may be written as products of rotations about the individual axes, of which there are three. If we consider an infinitesimal rotation, for example, about the z -axis we have

$$\mathbb{R}_{ij}^z = \delta_{ij} + \delta\theta^3 \mathbb{U}_{ij}^3 \quad (12.2.12)$$

where

$$\mathbb{U}_{ij}^3 = \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} = [\epsilon_3]_{ij} \quad (12.2.13)$$

is the “generator” of rotations about the z -axis. Here $[\epsilon_k]_{ij} = \epsilon_{kij}$ is, of course, the Levi-Civita tensor, and, for example, $[\epsilon_3]_{ij}$ is to be thought of as a matrix \hat{e}_3 , whose components are given by the Levi Civita symbol ϵ_{3ij} . For an infinitesimal rotation about the x -axis we would have

$$\mathbb{R}_{ij}^x = \delta_{ij} + \delta\theta^1 \mathbb{U}_{ij}^1 \quad (12.2.14)$$

where

$$\mathbb{U}_{ij}^1 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{bmatrix} = [\epsilon_1]_{ij} \quad (12.2.15)$$

and, finally about the y axis,

$$\mathbb{R}_{ij}^y = \delta_{ij} + \delta\theta^2 \mathbb{U}_{ij}^2 \quad (12.2.16)$$

where

$$\mathbb{U}_{ij}^2 = \begin{bmatrix} 0 & 0 & -1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} = [\epsilon_2]_{ij} \quad (12.2.17)$$

An arbitrary infinitesimal rotation of the coordinates x^i of our particle would therefore take the form (sum over repeated indices)

$$\delta x_i = \delta\theta^k [\epsilon_k]_{ij} x_j \quad (12.2.18)$$

and the time parameter remains unchanged, $t' = t$ or $\delta t = 0$. The parameters of the transformation are the angles θ^k and

$$\frac{\delta x_i}{\delta\theta^k} = \epsilon_{kij} x_j \quad (12.2.19)$$

It follows from Noether's theorem that, if the action for the particle is invariant under spatial rotations, the quantity

$$L_k = -\epsilon_{kij} p_i x_j = (\vec{r} \times \vec{p})_k \quad (12.2.20)$$

(which will be recognized as the angular momentum of the particle) is conserved. Conservation of angular momentum is a consequence of invariance of the action under spatial rotations. Similarly, if the action is invariant under Lorentz transformations (is a Lorentz scalar) then a four dimensional generalization of the above that is appropriate to the Lorentz group is obtained.

With these examples, the power of Noether's theorem should be evident. Let us turn now to another, equivalent, formulation of mechanics, which treats the Hamiltonian as the fundamental object of interest.

12.3 Hamiltonian description of Mechanics

In the previous section we had seen that when the Lagrangian has no explicit dependence on time then the quantity

$$\mathcal{H} = \frac{\partial \mathcal{L}}{\partial \dot{q}_i} \dot{q}_i - \mathcal{L} \quad (12.3.1)$$

is conserved and interpretable as the total energy of the system. The momentum conjugate to the coordinate q_i is defined to be

$$p_i = \frac{\partial \mathcal{L}}{\partial \dot{q}_i} \quad (12.3.2)$$

and the Hamiltonian is therefore written as

$$\mathcal{H} = p_i \dot{q}_i - \mathcal{L} \quad (12.3.3)$$

Equation (12.3.3) can be viewed as a Legendre transformation from the variables (\dot{q}_i, q_i) to the variables (p_i, q_i) . If we think of p_i as an independent variable and the definition (12.3.2) as giving the velocities in terms of the momenta, then in this form \mathcal{H} depends on (p_i, q_i) . This is easy to see, because

$$\frac{\partial \mathcal{H}}{\partial \dot{q}_i} = p_i - \frac{\partial \mathcal{L}}{\partial \dot{q}_i} \equiv 0 \quad (12.3.4)$$

We do something similar in thermodynamics: recall that all the thermodynamic potentials are obtained by Legendre transformations of the internal energy.³⁾

³From the first law of thermodynamics, the internal energy, U , may be thought of as a function of the state variables (S, V) with $T = \partial U / \partial S$ and $p = -\partial U / \partial V$. An example of a Legendre transformation is

The Hamiltonian function provides an alternative description of the dynamics. Consider the action functional expressed in terms of the Hamiltonian as follows

$$\mathcal{S} = \int_1^2 dt \mathcal{L}(q_i, \dot{q}_i, t) = \int_1^2 dt [p_i \dot{q}_i - \mathcal{H}(q_i, p_i, t)] \quad (12.3.5)$$

and perform a variation of the action, treating p_i and q_i (but not \dot{q}_i) as independent variables and holding *both* p_i and q_i fixed at the end points. It is easy to see that, if the action is stationary under variations w.r.t. p_i then

$$\dot{q}_i = \frac{\partial \mathcal{H}}{\partial p_i}, \quad (12.3.6)$$

and if it is stationary under variations w.r.t. q_i then (remember to hold q_i fixed at the end points)

$$\dot{p}_i = -\frac{\partial \mathcal{H}}{\partial q_i}. \quad (12.3.7)$$

These are the so-called Hamiltonian equations of motion or simply Hamilton's equations. The first equation confirms the fact that if $\mathcal{H}(q_i, p_i, t)$ is a Legendre transformation of $\mathcal{L}(q_i, \dot{q}_i, t)$ then $\mathcal{L}(q_i, \dot{q}_i, t)$ is a Legendre transformation of $\mathcal{H}(q_i, p_i, t)$. We see that the Hamiltonian equations of motion are a set of first order differential equations as opposed to the Euler-Lagrange equations which are second order differential equations. However, there are $2n$ coupled equations in the Hamiltonian description, whereas there are only n coupled equations in the Lagrangian description. The two descriptions are equivalent.

A convenient way to write these equations is by introducing the "Poisson bracket". Suppose $A(q_i, p_i, t)$ and $B(q_i, p_i, t)$ are two arbitrary functions on phase-space. The Poisson bracket between A and B is defined by

$$\{A, B\}_{P.B.} = \sum_i \left(\frac{\partial A}{\partial q_i} \frac{\partial B}{\partial p_i} - \frac{\partial A}{\partial p_i} \frac{\partial B}{\partial q_i} \right) \quad (12.3.8)$$

Now q_i and p_i are certainly functions on the phase space. We see that

$$\{q_i, p_j\}_{P.B.} = \delta_{ij} \quad (12.3.9)$$

the Helmholtz free energy, $F = U - TS$, of a system. The free energy is a function of (T, V) . This is seen from

$$F = U - TS \quad \rightarrow \quad dH = dU - TdS - SdT = -SdT - pdV$$

where we have used the first law, $dU = TdS - pdV$. Another example is the enthalpy $H = U + pV$. It should be clear that $H = H(S, p)$. Finally, making a double Legendre transformation, $G = U + pV - TS$ we get the Gibbs potential which must therefore be a function of the variables (p, T) . Verify the dependencies of that last two thermodynamic potentials explicitly by using the first law.

Moreover, a simple calculation will show that

$$\dot{q}_i = \{q_i, \mathcal{H}\}_{P.B.}, \quad \text{and} \quad \dot{p}_i = \{p_i, \mathcal{H}\}_{P.B.} \quad (12.3.10)$$

are different ways of expressing Hamilton's equations. Indeed, the time evolution of *any* phase-space function, $A(q_i, p_i, t)$ is seen to be given simply by⁴

$$\dot{A} = \{A, \mathcal{H}\}_{P.B.} + \frac{\partial A}{\partial t} \quad (12.3.11)$$

The Hamiltonian formulation of the classical dynamics of any physical theory is essential to the construction of its corresponding quantum theory.⁵

We wish to generalize these considerations to continuous systems, *i.e.*, systems with an infinite number of degrees of freedom. This is the subject of the following sections.

12.4 Lagrangian description of fields

The fundamental difference between a “field” and a system of point particles is that a field has an uncountably infinite number of degrees of freedom, a countable number at each point in space. In this sense the Lagrangian, \mathcal{L} , is itself a functional, an integral over space of some function (the Lagrange *density* function, \mathfrak{L}) of the fields and their derivatives. How this comes to be is best understood by an example.

Consider an infinitely long elastic rod laid along the x -axis, that is able to sustain oscillatory displacements of the particles composing it in a direction parallel to the rod. We “discretize” the rod by imagining that it is made up of a very large number of point like particles that are spaced a distance a apart and connected by massless springs (which simulate the interaction between neighboring atoms in the rod). Consider the n th particle and denote its displacement from equilibrium by η_n . Its kinetic energy is

$$T_n = \frac{1}{2} m_n \dot{\eta}_n^2 \quad (12.4.1)$$

⁴Problem: Prove this.

⁵The quantum theory can be thought of as arising from the non-commutativity of the phase-space variables q_i and p_i . This is realized by promoting the observables (phase-space functions like $A(q_i, p_i, t)$) to operators on a Hilbert space and replacing of the classical Poisson brackets by commutators. Thus, for example,

$$\{q_i, p_j\}_{P.B.} = \delta_{ij} \quad \rightarrow \quad [q_i, p_i] = i\delta_{ij} \quad (12.3.12)$$

and

$$\dot{A} = \{A, \mathcal{H}\}_{P.B.} + \frac{\partial A}{\partial t} \quad \rightarrow \quad \dot{A} = i[A, \mathcal{H}] + \frac{\partial A}{\partial t} \quad (12.3.13)$$

in the Heisenberg picture.

giving, for all the particles together, the total kinetic energy

$$T = \frac{1}{2} \sum_n m_n \dot{\eta}_n^2 \quad (12.4.2)$$

The total potential energy could be written as the sum of the potential energies of the springs which are stretched (or compressed) according to the displacements of neighboring particles:

$$V = \frac{1}{2} \sum_n k_n (\eta_{n+1} - \eta_n)^2 \quad (12.4.3)$$

The Lagrangean for the entire system follows from Hamilton's prescription:

$$\mathcal{L} = T - V = \frac{1}{2} \sum_n m_n [\dot{\eta}_n^2 - \omega_n^2 (\eta_{n+1} - \eta_n)^2] \quad (12.4.4)$$

which will be convenient to write the the following form

$$\mathcal{L} = \frac{1}{2} \sum_n a \left(\frac{m_n}{a} \right) \left[\dot{\eta}_n^2 - (a\omega_n)^2 \left(\frac{\eta_{n+1} - \eta_n}{a} \right)^2 \right] \quad (12.4.5)$$

Now consider taking the limit as $a \rightarrow 0$ of the above. The quantity m_n/a can be interpreted as the mass per unit length or the linear mass density of the rod, which we take to be some constant, μ . It is clear that in the same limit,

$$\lim_{a \rightarrow 0} \frac{\eta_{n+1} - \eta_n}{a} = \lim_{a \rightarrow 0} \frac{\eta(t, x+a) - \eta(t, x)}{a} = \frac{\partial}{\partial x} \eta(t, x) \quad (12.4.6)$$

and that

$$\dot{\eta}_n \rightarrow \frac{\partial}{\partial t} \eta(t, x), \quad (12.4.7)$$

but what about $a\omega_n^2$? A little thought will show that this is related to the Young's modulus, Y , of the rod by

$$am_n\omega_n^2 = Y \quad (12.4.8)$$

which we'll take to be independent of position. Thus, in the continuum limit, the Lagrangian function can be written as

$$\mathcal{L}(\dot{\eta}, \eta', x) = \int_{-\infty}^{\infty} dx \left[\frac{\mu}{2} \left(\frac{\partial \eta(t, x)}{\partial t} \right)^2 - Y \left(\frac{\partial \eta(t, x)}{\partial x} \right)^2 \right] \quad (12.4.9)$$

The quantity within square brackets is a density function (in this case a *linear* density) called the Lagrange density, \mathfrak{L} , and

$$\mathcal{S} = \int dt \int_{-\infty}^{\infty} dx \mathfrak{L}(\dot{\eta}(t, x), \eta'(t, x), x, t). \quad (12.4.10)$$

is the action. $\eta(t, x)$ is a “field” (the field of displacements from equilibrium of the rod’s constituents, or the elementary excitations of the rod)

We may now consider a general field theory as one that is described by a Lagrangian *functional*,

$$\mathcal{L}[\phi^A, \partial_\mu \phi^A, t] = \int d^3\vec{r} \mathfrak{L}(\phi^A(t, \vec{r}), \partial_\mu \phi^A(t, \vec{r}), t, \vec{r}) \quad (12.4.11)$$

where $\phi^A(t, \vec{r})$ is a symbol that denotes a field exhibiting definite transformation properties in space-time, *i.e.*, which transforms either as a scalar (eg. Higgs field), vector (eg. Electromagnetic field, other gauge bosons of the standard model), spinor (eg. Dirac/Majorana field for fermions) or tensor (eg. gravitational field) and “ A ” represents the collection of indices carried by it. The field theory is specified by an action

$$\mathcal{S} = \int dt \mathcal{L}[\phi^A, \partial_\mu \phi^A, t] = \int d^4x \mathfrak{L}(\phi^A(x), \partial_\mu \phi^A(x), x) \quad (12.4.12)$$

from which the equations of motion are to be derived using Hamilton’s principle.

Following the arguments given earlier for point particles, we realize at the onset that field variations are also of two types: (a) functional variations,

$$\phi^A(x) \rightarrow \phi'^A(x) = \phi^A(x) + \delta_0 \phi^A(x) \quad (12.4.13)$$

and (b) variations that arise because of a change in the parameters (in this case the four space-time coordinates),

$$x^\mu \rightarrow x'^\mu = x^\mu + \epsilon^\mu(x) \quad (12.4.14)$$

where ϵ^μ represents an infinitesimal change in x^μ . This induces a change in the field according to

$$\delta_1 \phi^A(x) = \phi^A(x') - \phi^A(x) = \phi^A(x + \epsilon) - \phi^A(x) = \epsilon^\mu \partial_\mu \phi^A \quad (12.4.15)$$

A general change is thus made up of both components, *i.e.*,

$$\begin{aligned} \delta \phi^A(x) &= \phi'^A(x') - \phi^A(x) = \phi'^A(x + \epsilon) - \phi^A(x) \\ &= \phi'^A(x) - \phi^A(x) + \epsilon \cdot \partial \phi^A = \delta_0 \phi^A(x) + \delta_1 \phi^A(x) \end{aligned} \quad (12.4.16)$$

up to first order in the variations, of course. To apply Hamilton’s principle to the action we first consider only functional variations that vanish at the boundary (usually taken to be at infinity). Then

$$\delta \mathcal{S} = \int d^4x \delta \mathfrak{L} = \int d^4x \left[\frac{\partial \mathfrak{L}}{\partial \phi^A} \delta_0 \phi^A + \frac{\partial \mathfrak{L}}{\partial (\partial_\mu \phi^A)} \delta_0 \partial_\mu \phi^A \right]$$

$$\begin{aligned}
&= \int d^4x \left[\frac{\partial \mathcal{L}}{\partial \phi^A} \delta_0 \phi^A + \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial (\partial_\mu \phi^A)} \delta_0 \phi^A \right) - \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial (\partial_\mu \phi^A)} \right) \delta_0 \phi^A \right] \\
&= \int d^4x \left[\frac{\partial \mathcal{L}}{\partial \phi^A} - \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial (\partial_\mu \phi^A)} \right) \right] \delta_0 \phi^A \tag{12.4.17}
\end{aligned}$$

where we have interchanged ∂_μ and δ_0 , because $[\partial_\mu, \delta_0]\phi = 0$, and used the fact that the integral of the total derivative is a surface term (an integral over the bounding surface) which vanishes exactly

$$\int_M d^4x \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial (\partial_\mu \phi^A)} \delta_0 \phi^A \right) = \int_{\partial M} d\sigma_\mu \left(\frac{\partial \mathcal{L}}{\partial (\partial_\mu \phi^A)} \delta_0 \phi^A \right) \equiv 0 \tag{12.4.18}$$

by our condition that the variation $\delta_0 \phi$ vanishes there. But, as $\delta_0 \phi$ is otherwise arbitrary, and the action is stationary ($\delta \mathcal{S} = 0$) we necessarily arrive at Euler's equations

$$\frac{\partial \mathcal{L}}{\partial \phi^A} - \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial (\partial_\mu \phi^A)} \right) = 0 \tag{12.4.19}$$

for a field theory governed by the Lagrange density \mathcal{L} . Of course, Lagrange multipliers and the Euler-Lagrange equations may be used when constraints are involved.

We may apply these equations to the action we wrote down earlier for the elastic rod. The field is $\eta(t, x)$ and the Lagrange density does not depend on $\eta(t, x)$ but only on its derivatives. Therefore we have

$$\partial_t \left(\frac{\partial \mathcal{L}}{\partial \dot{\eta}} \right) + \partial_x \left(\frac{\partial \mathcal{L}}{\partial \eta'} \right) = 0 = \mu \frac{\partial^2 \eta}{\partial t^2} - Y \frac{\partial^2 \eta}{\partial x^2} \tag{12.4.20}$$

which is a wave equation for perturbations that travel at the speed of sound, $v_s = \sqrt{\frac{Y}{\mu}}$, in the rod.

12.5 Noether's theorem for fields

Noether's theorem is quite a bit more powerful when applied to fields than when it is applied to a system of point particles. The theorem simply states that to every symmetry of the action there exists a corresponding current that is conserved. Its proof follows the general lines of reasoning that were introduced when the number of degrees of freedom was finite (point particles). Therefore, consider a general variation, including coordinate changes, of the action in (12.4.12)

$$\delta \mathcal{S} = \int [(\delta d^4x) \mathcal{L} + d^4x \delta \mathcal{L}] \tag{12.5.1}$$

and we need to determine, first, δd^4x . This is just an infinitesimal change in coordinates

$$x'^\mu = x^\mu + \epsilon^\mu(x) \quad \rightarrow \quad \frac{\partial x'^\mu}{\partial x^\nu} = \delta_\nu^\mu + \partial_\nu \epsilon^\mu(x) \quad (12.5.2)$$

The change in measure, δd^4x is determined via the Jacobian

$$d^4x \rightarrow d^4x' = d^4x \det\left(\frac{\partial x'^\mu}{\partial x^\nu}\right) = d^4x \det(\delta_\nu^\mu + \partial_\nu \epsilon^\mu) = d^4x \det \hat{J} \quad (12.5.3)$$

where $\hat{J}_\nu^\mu = (\widehat{1 + \partial\epsilon})_\nu^\mu$ is the Jacobian of the transformation. But the determinant of a matrix \hat{J} is related to its trace according to⁶

$$\ln \det \hat{J} = \text{tr} \ln \hat{J} \quad (12.5.4)$$

so that

$$\det \hat{J} = e^{\text{tr} \ln \hat{J}} = e^{\text{tr} \ln(\widehat{1 + \partial\epsilon})} \approx e^{\text{tr} \widehat{\partial\epsilon}} \approx 1 + \partial_\mu \epsilon^\mu \quad (12.5.5)$$

It follows that

$$\delta d^4x = d^4x' - d^4x = d^4x(\partial \cdot \epsilon) \quad (12.5.6)$$

and therefore that

$$\begin{aligned} \delta \mathcal{S} &= \int d^4x [(\partial \cdot \epsilon) \mathcal{L} + \delta \mathcal{L}] = \int d^4x [(\partial \cdot \epsilon) \mathcal{L} + \epsilon \cdot \partial \mathcal{L} + \delta_0 \mathcal{L}] \\ &= \int d^4x \left[\partial_\mu (\mathcal{L} \epsilon^\mu) + \frac{\partial \mathcal{L}}{\partial \phi^A} \delta_0 \phi^A + \frac{\partial \mathcal{L}}{\partial (\partial_\mu \phi^A)} \delta_0 \partial_\mu \phi^A \right] \end{aligned} \quad (12.5.7)$$

Interchanging δ_0 and ∂_μ ($[\delta_0, \partial_\mu] \phi^A = 0$), we find

$$\delta \mathcal{S} = \int d^4x \left[\partial_\mu (\mathcal{L} \epsilon^\mu) + \frac{\partial \mathcal{L}}{\partial \phi^A} \delta_0 \phi^A + \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial (\partial_\mu \phi^A)} \delta_0 \phi^A \right) - \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial (\partial_\mu \phi^A)} \right) \delta_0 \phi^A \right] \quad (12.5.8)$$

which, if we use the equations of motion (Euler's equations), simplifies to

$$\delta \mathcal{S} = \int d^4x \partial_\mu \left[\mathcal{L} \epsilon^\mu + \frac{\partial \mathcal{L}}{\partial (\partial_\mu \phi^A)} \delta_0 \phi^A \right] \quad (12.5.9)$$

As before, exchange the functional variation of $\phi(x)$ above for a total variation and find

$$\delta \mathcal{S} = \int d^4x \partial_\mu \left[\left(\mathcal{L} \delta_\nu^\mu - \frac{\partial \mathcal{L}}{\partial (\partial_\mu \phi^A)} \partial_\nu \phi^A \right) \epsilon^\nu + \frac{\partial \mathcal{L}}{\partial (\partial_\mu \phi^A)} \delta \phi^A \right] \quad (12.5.10)$$

⁶This is easy to prove: let λ_n be the eigenvalues of \hat{J} , so

$$\det \hat{J} = \prod_n \lambda_n \rightarrow \ln \det \hat{J} = \ln \left(\prod_n \lambda_n \right) = \sum_n \ln \lambda_n = \text{tr} \ln \hat{J}$$

We imagine again that the variations result from some global transformations that depend on a certain set of constant parameters, ω^a , $a \in \{1, 2, \dots\}$. Then

$$\delta\mathcal{S} = \int d^4x \partial_\mu \left[\left(\mathfrak{L}\delta_\nu^\mu - \frac{\partial\mathfrak{L}}{\partial(\partial_\mu\phi^A)}\partial_\nu\phi^A \right) \frac{\delta x^\nu}{\delta\omega^a} + \frac{\partial\mathfrak{L}}{\partial(\partial_\mu\phi^A)} \frac{\delta\phi^A}{\delta\omega^a} \right] \delta\omega^a \quad (12.5.11)$$

If the action is invariant under the transformations, $\delta\mathcal{S} = 0$, then because the parameters are arbitrary it follows that

$$j_a^\mu = \left(\frac{\partial\mathfrak{L}}{\partial(\partial_\mu\phi^A)}\partial_\nu\phi^A - \mathfrak{L}\delta_\nu^\mu \right) \frac{\delta x^\nu}{\delta\omega^a} - \frac{\partial\mathfrak{L}}{\partial(\partial_\mu\phi^A)} \frac{\delta\phi^A}{\delta\omega^a} \quad (12.5.12)$$

is a *conserved* current, *i. e.*,

$$\partial_\mu j_a^\mu \equiv 0. \quad (12.5.13)$$

This is Noether's first theorem applied to fields.

As a special application, let's consider space-time translations: $x^\mu \rightarrow x'^\mu = x^\mu + \delta\epsilon^\mu$ (we can think of the constant parameters $\delta\omega^a$ as being $\delta\epsilon^\mu$) and suppose that $\delta\phi^A(x) = 0$ (since fields are Lorentz scalars, vectors or tensors, they must be invariant under constant translations) then

$$\frac{\delta x^\mu}{\delta\epsilon^\sigma} = \delta_\sigma^\mu \quad (12.5.14)$$

and

$$j_\sigma^\mu = \mathfrak{L}\delta_\sigma^\mu - \frac{\partial\mathfrak{L}}{\partial(\partial_\mu\phi^A)}\partial_\sigma\phi^A \quad (12.5.15)$$

is conserved. What we have obtained in an important conserved quantity, a (mixed) second rank tensor, obeying

$$\partial_\mu j_\sigma^\mu = 0 \quad (12.5.16)$$

called the "stress-energy" or "energy-momentum" tensor of the field.⁷

If it turns out that $j^{\mu\nu} = j_\sigma^\mu \eta^{\sigma\nu}$ is symmetric, the fact that it is conserved allows us to define a conserved 4-vector current that we can identify with the energy-momentum flow of the field,

$$P_\sigma = \int d^3\vec{r} \vec{\rho}_\sigma = \int d^3\vec{r} j_\sigma^0, \quad P^\mu = \int d^3\vec{r} \rho^\mu = \int d^3\vec{r} j^{0\mu} \quad (12.5.17)$$

where $j^{0\mu} = j_\sigma^0 \eta^{\sigma\mu}$. It is easy to check that⁸

$$\frac{dP^\mu}{dt} \equiv 0, \quad (12.5.18)$$

⁷Problem: Compute the stress-energy tensor of the field $\eta(t, x)$, representing the excitations of an elastic rod.

⁸Problem: Show that P_μ transforms as a 4-vector.

if the fields fall off sufficiently fast, because

$$\frac{dP^\mu}{dt} = \frac{d}{dt} \int d^3\vec{r} j^{0\mu} = \int d^3\vec{r} \partial_t j^{0\mu} = - \int d^3\vec{r} \partial_i j^{i\mu} = - \oint_S dS \hat{n}_i j^{i\mu}, \quad (12.5.19)$$

where S is a bounding surface at infinity and \hat{n}_i is normal to it. This is just Gauss' law. Assuming that the fields fall off fast enough that there is no energy-momentum flux across the surface at infinity, the surface integral vanishes and we have the desired result.

As mentioned, the stress tensor, $j^{\mu\nu}$, as defined above, is not generally symmetric in its indices (although it may be in some important cases). It is also not uniquely defined as a conserved quantity because we could add to it *any* divergence free, second rank tensor. In particular, the divergence of a third rank tensor which is antisymmetric in two indices, eg., a tensor $k^{\lambda\mu\nu}$ which is *antisymmetric* in (λ, μ) , gives a second rank tensor, $\Delta^{\mu\nu} = \partial_\lambda k^{\lambda\mu\nu}$ which satisfies this criterion. In other words, if we define

$$t^{\mu\nu} = j^{\mu\nu} + \Delta^{\mu\nu} = j^{\mu\nu} + \partial_\lambda k^{\lambda\mu\nu} \quad (12.5.20)$$

then $t^{\mu\nu}$ is still trivially conserved by the antisymmetry of $k^{\lambda\mu\nu}$.⁹

Suppose that, by doing the above, we could find a tensor, $t^{\mu\nu}$, that is *symmetric* in its indices, (μ, ν) , and conserved as well. It turns out that we can generally do this and we will see an example of it when we work with the electromagnetic field in the following chapter. Consider an infinitesimal Lorentz transformation, $\delta x^\mu = \delta\omega^{\mu\nu} x_\nu$. Then

$$\frac{\delta x^\mu}{\delta\omega^{\alpha\beta}} = \frac{1}{2} \left(\delta_\alpha^\mu \delta_\beta^\nu - \delta_\beta^\mu \delta_\alpha^\nu \right) x_\nu, \quad (12.5.21)$$

where we have exploited the antisymmetry of $\delta\omega^{\mu\nu}$. Suppose also that under the Lorentz transformation,

$$\frac{\delta\phi^A(x)}{\delta\omega^{\mu\nu}} = G_{\mu\nu}^A \quad (12.5.22)$$

where $G_{\mu\nu}^A$ is antisymmetric in (μ, ν) and depends on the transformation properties of ϕ^A . We now find that

$$M_{\alpha\beta}^\mu = \frac{1}{2} (j_\alpha^\mu x_\beta - j_\beta^\mu x_\alpha) + S_{\alpha\beta}^\mu \quad (12.5.23)$$

is conserved, where

$$S_{\alpha\beta}^\mu = \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi^A)} G_{\alpha\beta}^A \quad (12.5.24)$$

$M_{\alpha\beta}^\mu$ is called the "total angular momentum" tensor density and

$$L_{\alpha\beta}^\mu = \frac{1}{2} (j_\alpha^\mu x_\beta - j_\beta^\mu x_\alpha) \quad (12.5.25)$$

⁹Observe that $\partial_\mu \Delta^{\mu\nu} = \partial_\mu \partial_\lambda k^{\lambda\mu\nu} \equiv 0$ because the partial derivatives commute whereas k is antisymmetric in those indices.

is called the “orbital angular momentum” tensor density. $S_{\alpha\beta}^{\mu}$ is called the “intrinsic angular momentum” tensor density.

From the conservation of $M_{\alpha\beta}^{\mu}$ follows

$$\partial_{\mu} S_{\alpha\beta}^{\mu} = \frac{1}{2} (j_{\alpha\beta} - j_{\beta\alpha}) \quad (12.5.26)$$

which says that if $j_{\alpha\beta}$ is a symmetric tensor then \widehat{S} is separately conserved. Otherwise the intrinsic angular momentum can be exchanged for the orbital angular momentum and vice-versa. The extent to which this occurs is measured by the antisymmetric part of the canonical energy momentum tensor.

With the symmetrized stress-energy tensor, $t^{\mu\nu}$, let us define the modified angular momentum tensor density

$$\widetilde{L}_{\alpha\beta}^{\mu} = t_{\alpha}^{\mu} x_{\beta} - t_{\beta}^{\mu} x_{\alpha}, \quad (12.5.27)$$

and the modified intrinsic angular momentum tensor density

$$\widetilde{S}_{\alpha\beta}^{\mu} = S_{\alpha\beta}^{\mu} - \frac{1}{2} (\Delta_{\alpha}^{\mu} x_{\beta} - \Delta_{\beta}^{\mu} x_{\alpha}). \quad (12.5.28)$$

Both would be separately conserved.¹⁰ We then have

$$\int d^3\vec{r} \widetilde{L}^{0\alpha\beta} = \int d^3\vec{r} (x^{\beta} \varphi^{\alpha} - x^{\alpha} \varphi^{\beta}) \stackrel{\text{def}}{=} L^{\alpha\beta} \quad (12.5.29)$$

where $\varphi^{\mu} = t^{0\mu}$ is the momentum density. The formula is reminiscent of the way in which ordinary angular momentum is defined in point particle mechanics and is how we define the angular momentum of a field. We see once again that

$$\frac{dL^{\mu\nu}}{dt} \equiv 0, \quad (12.5.30)$$

if the fields fall off fast enough at infinity. We will have occasion to use these concepts in the next chapter when we work with the electromagnetic field and attempt to determine its properties. Let us now turn to the Hamiltonian description of field theories.

12.6 Hamiltonian description of fields

Our treatment of the Hamiltonian dynamics of fields may parallel the treatment of point particles. However, we need to be a bit more careful now than we were before, because the Lagrangian is really a functional and not a function as it was for point particles. Because we need to make a Legendre transformation on the Lagrangian to arrive at the

¹⁰Problem: Show this!

Hamiltonian, we will be required to define functional derivatives. We have actually already done this, but implicitly. Let us make this more explicit. Suppose that we are given a Lagrangian density function

$$\mathfrak{L} = \mathfrak{L}(\phi^A(x), \partial_\mu \phi^A(x), x), \quad (12.6.1)$$

and a Lagrangian functional

$$\mathcal{L}[\phi^A, \partial_\mu \phi^A, t] = \int d^3\vec{r} \mathfrak{L}(\phi^A(\vec{r}, t), \partial_\mu \phi^A(\vec{r}, t), \vec{r}, t). \quad (12.6.2)$$

We will define the n^{th} functional (or variational) derivative of \mathcal{L} w.r.t. $\phi^A(\vec{r}, t)$ by

$$\delta^n \mathcal{L}[\phi^A] = \int d^3\vec{r}_1 \dots d^3\vec{r}_n \frac{\delta^n \mathcal{L}[\phi^A]}{\delta \phi^{A_1}(\vec{r}_1, t) \dots \delta \phi^{A_n}(\vec{r}_n, t)} \delta \phi^{A_1}(\vec{r}_1, t) \dots \delta \phi^{A_n}(\vec{r}_n, t) \equiv \left. \frac{\partial^n \mathcal{L}[\phi^A + \alpha \eta^A]}{\partial \alpha^n} \right|_{\alpha=0} \quad (12.6.3)$$

This definition is purely formal. $\delta^n \mathcal{L}$ represents the n^{th} functional variation of the Lagrangian functional, each variation being performed precisely as we did for the action functional we have been working with till now. For instance we will recognize the first functional derivative from the definition above as

$$\delta \mathcal{L} = \int d^3\vec{r} \frac{\delta \mathcal{L}}{\delta \phi^A(\vec{r}, t)} \delta \phi^A(\vec{r}, t) \quad (12.6.4)$$

If, for $\phi^A(x) \in \mathbb{C}$ and all C^∞ variations $\delta \phi^A$, all the successive variations $\delta^n \mathcal{L}$ can be expressed in the above form up to some $k \in \mathbb{N}$, then the functional \mathcal{L} is said to be a C^k functional on \mathbb{C} .

Now define the momentum conjugate to the field (variable) $\phi(x)$ as

$$\pi_A(x) = \frac{\delta \mathcal{L}}{\delta \dot{\phi}^A(x)}, \quad (12.6.5)$$

and the Hamiltonian by a Legendre transformation as

$$\mathcal{H}[\pi_A, \phi^A, \partial_i \phi^A, t] = \int d^3\vec{r} \pi_A(\vec{r}, t) \dot{\phi}^A(\vec{r}, t) - \mathcal{L} = \int d^3x \mathfrak{H}(\pi_A, \phi^A, \partial_i \phi^A, t), \quad (12.6.6)$$

where \mathfrak{H} is the Hamiltonian density function. It is not difficult to see that \mathcal{H} is indeed independent of $\dot{\phi}^A, t$, because taking a functional derivative of \mathcal{H} w.r.t. $\dot{\phi}^A$ yields

$$\frac{\delta \mathcal{H}}{\delta \dot{\phi}^A(x)} = \pi_A(x) - \frac{\delta \mathcal{L}}{\delta \dot{\phi}^A(x)} \equiv 0 \quad (12.6.7)$$

As was the case for point particles, the equations of motion are recovered by varying the action

$$\mathcal{S} = \int_1^2 d^4x \left[\pi_A(x) \dot{\phi}^A(x) - \mathfrak{H}(\pi_A(x), \phi^A(x), \partial_i \phi^A(x), x) \right] \quad (12.6.8)$$

independently w.r.t. each of variables $(\pi_A(x), \phi^A(x))$, keeping *both* fixed at the boundaries “1” and “2”. One finds that

$$\dot{\phi}^A(x) = \frac{\delta \mathcal{H}}{\delta \pi_A(x)} = \frac{\partial \mathfrak{H}}{\partial \pi_A}, \quad \dot{\pi}_A(x) = -\frac{\delta \mathcal{H}}{\delta \phi^A(x)} = -\frac{\partial \mathfrak{H}}{\partial \phi^A} + \partial_i \left(\frac{\partial \mathfrak{H}}{\partial (\partial_i \phi^A)} \right) \quad (12.6.9)$$

For example, for the Lagrangian density (12.4.9) describing the elementary excitations of the rod, we have (suppressing the x dependence)

$$\pi = \mu \dot{\eta}, \quad \mathfrak{H} = \pi \dot{\eta} - \mathcal{L} = \frac{\pi^2}{2\mu} + Y \eta'^2 \quad (12.6.10)$$

giving the equations of motion

$$\dot{\eta} = \frac{\partial \mathfrak{H}}{\partial \pi} = \frac{\pi}{\mu}, \quad \dot{\pi} = -\frac{\partial \mathfrak{H}}{\partial \eta} + \left(\frac{\partial \mathfrak{H}}{\partial \eta'} \right)' = Y \eta'' \quad (12.6.11)$$

To obtain the Lagrangian equation of motion from the above canonical ones, use the first to replace π with $\mu \dot{\eta}$ in the second and get

$$\mu \ddot{\eta} - Y \eta'' = 0 \quad (12.6.12)$$

as we had before.

As in the case of point particles, the canonical equations can also be given in terms of Poisson brackets. Let $\mathcal{A}[\pi, \phi, t]$ and $\mathcal{B}[\pi, \phi, t]$ be two functionals of the phase space, $\{\phi, \pi\}$, then define the Poisson brackets between them as

$$\{\mathcal{A}, \mathcal{B}\}_{P.B.}^{t=t'} = \int d^3 \vec{r} \left[\frac{\delta \mathcal{A}}{\delta \phi^A(\vec{r}, t)} \frac{\delta \mathcal{B}}{\delta \pi_A(\vec{r}, t)} - \frac{\delta \mathcal{A}}{\delta \pi_A(\vec{r}, t)} \frac{\delta \mathcal{B}}{\delta \phi^A(\vec{r}, t)} \right] \quad (12.6.13)$$

If we compare the above expression with the definition we used for point particles, we will notice that only two things have changed, *viz.*, (a) there is an additional integration over space and (b) the ordinary derivatives have been replaced by functional derivatives. Both changes are dictated by the fact that we are now dealing with Lagrangian *functionals*. Had we more than one field, say ϕ_a , then a sum over all fields present (*i.e.*, \sum_a) would also have to be thrown into the definition, just as we sum over all the degrees of freedom for point particles. Note, finally, that the Poisson brackets are defined *at equal times*.

Consider the functional

$$\mathcal{F}[\phi^A] = \int d^3 \vec{r} \phi^A(\vec{r}, t) \quad (12.6.14)$$

then

$$\{\mathcal{F}, \mathcal{H}\}_{P.B.}^{t=t'} = \int d^3 \vec{r} \frac{\delta \mathcal{H}}{\delta \pi_A(\vec{r}, t)} = \int d^3 \vec{r} \dot{\phi}^A(\vec{r}, t) \quad (12.6.15)$$

and likewise,

$$\mathcal{F}[\pi] = \int d^3\vec{r} \pi_A(\vec{r}, t) \quad (12.6.16)$$

leads to

$$\{\mathcal{F}, \mathcal{H}\}_{P.B.}^{t=t'} = - \int d^3\vec{r} \frac{\delta \mathcal{H}}{\delta \phi^A(\vec{r}, t)} = \int d^3\vec{r} \dot{\pi}_A(\vec{r}, t) \quad (12.6.17)$$

From these equations it will be clear that we could also define the fundamental Poisson bracket relations

$$\begin{aligned} \{\phi^A(\vec{r}, t), \phi^B(\vec{r}', t')\}_{P.B.}^{t'=t} &= 0 = \{\pi_A(\vec{r}, t), \pi_B(\vec{r}', t')\}_{P.B.}^{t'=t} \\ \{\phi^A(\vec{r}, t), \pi_B(\vec{r}', t')\}_{P.B.}^{t'=t} &= \delta_B^A \delta^3(\vec{r} - \vec{r}') \end{aligned} \quad (12.6.18)$$

and write

$$\begin{aligned} \dot{\phi}^A(\vec{r}, t) &= \{\phi^A(\vec{r}, t), \mathcal{H}(t')\}_{P.B.}^{t'=t} \\ \dot{\pi}_B(\vec{r}, t) &= \{\pi_B(\vec{r}, t), \mathcal{H}(t')\}_{P.B.}^{t'=t} \end{aligned} \quad (12.6.19)$$

They would yield the canonical equations of (12.6.9).¹¹

¹¹Problem: Show this!

Chapter 13

The Electromagnetic Field and Sources*

13.1 Action for massive, free particles

The relativistic point particle extremizes its proper time (this can be thought of as a generalization of Fermat's principle, which was originally enunciated for the motion of light "corpuscles"),

$$\mathcal{S}_p = -mc^2 \int d\tau = -mc \int_1^2 \sqrt{-\eta_{\mu\nu} dx^\mu dx^\nu} = -mc^2 \int_1^2 dt \sqrt{1 - \frac{\vec{v}^2}{c^2}} \quad (13.1.1)$$

where $d\tau = ds/c = \frac{1}{c} \sqrt{-\eta_{\mu\nu} dx^\mu dx^\nu}$ and the constant " mc^2 " is chosen so that \mathcal{S}_p has the dimension of action (or angular momentum: J·s). One sees quite easily that this action principle reduces to Hamilton's principle ($V = 0$) when the velocity of the particle relative to the observer is small compared with the velocity light, for then

$$\sqrt{1 - \frac{\vec{v}^2}{c^2}} \approx 1 - \frac{1}{2} \frac{\vec{v}^2}{c^2} \quad (13.1.2)$$

which, when inserted into (13.1.1) gives

$$\mathcal{S}_p \approx \int_1^2 dt \left[\frac{1}{2} m \vec{v}^2 - mc^2 \right] \quad (13.1.3)$$

The second term is, of course, just a constant (later to be identified with the rest mass energy of the particle) and can be dropped without affecting either the equations of motion or the conservation laws. The first term is the non-relativistic kinetic energy of the particle and the action is therefore just that of a free non-relativistic point particle.

The momentum conjugate to x^i is

$$p_i = \frac{\partial \mathcal{L}}{\partial \dot{x}^i} = \frac{mv_i}{\sqrt{1 - \vec{v}^2/c^2}} = \gamma m v_i, \quad (13.1.4)$$

which reduces to $p_i = m v_i$ when $|\vec{v}| \ll c$, and Euler's equations give

$$\frac{d\vec{p}}{dt} = \frac{d}{dt}(\gamma m \vec{v}) = \frac{d}{dt} \frac{m\vec{v}}{\sqrt{1 - \vec{v}^2/c^2}} = 0 \quad (13.1.5)$$

which are the equations of motion of the particle. The Lagrangian does not depend explicitly on time, so we expect that the Hamiltonian is the total energy and is conserved,

$$E = \mathcal{H} = p_i \dot{x}^i - \mathcal{L} = \frac{m\vec{v}^2}{\sqrt{1 - \vec{v}^2/c^2}} + mc^2 \sqrt{1 - \vec{v}^2/c^2} = \frac{mc^2}{\sqrt{1 - \vec{v}^2/c^2}} \quad (13.1.6)$$

The quantity

$$m_R = \frac{m}{\sqrt{1 - \vec{v}^2/c^2}} \quad (13.1.7)$$

is generally called the “relativistic mass” or simply “mass” of the particle, whereas the parameter m we used initially is called the “rest” mass of the particle and can be thought of as its mass when measured in its proper frame ($\vec{v} = 0$). We have just obtained the popular Einstein relation,

$$E = m_R c^2. \quad (13.1.8)$$

Notice that the energy of the particle is not zero in the rest frame. Associated with the proper (rest) mass, is the proper energy, $E = mc^2$.

The Hamiltonian is rightly thought of as obtained by a Legendre transformation of the Lagrangian and should be expressed in terms of the momenta and coordinates but not the velocities. This is easily accomplished by noting that (13.1.4) gives

$$\vec{p}^2 = \frac{m^2 \vec{v}^2}{1 - \vec{v}^2/c^2} \rightarrow \frac{\vec{v}}{c} = \frac{\vec{p}}{\sqrt{\vec{p}^2 + m^2 c^2}} \quad (13.1.9)$$

Thus we get

$$1 - \frac{\vec{v}^2}{c^2} = \frac{\vec{p}^2}{\vec{p}^2 + m^2 c^2}, \quad (13.1.10)$$

which, when inserted into (13.1.6), gives another well known result,

$$\mathcal{H} = E = \sqrt{\vec{p}^2 c^2 + m^2 c^4}. \quad (13.1.11)$$

Again we recover the rest mass energy, $E = mc^2$ when we set $\vec{p} = 0$.

Let us note that the momentum

$$p_i = m\gamma v_i = m \frac{dt}{d\tau} \frac{dx_i}{dt} = m \frac{dx_i}{d\tau} \quad (13.1.12)$$

is quite manifestly the spatial component of the four-vector¹

$$p_\mu = m \frac{dx_\mu}{d\tau} = mU_\mu. \quad (13.1.13)$$

where we have defined the “four velocity vector” $U^\mu = dx^\mu/d\tau$. What is its time component? We have

$$p_0 = -mc^2 \frac{dt}{d\tau} = -\frac{mc^2}{\sqrt{1 - \vec{v}^2/c^2}} = -E \quad (13.1.14)$$

so we see that the spatial momentum and the energy are components of one four-vector momentum,

$$p^\mu = m \frac{dx^\mu}{d\tau}, \quad p^0 = \frac{E}{c^2}, \quad p^i = \frac{mv^i}{\sqrt{1 - \vec{v}^2/c^2}} \quad (13.1.15)$$

With this definition of the four-vector momentum, formula (13.1.11) for the energy is seen to result from a purely kinematic relation, because

$$p^2 = \eta_{\mu\nu} p^\mu p^\nu = m^2 \eta_{\mu\nu} \frac{dx^\mu}{d\tau} \frac{dx^\nu}{d\tau} = -m^2 \left[\frac{ds}{d\tau} \right]^2 = -m^2 c^2 \quad (13.1.16)$$

(the kinematic relation being, of course $U_\mu U^\mu = -c^2$, *remember it*). Therefore, expanding the l.h.s.,

$$p^2 = -\frac{E^2}{c^2} + \vec{p}^2 = -m^2 c^2, \quad \rightarrow \quad E^2 = \vec{p}^2 c^2 + m^2 c^4 \quad (13.1.17)$$

Interestingly, taking the square root allows for both positive and negative energies but we have chosen the positive sign, thereby excluding negative energy free particles by *fiat*.

Euler’s equations as given in (13.1.5) are not in a manifestly covariant form. They can, however, be put in such a form if we multiply by γ , expressing them as

$$\gamma \frac{d\vec{p}}{dt} = \frac{dt}{d\tau} \frac{d\vec{p}}{dt} = \frac{d\vec{p}}{d\tau} = 0 \quad (13.1.18)$$

This is the equation of motion for a *free* particle, so the r.h.s. is zero. The l.h.s. transforms as the spatial components of a four-vector and we need not worry about the transformation properties of the r.h.s., since it vanishes. In the presence of an external force the

¹Problem: Convince yourself that $p^\mu = m dx^\mu/d\tau$ is indeed a four-vector under Lorentz transformations. Remember that the proper time, τ , is a scalar. Determine the transformation properties of the four-vector momentum.

r.h.s. should not vanish and the principle of covariance requires that both sides of the equations of motion should transform in the same way under Lorentz transformations. Let us tentatively write a covariant equation of motion as

$$\frac{dp^\mu}{d\tau} = f^\mu \quad (13.1.19)$$

where f^μ is a four-vector. It must be interpreted as the relativistic equivalent of Newton's "force". But what is the connection between the four-vector force, f^μ , and our beloved Newtonian force, which we will call \vec{F}_N ? To find the relationship, consider the instantaneous rest-frame of the particle (quantities in the rest frame will be represented by an over-bar), in which $\tau = t$ and $\bar{p}^0 = m$, $\bar{p}^i = 0$. In this frame, the time component of the l.h.s of (13.1.19) is evidently zero and therefore so is \bar{f}^0 . The spatial part of the force equation in this (proper) frame reads

$$\frac{d\bar{p}^i}{dt} = \bar{f}^i, \quad (13.1.20)$$

which identifies the spatial components, \bar{f}^i , of \bar{f}^μ with the *Newtonian* force, \vec{F}_N , on the particle. Thus we have found that *in the instantaneous rest frame of the particle*

$$\bar{f}^\mu = (0, \vec{F}_N). \quad (13.1.21)$$

To determine f^μ in an arbitrary frame we merely need to perform a boost. Therefore, in a frame in which the instantaneous velocity of the particle is \vec{v} , we find in particular that

$$f^0 = -\gamma \frac{\vec{v} \cdot \vec{F}_N}{c^2} \quad (13.1.22)$$

i.e.,

$$\frac{dE}{dt} = -\vec{v} \cdot \vec{F}_N \quad (13.1.23)$$

The equation makes perfect sense, saying simply that the rate of energy gain (loss) of the particle is simply the power transferred (removed) to (from) the system by the external Newtonian forces. The same boost also gives the spatial components of the relativistic force in an arbitrary frame as

$$\vec{f} = \vec{F}_N + (\gamma - 1) \frac{\vec{v}}{v^2} (\vec{v} \cdot \vec{F}_N) \quad (13.1.24)$$

Notice that it has the non-relativistic limit ($\gamma \approx 1$) $\vec{f} \approx \vec{F}_N$, as it should.

We have given two forms of the action for the massive point particle in (13.1.1) although we have concentrated so far on the last of these. The first form is actually quite interesting. We had,

$$\mathcal{S}_p = \int_1^2 \sqrt{-\eta_{\mu\nu} dx^\mu dx^\nu}. \quad (13.1.25)$$

If λ is any parameter describing the particle trajectories then we could write this in the form

$$\mathcal{S}_p = \int_1^2 d\lambda \sqrt{-\eta_{\mu\nu} U_{(\lambda)}^\mu U_{(\lambda)}^\nu} \quad (13.1.26)$$

where

$$U^\mu(\lambda) = \frac{dx^\mu(\lambda)}{d\lambda} \quad (13.1.27)$$

is tangent to the trajectories $x^\mu(\lambda)$ and λ is an arbitrary parameter. The action is therefore reparameterization invariant, *i.e.*, it does not depend on the choice of parameter. All that we have said earlier corresponds to a particular choice of λ ($= t$) which, loosely speaking, is like fixing a “gauge”. After all, isn’t this essentially what we did with the electromagnetic field? We exploited the gauge freedom to fix the representative field configuration.

13.2 Action for the Electromagnetic field

The electromagnetic field equations of motion in the previous section are derivable from an action principle. Consider the following (gauge invariant) action,

$$\mathcal{S}_A = -\frac{1}{4\mu_o} \int d^4x F_{\mu\nu} F^{\mu\nu} \quad (13.2.1)$$

and ask what field configuration would extremize it. Varying w.r.t. the fields, A_μ , gives the Euler equations of motion,

$$\frac{\delta \mathcal{L}}{\delta A_\mu} - \partial_\nu \left[\frac{\delta \mathcal{L}}{\delta(\partial_\nu A_\mu)} \right] \equiv 0 \quad (13.2.2)$$

where $\mathcal{L}(A, \partial A)$ is the Lagrangian functional,

$$\mathcal{L}(A, \partial A) = -\frac{1}{4\mu_o} \int d^3\vec{r} F_{\mu\nu} F^{\mu\nu} \quad (13.2.3)$$

of the system. The Euler equations give precisely the dynamical equations

$$\partial_\mu F^{\mu\nu} = 0 \quad (13.2.4)$$

in the absence of sources.²

Let us compute the Hamiltonian of the electromagnetic field. Write the action as

$$\begin{aligned}
 \mathcal{S}_A &= -\frac{1}{4\mu_o} \int d^4x [F_{0i}F^{0i} + F_{ij}F^{ij}] \\
 &= -\frac{1}{4\mu_o} \int d^4x \left[-\frac{2}{c^2} \eta^{ij} F_{0i}F_{0j} + F_{ij}F^{ij} \right] \\
 &= -\frac{1}{4\mu_o} \int d^4x \left[-\frac{2}{c^2} \eta^{ij} (\partial_0 A_i - \partial_i A_0)(\partial_0 A_j - \partial_j A_0) + F_{ij}F^{ij} \right] \quad (13.2.5)
 \end{aligned}$$

First we notice that there are no time derivatives of the scalar potential, A_0 . Therefore the momentum conjugate to it is identically zero. The momenta conjugate to the space components of the electromagnetic field are

$$\pi^i = \frac{\partial \mathcal{L}}{\partial(\partial_0 A_i)} = -\frac{F^{0i}}{\mu_o} = \frac{F^{i0}}{\mu_o}, = -\frac{E^i}{\mu_o c^2} \quad (13.2.6)$$

so that the “velocities” are given in terms of the momenta as

$$\partial_0 A_i = \dot{A}_i = \mu_o c^2 \pi_i + \partial_i A_0 \quad (13.2.7)$$

It follows that the Hamiltonian density function may be written as

$$\begin{aligned}
 \mathfrak{H} = \pi^i \dot{A}_i - \mathcal{L} &= \pi^i (\mu_o c^2 \pi_i + \partial_i A_0) - \frac{1}{2} \mu_o c^2 \pi_i \pi^i + \frac{1}{4\mu_o} F_{ij} F^{ij} \\
 &= \frac{1}{2} \mu_o c^2 \pi^i \pi_i + \frac{1}{4\mu_o} F_{ij} F^{ij} + \pi^i \partial_i A_0 \quad (13.2.8)
 \end{aligned}$$

and the Hamiltonian functional as

$$\mathcal{H} = \int d^3\vec{r} \left[\frac{1}{2} \mu_o c^2 \pi^i \pi_i + \frac{1}{4\mu_o} F_{ij} F^{ij} - A_0 \partial_i \pi^i \right] \quad (13.2.9)$$

where we have performed an integration by parts in writing the last term and assumed that the fields fall off rapidly enough at infinity that the surface terms disappear. In the form above it becomes clear that the time-component of the electromagnetic field is a Lagrange multiplier field (this is why $\pi^0 \equiv 0$) while

$$\partial_i \pi^i = 0 = \vec{\nabla} \cdot \vec{E} \quad (13.2.10)$$

(which is just Gauss’ Law in the vacuum) is a constraint. Therefore, given that the gauge freedom allows us to choose one extra constraint in the form of a gauge condition, we learn that the electromagnetic field in the vacuum has two physical degrees of freedom.

²Prove that (13.2.2) actually gives the correct dynamical equations in the absence of sources ($j^\mu = 0$).

The remaining two terms in the Hamiltonian density will be recognized as the energy density of the electromagnetic field,

$$\frac{1}{2}\mu_0 c^2 \pi^i \pi_i + \frac{1}{4\mu_0} F_{ij} F^{ij} = \frac{1}{2} \left[\epsilon_0 \vec{E}^2 + \frac{\vec{B}^2}{\mu_0} \right]. \quad (13.2.11)$$

We may define the fundamental Poisson brackets by

$$\{A_i(t, \vec{r}), \pi^j(t, \vec{r}')\}_{P.B.} = \delta_i^j \delta^{(3)}(\vec{r} - \vec{r}') \quad (13.2.12)$$

(equal time!) and derive the canonical equations of motion from these fundamental relations. Thus,

$$\dot{A}_i(t, \vec{r}) = \{A_i(t, \vec{r}), \mathcal{H}\}_{P.B.} = \mu_0 c^2 \pi_i + \partial_i A_0 \quad (13.2.13)$$

which will be recognized as the expression we had earlier for the velocities in terms of the momenta, and

$$\dot{\pi}^i(t, \vec{r}) = \{\pi^i(t, \vec{r}), \mathcal{H}\}_{P.B.} = \frac{1}{\mu_0} \partial_j F^{ji} \quad (13.2.14)$$

which are precisely the Euler equations in (13.2.4)³

13.3 Inclusion of Sources

How shall we get a suitable coupling of the electromagnetic field to sources? The simplest Lorentz invariant integral we might construct which involves a 4-vector potential is the line integral,

$$e \int dx^\mu A_\mu. \quad (13.3.1)$$

where we have thrown in an electric charge which is a scalar and makes no difference to the Lorentz invariance of the expression. e is the electric charge. It is certainly a Lorentz scalar, but is it gauge invariant? Consider the gauge transformation $A_\mu \rightarrow A'_\mu = A_\mu + \partial_\mu \Lambda$. Our line integral goes to

$$\begin{aligned} \mathcal{S}_{\text{int}} &= e \int_1^2 dx^\mu A_\mu \rightarrow e \int_1^2 dx^\mu A'_\mu = e \int_1^2 dx^\mu (A_\mu + \partial_\mu \Lambda) \\ &= e \int_1^2 dx^\mu A_\mu + e(\Lambda_2 - \Lambda_1) \end{aligned} \quad (13.3.2)$$

Thus, under a gauge transformation, our line integral has changed by a constant. However, an action is defined only up to the addition of a constant, because the addition of a constant

³Check this!

makes absolutely no difference to the equations of motion. Therefore, although our line integral is not gauge invariant its change by a gauge transformation is only by a constant and we expect that the equations of motion that would follow from it will stay gauge invariant. Rewrite the integral as

$$\begin{aligned}\mathcal{S}_{\text{int}} &= e \int_1^2 dx^\mu A_\mu = e \int_1^2 dt v^\mu(t) A_\mu(t, \mathbf{r}(t)) \\ &= e \int d^4x v^\mu(t) A_\mu(\vec{r}, t) \delta^{(3)}(\vec{r} - \mathbf{r}(t))\end{aligned}\quad (13.3.3)$$

where $v^\mu = dx^\mu/dt$ is the velocity (*not* the 4-velocity vector!) and $\mathbf{r}(t)$ is the particle's trajectory. In the last rewrite we have replaced the integral over time by an integral over space-time and introduced a δ -function to eliminate the extra integral over the spatial volume. This, of course, is just a trick in order to be able to easily combine the two integrals (13.2.1) and (13.3.1) into one. Consider then the modified action

$$\mathcal{S}' = \int d^4x \left[-\frac{1}{4\mu_o} F_{\mu\nu} F^{\mu\nu} + ev^\mu A_\mu \delta^{(3)}(\vec{r} - \mathbf{r}(t)) \right] \quad (13.3.4)$$

and extremize it w.r.t. variations of A_μ . We easily find

$$\partial_\nu F^{\nu\mu} = -e\mu_o v^\mu \delta^{(3)}(\vec{r} - \mathbf{r}(t)) \quad (13.3.5)$$

which identifies the current 4-vector as $j^\mu = e\mu_o v^\mu \delta^{(3)}(\vec{r} - \mathbf{r}(t))$.⁴ Thus, for the spatial components we have

$$j^i = ev^i \delta^{(3)}(\vec{r} - \mathbf{r}(t)) \quad (13.3.6)$$

and for the time component ($v^0 = 1$)

$$\rho = e\delta^{(3)}(\vec{r} - \mathbf{r}(t)), \quad (13.3.7)$$

precisely as we have used in the past for the current corresponding to a single point particle! Indeed it will also yield the Lorentz force law via the Euler equations for the particle, as we shortly see. The line integral must be thought of as an “interaction” action for a charged particle (charge e) interacting with the electromagnetic field.

To complete the action for the charged particle and the electromagnetic field, we simply add to (13.3.5) the action for the free particle. The complete action, \mathcal{S} , will then be $\mathcal{S} = \mathcal{S}_A + \mathcal{S}_{\text{int}} + \mathcal{S}_p$, *i.e.*,

$$\mathcal{S} = \int d^4x \left[-\frac{1}{4\mu_o} F_{\mu\nu} F^{\mu\nu} + (ev^\mu A_\mu - mc\sqrt{-\eta_{\mu\nu}v^\mu v^\nu}) \delta^{(3)}(\vec{r} - \mathbf{r}(t)) \right] \quad (13.3.8)$$

⁴For j^μ to be a satisfactory current density, it must be conserved, *i.e.*, $\partial_\mu j^\mu = 0$. (a) Satisfy yourself that this is indeed a necessary condition. (b) Prove that the current defined in this way is indeed conserved. If you cannot do so, see Chapter 10.

Because the last term does not contain the electromagnetic field, it is transparent to variations of the field and so does not affect Maxwell's dynamical equation. However, to describe the particle's motion we must vary w.r.t. the particle's position and for this it is necessary to consider the second two terms of (13.3.8)

$$\mathcal{S}'' = \int_1^2 dt [ev^\mu A_\mu - mc\sqrt{-\eta_{\mu\nu}v^\mu v^\nu}] = \int_1^2 d\tau [eU^\mu A_\mu - mc\sqrt{-\eta_{\mu\nu}U^\mu U^\nu}] \quad (13.3.9)$$

We find

$$\frac{\partial \mathcal{L}''}{\partial x^\mu} - \frac{d}{d\tau} \left(\frac{\partial \mathcal{L}''}{\partial U^\mu} \right) = 0 = eU^\nu \partial_\mu A_\nu - \frac{d}{d\tau} [eA_\mu + mU_\mu] \quad (13.3.10)$$

where we have used $U_\mu U^\mu = -c^2$. But we can write

$$\frac{dA_\mu}{d\tau} = \frac{dx^\nu}{d\tau} \partial_\nu A_\mu = U^\nu \partial_\nu A_\mu \quad (13.3.11)$$

so that the above equation of motion for the particle becomes

$$m \frac{dU_\mu}{d\tau} = e(\partial_\mu A_\nu - \partial_\nu A_\mu)U^\nu = eF_{\mu\nu}U^\nu \quad (13.3.12)$$

We see that, because of the interaction Lagrangian, the electromagnetic field is no longer source free and the particle is no longer free but subject to an external force. What is this force? Consider a non-relativistic particle, so that $U^0 \approx 1$ and $U^i \approx v^i$. Then ($\tau \approx t$)

$$m \frac{dv_i}{dt} \approx e[F_{i0} + F_{ij}v^j] = e[E_i + \epsilon_{ijk}v^j B^k] = eE_i + e(\vec{v} \times \vec{B})_i \quad (13.3.13)$$

which must be recognized as precisely the Lorentz force on the particle. Equation (13.3.12) is known as the Minkowski equation and is the correct generalization of the Newton-Lorentz force equation of non-relativistic electrodynamics. The particular line integral that gave us the interaction action is said to correspond to a *minimal* coupling between the electromagnetic field and the particle.

The generalization to a system of non-interacting particles is trivial. Simply sum over the particles in the free action and the interaction action, *i.e.*, consider the action

$$\mathcal{S} = \int d^4x \left[-\frac{1}{4\mu_0} F_{\mu\nu} F^{\mu\nu} + \sum_{n=1}^N \left(e_{(n)} v_{(n)}^\mu A_\mu - m_{(n)} c \sqrt{-\eta_{\mu\nu} v_{(n)}^\mu v_{(n)}^\nu} \right) \delta^{(3)}(\vec{r} - \mathbf{r}_{(n)}(t)) \right] \quad (13.3.14)$$

It gives the following expression for the current that couples to the electromagnetic field

$$\partial_\nu F^{\nu\mu} = -j^\mu, \quad j^\mu(\vec{r}, t) = \sum_{(n)} \mu_0 e_n v_{(n)}^\mu(t) \delta^{(3)}(\vec{r} - \mathbf{r}_{(n)}(t)) \quad (13.3.15)$$

and for each particle it gives the expected relativistic equation of motion

$$m_{(n)} \frac{dU_{\mu}^{(n)}}{d\tau} = e_{(n)} F_{\mu\nu} U_{(n)}^{\nu}. \quad (13.3.16)$$

Having an action that clearly is correct as far as the equations of motion go, let us compute the Hamiltonian for the particle in an electromagnetic field and then the total Hamiltonian of the system. We have already done so for the electromagnetic field in the absence of sources and separately for the free particle. The only addition then is the interaction term, but this will be seen to change things quite dramatically. Let us begin with the particle action (including the interaction term),

$$\begin{aligned} \mathcal{S}'' &= \int_1^2 dt [-mc\sqrt{-\eta_{\mu\nu}v^{\mu}v^{\nu}} + eA_{\mu}v^{\mu}] \\ &= \int_1^2 dt \left[-mc^2\sqrt{1 - \frac{\vec{v}^2}{c^2}} + e\vec{v} \cdot \vec{A} + eA_0 \right] \end{aligned} \quad (13.3.17)$$

and compute the particle conjugate momenta,

$$\vec{p} = \frac{m\vec{v}}{\sqrt{1 - \frac{v^2}{c^2}}} + e\vec{A} \quad (13.3.18)$$

This expression is easily inverted, as before, to yield the velocities in terms of the conjugate momenta

$$\frac{\vec{v}}{c} = \frac{(\vec{p} - e\vec{A})}{\sqrt{(\vec{p} - e\vec{A})^2 + m^2c^2}} \quad (13.3.19)$$

Performing a Legendre transformation then yields the Hamiltonian of the particle in the presence of the electromagnetic field,

$$\mathcal{H} = p_i \dot{x}^i - \mathcal{L} = \sqrt{(\vec{p} - e\vec{A})^2 c^2 + m^2 c^4} - eA_0. \quad (13.3.20)$$

The square-root has the same structure as the Hamiltonian for the free particle: the difference is that the momentum has been changed

$$\vec{p} \rightarrow \vec{p} - e\vec{A} \quad (13.3.21)$$

by the presence of the electromagnetic field. What are the canonical equations of motion for the particle? We evaluate the Poisson brackets,

$$\vec{v} = \dot{\vec{r}} = \{\vec{r}, \mathcal{H}\}_{P.B.} = \frac{(\vec{p} - e\vec{A})c}{\sqrt{(\vec{p} - e\vec{A})^2 + m^2c^2}} \quad (13.3.22)$$

which is precisely the expression already obtained above, and

$$\dot{p}_i = \{p_i, \mathcal{H}\}_{P.B.} = \frac{ec(p^j - eA^j)\partial_i A_j}{\sqrt{(\vec{p} - e\vec{A})^2 + m^2c^2}} + e\partial_i A_0 = ev^j\partial_i A_j + e\partial_i A_0 \quad (13.3.23)$$

where, in the last expression, we have used the relation between the velocity and the momenta. This does not *look* the same as the Minkowski equation (13.3.12), but indeed it is. In showing that the equations are the same one should note that the derivative in \vec{p} is w.r.t. coordinate time whereas, in (13.3.12), the momentum derivative is w.r.t. proper time. Also, the relationship between the momenta and velocities in (13.3.18) involves the electromagnetic field. Write

$$\frac{dp_i}{dt} = \frac{d\tau}{dt} \frac{dp_i}{d\tau} = \frac{1}{\gamma} \frac{d}{d\tau} (mU_i + eA_i), \quad (13.3.24)$$

where we have used (13.3.18), then the canonical equations may be written as

$$m \frac{dU_i}{d\tau} = eU^j \partial_i A_j + eU^0 \partial_i A_0 - e \frac{dA_i}{d\tau} = e(\partial_i A_\nu - \partial_\nu A_i)U^\nu = eF_{i\nu}U^\nu \quad (13.3.25)$$

which is the spatial part of (13.3.12)

To get the total Hamiltonian for the system, simply add the Hamiltonian for the electromagnetic field itself (in the absence of sources, since we have already taken the source action with the interaction term into account). We find

$$\mathcal{H} = \frac{1}{2} \int d^3\vec{r} \left[\epsilon_o \vec{E}^2 + \frac{\vec{B}^2}{\mu_o} \right] + \sqrt{(\vec{p} - e\vec{A})^2 c^2 + m^2 c^4} - eA_0. \quad (13.3.26)$$

These considerations are readily generalized to a system with many charged particles.

We must now turn to the symmetries of our Lagrangian and what these symmetries tell us about the conserved quantities.

13.4 Conservation Laws

There are three conservation laws, *viz.*, conservation of (a) energy and momentum (b) orbital angular momentum and (c) intrinsic angular momentum, which we will examine in this section

13.4.1 Energy and Momentum

Let's consider the effects of a global translation, $x^\mu \rightarrow x^\mu + \epsilon^\mu$, (ϵ^μ const.) such that $\delta A_\mu = 0$. Invariance of the action under this transformation leads to the (conserved)

electromagnetic stress tensor:

$$j_\nu^\mu = \mathfrak{L}\delta_\nu^\mu - \frac{\partial \mathfrak{L}}{\partial(\partial_\mu A_\alpha)} \partial_\nu A_\alpha \quad (13.4.1)$$

where

$$\mathfrak{L} = -\frac{1}{4\mu_o} F_{\alpha\beta} F^{\alpha\beta} \quad (13.4.2)$$

A simple calculation gives

$$\mu_o j_\nu^\mu = F^{\mu\alpha} \partial_\nu A_\alpha - \frac{1}{4} \delta_\nu^\mu F^2, \quad \mu_o j^{\mu\nu} = F^{\mu\alpha} \partial^\nu A_\alpha - \frac{1}{4} \eta^{\mu\nu} F^2 \quad (13.4.3)$$

where $F^2 = F_{\alpha\beta} F^{\alpha\beta}$. While the second term is symmetric in its indices, the first term is not, so we must try to find a divergence free tensor that, when added to the expression on the r.h.s. of the above equation, gives a symmetric tensor. In the previous chapter, the prescribed way was to add the divergence of a third rank tensor, *i.e.*, $\Delta^{\mu\nu} = \partial_\lambda k^{\lambda\mu\nu}$, where k is antisymmetric in (λ, μ) , to the expression obtained directly from Noether's theorem.

Consider then the tensor

$$\Delta^{\mu\nu} = -\frac{1}{\mu_o} \partial_\alpha (F^{\alpha\mu} A^\nu) \quad (13.4.4)$$

Thus we have chosen $k^{\lambda\mu\nu} = F^{\lambda\mu} A^\nu$, which is antisymmetric in (λ, μ) so $\partial_\mu \Delta^{\mu\nu} \equiv 0$ by construction. Moreover,

$$\partial_\alpha (F^{\alpha\mu} A^\nu) = (\partial_\alpha F^{\alpha\mu}) A^\nu + F^{\alpha\mu} (\partial_\alpha A^\nu) = -j^\mu A^\nu + F^{\alpha\mu} (\partial_\alpha A^\nu) \quad (13.4.5)$$

by Maxwell's equations. The first term vanishes at all points where there are no sources, *i.e.*, at all points at which $j^\mu = 0$. At such points, adding $\Delta^{\mu\nu}$ to $j^{\mu\nu}$ yields a symmetric tensor. We may therefore define the vacuum *electromagnetic* stress tensor by

$$\begin{aligned} t^{\mu\nu} &= j^{\mu\nu} + \Delta^{\mu\nu} = \frac{1}{\mu_o} \left[F^{\mu\alpha} (\partial^\nu A_\alpha - \partial_\alpha A^\nu) - \frac{1}{4} \eta^{\mu\nu} F^2 \right] \\ &= \frac{1}{\mu_o} \left[-F^{\mu\alpha} F_\alpha{}^\nu - \frac{1}{4} \eta^{\mu\nu} F^2 \right] \end{aligned} \quad (13.4.6)$$

It is both symmetric and conserved at points where there are no sources. With it, we can compute the linear and angular momenta of the electromagnetic field. The linear momentum is simply

$$P^\mu = \int d^3\vec{r} \, t^{\mu 0} \quad (13.4.7)$$

so that the momentum density, $\wp^\mu = t^{\mu 0}$, has components,

$$\wp^0 = \frac{1}{\mu_o} \left[F^{0j} F_j{}^0 + \frac{1}{4c^2} F^2 \right] = \frac{1}{2\mu_o c^2} \left[\frac{\vec{E}^2}{c^2} + \vec{B}^2 \right] = \frac{1}{2c^2} \left[\epsilon_o \vec{E}^2 + \frac{\vec{B}^2}{\mu_o} \right] = \frac{\mathcal{E}}{c^2} \quad (13.4.8)$$

(where \mathcal{E} , as we expect, this is the Hamiltonian (energy) density of the electromagnetic field) and

$$\wp^i = \frac{1}{\mu_o} F^{ij} F_j^0 = -\frac{1}{\mu_o} \epsilon^{ijk} B_k \frac{E_j}{c^2} = \frac{1}{\mu_o c^2} (\vec{E} \times \vec{B})^i, \quad (13.4.9)$$

which we will recognize as the famous ‘‘Poynting vector’’ whose integral on a surface does, in fact, represent the rate of electromagnetic energy flow across that surface. Note the fact that when \wp^μ is evaluated for the plane wave of Chapter I we get

$$\vec{\wp} = \pm \frac{\mathcal{E}}{c} \hat{k} \quad (13.4.10)$$

where \hat{k} is the unit vector in the direction of propagation. In words, the plane wave satisfies a dispersion relation that is typical of a massless particle!

13.4.2 Orbital Angular Momentum

With our (symmetric) $t^{\mu\nu}$, we may proceed to construct the modified angular momentum tensor density of the electromagnetic field in terms of the conserved tensor density

$$\tilde{L}^{\lambda\mu\nu} = x^\nu t^{\lambda\mu} - x^\mu t^{\lambda\nu}. \quad (13.4.11)$$

As we learned in the previous chapter, the angular momentum density is

$$\tilde{L}^{0\mu\nu} = \tilde{L}^{0\mu\nu} = x^\nu \wp^\mu - x^\mu \wp^\nu \quad (13.4.12)$$

and the *total* angular momentum

$$L^{\mu\nu} = \int d^3\vec{r} \tilde{L}^{0\mu\nu} \quad (13.4.13)$$

is conserved. The angular momentum tensor is antisymmetric in its indices and so it has six independent components. Of these, the three spatial components, l^{ij} , are of particular interest to us at this point. Define the vectors

$$l_i = -\frac{1}{2} \epsilon_{ijk} l^{jk} \quad (13.4.14)$$

so that $\vec{l} = (\vec{r} \times \vec{\wp})$, which parallels the definition of the angular momentum in classical particle mechanics, except that we must bear in mind that we are dealing here with fields and that l^{ij} is really a *density*. $\vec{\wp}$ is the Poynting vector, therefore we find

$$\vec{l} = \frac{1}{\mu_o c^2} \vec{r} \times (\vec{E} \times \vec{B}), \quad (13.4.15)$$

and

$$\vec{L} = \frac{1}{\mu_0 c^2} \int d^3 \vec{r} \left[\vec{r} \times (\vec{E} \times \vec{B}) \right] \quad (13.4.16)$$

The remaining three components of the angular momentum density of the electromagnetic field are the time-space components, l^{0i} . These three components are necessary to have a covariant generalization of angular momentum. Their conservation is a statement on the center of mass motion.

13.4.3 Intrinsic Angular Momentum

Turn next to the modified intrinsic angular momentum in (12.5.28). To compute $S_{\alpha\beta}^\mu$ we require the transformation of the electromagnetic field under Lorentz transformations,

$$A^\mu(x) \rightarrow A'^\mu(x') = \frac{\partial x'^\mu}{\partial x^\lambda} A^\lambda(x) \Rightarrow \delta A^\mu = \frac{1}{2} (\delta_\alpha^\mu \eta_{\beta\lambda} - \delta_\beta^\mu \eta_{\alpha\lambda}) A^\lambda \delta \omega^{\alpha\beta} \quad (13.4.17)$$

so

$$G_{\alpha\beta}^\mu = \frac{1}{2} (\delta_\alpha^\mu \eta_{\beta\lambda} - \delta_\beta^\mu \eta_{\alpha\lambda}) A^\lambda \quad (13.4.18)$$

and

$$S_{\alpha\beta}^\mu = -\frac{1}{2\mu_0} (F^\mu{}_\alpha A_\beta - F^\mu{}_\beta A_\alpha) = -\frac{1}{2\mu_0} F^\mu{}_{[\alpha} A_{\beta]} \quad (13.4.19)$$

and from this we should subtract $\frac{1}{2} \Delta_{[\alpha}^\mu x_{\beta]}$, where the brackets indicate the required antisymmetrization. The result can be given as

$$\tilde{S}_{\alpha\beta}^\mu = -\frac{1}{\mu_0} F^\mu{}_{[\alpha} A_{\beta]} + \frac{1}{2\mu_0} \partial_\sigma (F^{\sigma\mu} A_{[\alpha} x_{\beta]}) \quad (13.4.20)$$

It is easy to see that \tilde{S} is not gauge invariant. However, one can define a gauge invariant spin density vector by simply restricting \vec{A} above to its solenoidal part (using the Helmholtz decomposition⁵).

⁵The Helmholtz decomposition of a vector is the statement that *any* vector, \vec{F} in three dimensions can be decomposed into a solenoidal component ($\vec{\nabla} \cdot \vec{F} = 0$) and an irrotational component ($\vec{\nabla} \times \vec{F} = 0$) as follows:

$$\begin{aligned} \vec{F} &= -\vec{\nabla} \varphi(\vec{r}) + \vec{\nabla} \times \vec{A}(\vec{r}) = \vec{F}_{\text{irr}} + \vec{F}_{\text{sol}} \\ \varphi &= \int_V d^3 \vec{r}' \frac{\vec{\nabla}' \cdot \vec{F}(\vec{r}')}{4\pi |\vec{r} - \vec{r}'|} - \oint_S \frac{\vec{F}(\vec{r}') \cdot d\vec{S}'}{4\pi |\vec{r} - \vec{r}'|} \\ \vec{A} &= \int_V d^3 \vec{r}' \frac{\vec{\nabla}' \times \vec{F}(\vec{r}')}{4\pi |\vec{r} - \vec{r}'|} + \oint_S \frac{\vec{F}(\vec{r}') \times d\vec{S}'}{4\pi |\vec{r} - \vec{r}'|} \end{aligned}$$

where S bounds V .

13.5 Energy and Momentum with sources

Finally, let us see what the divergence of $t^{\mu\nu}$ yields when sources are present and what this means for energy and momentum conservation (we have, after all, called $t^{\mu\nu}$ a stress-energy or energy momentum tensor, so its divergence law should tell us something about this!). Taking the divergence of $t^{\mu\nu}$ we find

$$\begin{aligned}\partial_\mu t^{\mu\nu} &= \frac{1}{\mu_o} \left[-\partial_\mu F^{\mu\alpha} F_\alpha{}^\nu - F^{\mu\alpha} \partial_\mu F_\alpha{}^\nu - \frac{1}{2} \eta^{\mu\nu} F_{\alpha\beta} \partial_\mu F^{\alpha\beta} \right] \\ &= \frac{1}{\mu_o} \left[j^\alpha F_\alpha{}^\nu - F^{\mu\alpha} \partial_\mu F_\alpha{}^\nu - \frac{1}{2} \eta^{\mu\nu} F_{\alpha\beta} \partial_\mu F^{\alpha\beta} \right]\end{aligned}\quad (13.5.1)$$

so that, taking the first term on the right hand side to the left and re-grouping the remaining terms on the right, we have

$$\partial_\mu t^{\mu\nu} - \frac{1}{\mu_o} j^\alpha F_\alpha{}^\nu = -\frac{1}{\mu_o} \left[F_{\alpha\beta} \left(\partial^\alpha F^{\beta\nu} + \frac{1}{2} \partial^\nu F^{\alpha\beta} \right) \right]\quad (13.5.2)$$

Now, because $F^{\mu\nu}$ is antisymmetric in its indices, the following

$$\partial^\nu F^{\alpha\beta} + \partial^\beta F^{\nu\alpha} + \partial^\alpha F^{\beta\nu} \equiv 0\quad (13.5.3)$$

is an *algebraic* identity that can be verified simply by expansion (it is sometimes called the homogeneous Maxwell equations). It allows us to write the right hand side of 13.5.2 as

$$\begin{aligned}& -\frac{1}{\mu_o} F_{\alpha\beta} \left(\partial^\alpha F^{\beta\nu} - \frac{1}{2} (\partial^\beta F^{\nu\alpha} + \partial^\alpha F^{\beta\nu}) \right) \\ &= -\frac{1}{\mu_o} F_{\alpha\beta} \left(\partial^\alpha F^{\beta\nu} + \partial^\beta F^{\alpha\nu} \right)\end{aligned}\quad (13.5.4)$$

where, in the last equation above we have written

$$F_{\alpha\beta} \partial^\beta F^{\nu\alpha} = -F_{\alpha\beta} \partial^\beta F^{\alpha\nu}\quad (13.5.5)$$

and

$$F_{\alpha\beta} \partial^\alpha F^{\beta\nu} = F_{\beta\alpha} \partial^\beta F^{\alpha\nu} = -F_{\alpha\beta} \partial^\beta F^{\alpha\nu}\quad (13.5.6)$$

to simplify the terms. We see that the right hand side must be identically zero because it is the product of an antisymmetric tensor ($F_{\alpha\beta}$) and a symmetric one. Therefore,

$$\partial_\mu t^{\mu\nu} = \frac{1}{\mu_o} j^\alpha F_\alpha{}^\nu\quad (13.5.7)$$

(it is conserved in the absence of sources⁶). In the presence of sources, therefore, we get, for the time-component ($\nu = 0$),

$$\partial_t t^{00} + \partial_i t^{i0} = \frac{1}{\mu_o} j^i F_i^0 \quad (13.5.8)$$

or

$$\frac{\partial \wp^0}{\partial t} + \vec{\nabla} \cdot \vec{\wp} = -\frac{1}{\mu_o c^2} \vec{j} \cdot \vec{E} \quad (13.5.9)$$

or

$$\frac{\partial \mathcal{E}}{\partial t} + \vec{\nabla} \cdot \vec{S} = -\frac{1}{\mu_o} \vec{j} \cdot \vec{E} \quad (13.5.10)$$

where we have defined the Poynting vector, \vec{S} , as $\vec{S} = \frac{1}{\mu_o} (\vec{E} \times \vec{B})$. If the right hand side vanishes, this is a continuity equation and expresses the conservation of energy momentum. If it does not vanish, then it represents the rate at which energy is transferred either to or from the electromagnetic field by the charge currents present in the system.

For the space components ($\nu = i$), we have

$$\partial_t t^{0i} + \partial_k t^{ki} = \frac{1}{\mu_o} (j^0 F_0^i + j^k F_k^i) = -\left[\rho \vec{E} + \frac{1}{\mu_o} \vec{j} \times \vec{B} \right]^i \quad (13.5.11)$$

or

$$\frac{\partial \wp^i}{\partial t} - \partial_k t^{ki} = \left[\rho \vec{E} + \frac{1}{\mu_o} \vec{j} \times \vec{B} \right]^i \quad (13.5.12)$$

which represents the conservation of stress. The four vector on the right hand side of equation 13.5.7, *i.e.*,

$$f^\mu = -\frac{1}{\mu_o} j^\alpha F_\alpha^\nu = \left(\frac{1}{\mu_o} \vec{j} \cdot \vec{E}, \rho \vec{E} + \frac{1}{\mu_o} \vec{j} \times \vec{B} \right) \quad (13.5.13)$$

is called the ‘‘Lorentz force density’’.

The fact that the electromagnetic stress energy tensor is not divergence free in the presence of sources means that we’re missing something: after all, we know that energy and momentum must be *globally* conserved. This means that the stress energy tensor of the sources must compensate in some way, leading to a conserved *total* stress energy tensor. The most obvious candidate for the stress energy tensor of a single particle is just

$$\tau^{\mu\nu} = \frac{m}{E} U^\mu U^\nu \delta^{(3)}(\vec{r} - \mathbf{r}(t)) = p^\nu \frac{dx^\mu}{dt} \delta^{(3)}(\vec{r} - \mathbf{r}(t)) = \frac{1}{e} j^\mu p^\nu \quad (13.5.14)$$

⁶Of course we do not expect it to be conserved in the presence of sources because, in deriving the stress tensor, we used only the Lagrangian density of the electromagnetic field.

where E is the particle energy and j^μ is the source current. It is symmetric in its indices, as is apparent from the first expression on the right, although this is not obvious in the last expression. Its generalization to many particles should be obvious by now.⁷ Take the divergence of this tensor, remembering that the current, j^μ , is conserved. This means that the only contribution to the divergence comes from a derivative of p^ν . However, p^ν involves just time – it is not a function of space – so $\partial_\mu p^\nu = dp^\nu/dt$. Then use (13.3.12) to find

$$\begin{aligned}\partial_\mu \tau^{\mu\nu} &= \frac{1}{e} \frac{dp^\nu}{dt} j^0 = \frac{dp^\nu}{d\tau} \frac{d\tau}{dt} \delta^{(3)}(\vec{r} - \mathbf{r}(t)) \\ &= e F^{\nu\alpha} U_\alpha \frac{d\tau}{dt} \delta^{(3)}(\vec{r} - \mathbf{r}(t)) = F^{\nu\alpha} v_\alpha \delta^{(3)}(\vec{r} - \mathbf{r}(t)) \\ \rightarrow \partial_\mu \tau^{\mu\nu} &= F^{\nu\alpha} j_\alpha = -j^\alpha F_\alpha{}^\nu.\end{aligned}\tag{13.5.15}$$

Thus

$$\Theta^{\mu\nu} = t^{\mu\nu} + \tau^{\mu\nu}\tag{13.5.16}$$

is indeed conserved. It is the *full* energy momentum tensor we have talked about and represents the total energy and momentum of the field plus matter. Its conservation declares that the total energy-momentum is neither created nor destroyed.

⁷Yes, construct it and go through the discussion that follows for a single particle! Make sure it works for the multiparticle system too!

Chapter 14

The Homogeneous Wave Equation

14.1 Isotropic, linear, non-dispersive media

We will now make two somewhat restrictive assumptions, *viz.*,

- the scales of interest are large compared to the interatomic distances (a typical electric or magnetic dipole length). Thus we shall use the macroscopic Maxwell's equations as given in (8.1.1).
- The media that enter into our problems are all isotropic, linear and non-dispersive.

Recall that in most dielectrics the electric polarization vector is given by the phenomenological relation¹

$$\vec{P} = \chi_e(E)\vec{E} \quad (14.1.1)$$

where the quantity $\chi_e(E)$ is called the electric susceptibility of the material. If we write $\vec{D} = \epsilon(E)\vec{E}$, then from the definition of \vec{D} we find $\epsilon(E) = \epsilon_o + \chi_e(E)$ and $\epsilon(E)$ is called the permittivity of the material. Frequently, and particularly for weak electric fields, it is found that the permittivity (or susceptibility) is *independent* of the electric field. Such media are said to be linear dielectrics. The notion of a linear magnetic material is the same. In a large class of materials, if they are isotropic, one has the phenomenological relation²

$$\vec{M} = \chi_m(H)\vec{H} \quad (14.1.2)$$

¹If a dielectric is anisotropic then we replace (14.1.1) by $P_i = \chi_{ij}^e(E)E_j$ (sum over repeated indices!) and \vec{P} does not necessarily have the same direction as \vec{E} . χ_{ij}^e is the electric susceptibility *tensor*. It is a tensor w.r.t. spatial rotations.

²Again, if a medium is anisotropic then we replace (14.1.2) by $M_i = \chi_{ij}^m(H)H_j$ and \vec{M} does not necessarily have the same direction as \vec{H} . χ_{ij}^m is the magnetic susceptibility tensor.

where $\chi_m(H)$ is called the magnetic susceptibility of the medium. Again, if we write $\vec{B} = \mu(H)\vec{H}$ then as $\mu(H) = \mu_o(1 + \chi_m(H))$. The quantity μ is called the permeability of the magnetic medium ($K_m = \mu/\mu_o = 1 + \chi_m$ is called the relative permeability). If the susceptibility (or permeability) does not depend on \vec{H} then the medium is said to be linear. In dispersive media such simple relationships between the polarization vector and the electric field or the magnetization vector and the magnetic field are not possible: the permittivity and permeability may also depend on the frequency of the electric and magnetic fields.

Thus, for the present, $\vec{B} = \mu\vec{H}$ and $\vec{D} = \epsilon\vec{E}$ and both ϵ and μ are constants. In this case, defining

$$\vec{B} = \vec{\nabla} \times \vec{A}, \quad \vec{E} = -\vec{\nabla}\phi - \frac{\partial\phi}{\partial t} \quad (14.1.3)$$

as before, our dynamical equations turn into

$$\begin{aligned} -\vec{\nabla}^2\phi - \frac{\partial}{\partial t}\vec{\nabla} \cdot \vec{A} &= \frac{\rho_f}{\epsilon} \\ \frac{1}{v^2}\frac{\partial^2\vec{A}}{\partial t^2} - \vec{\nabla}^2\vec{A} + \vec{\nabla}\left(\vec{\nabla} \cdot \vec{A} + \frac{1}{v^2}\frac{\partial\phi}{\partial t}\right) &= \mu\vec{j}_f \end{aligned} \quad (14.1.4)$$

where $v = 1/\sqrt{\mu\epsilon}$ is the speed of light in the medium. We notice that all our previous work is appropriate for and can be translated to the case of isotropic, linear and non-dispersive media by simply changing (a) the speed of light ($c \rightarrow v$), (b) the permittivity and permeability ($\epsilon_o \rightarrow \epsilon$ and $\mu_o \rightarrow \mu$)³ and (c) replacing the sources by the transport charge and current densities. (In terms of the electric and magnetic fields, the dynamical equations read

$$\begin{aligned} \vec{\nabla} \cdot \vec{E} &= \frac{\rho_f}{\epsilon} \\ \vec{\nabla} \times \vec{B} - \frac{1}{v^2}\frac{\partial\vec{E}}{\partial t} &= \mu\vec{j}_f \end{aligned} \quad (14.1.5)$$

and the Bianchi identities are, naturally, unchanged.) In particular, imposing the Lorentz condition appropriate to the medium we should find that the potentials obey the wave equation

$$\left[\frac{1}{v^2}\frac{\partial^2}{\partial t^2} - \vec{\nabla}^2\right]\psi = f \quad (14.1.6)$$

with the difference that the velocity of light in the vacuum, c , is replaced by its velocity in the medium, v , and the source, f , by the transport charge and current densities.

³The refractive index of the medium is defined by $n = \sqrt{\mu\epsilon/\mu_o\epsilon_o} = c/v$, where c is the velocity of light in the vacuum. This number is experimentally greater than unity which means that light travels *slower* in a medium than it does in a vacuum. This behavior is contrary to that of mechanical waves (waves that use the properties of the medium to propagate, like sound) whose velocity normally increases with the medium's density.

14.2 Non-conducting medium

For example, the source-free, plane, monochromatic electromagnetic wave in an isotropic, linear and non-dispersive medium looks identical to its counterpart in the vacuum (with the appropriate changes as indicated above):

$$\begin{aligned}\vec{A} &= \vec{A}_o e^{i(\vec{\kappa} \cdot \vec{r} - \omega t)} \\ \phi &= \phi_o e^{i(\vec{\kappa} \cdot \vec{r} - \omega t)}\end{aligned}\quad (14.2.1)$$

where $\vec{\kappa}^2 = \omega^2/v^2 = n^2\omega^2/c^2$ and $|\vec{\kappa}|$ is the wave-number. A scalar relation between the wave-vector and the frequency is called a “dispersion” relation. We will take both $\vec{\kappa}$ and ω to be real, but \vec{A}_o and ϕ_o may be complex. The electric and magnetic fields are related to the potentials by

$$\begin{aligned}\vec{B} = \vec{\nabla} \times \vec{A} &\rightarrow \vec{B} = i\vec{\kappa} \times \vec{A} \\ \vec{E} = -\vec{\nabla}\phi - \frac{\partial \vec{A}}{\partial t} &\rightarrow \vec{E} = -i\vec{\kappa}\phi + i\omega\vec{A}\end{aligned}\quad (14.2.2)$$

and one must supplement these relations, as before, with the gauge condition

$$i\vec{\kappa} \cdot \vec{A} - \frac{i\omega}{v^2}\phi = 0. \quad (14.2.3)$$

Thus,

$$\begin{aligned}\vec{E} \times \vec{B} &= \frac{v^2}{\omega} \vec{B}^2 \vec{\kappa} = \frac{c^2}{n^2\omega} \vec{B}^2 \vec{\kappa} \\ \vec{B} \times \vec{\kappa} &= \frac{\omega}{v^2} \vec{E} = \frac{n^2\omega}{c^2} \vec{E} \\ \vec{\kappa} \times \vec{E} &= \omega \vec{B}\end{aligned}\quad (14.2.4)$$

The electric and magnetic fields oscillate as before within the constant phase planes while the wave itself propagates perpendicular to them. We could write the solutions for the electric and magnetic fields as

$$\begin{aligned}\vec{E} &= \vec{E}_o e^{i(\vec{\kappa} \cdot \vec{r} - \omega t)} \\ \vec{B} &= \vec{B}_o e^{i(\vec{\kappa} \cdot \vec{r} - \omega t)}\end{aligned}\quad (14.2.5)$$

It should be remembered that no restrictions have been placed on the vectors \vec{E}_o and \vec{B}_o . In particular, they should *ab initio* be considered as complex vectors. However, the

physical \vec{E} and \vec{B} fields must be real, so in the end we are interested in the real part of the solutions above. As we take κ to be real, it follows that $\vec{B}_o = n/c(\hat{\kappa} \times \vec{E}_o)$ (from the last equation in (14.2.4)). A wave that obeys the condition that both the electric field and the magnetic field are perpendicular to $\vec{\kappa}$ is called a transverse wave.⁴

14.2.1 Energy and momentum density

We will now examine these solutions. To start with we will be interested in the time averaged values of the electromagnetic energy density and the Poynting vector. Let us prove a little theorem.

Theorem:

Let $f(\vec{r}, t) = f_o(\vec{r})e^{i\omega t}$ and $g(\vec{r}, t) = g_o(\vec{r})e^{i\omega t}$ be two arbitrary complex functions such that f_o and g_o (both also complex) do not depend on t . Then

$$\langle \text{Re}(f(\vec{r}, t)), \text{Re}(g(\vec{r}, t)) \rangle_\tau = \frac{1}{2} \text{Re}(f_o^* g_o) \quad (14.2.6)$$

where f^* is the complex conjugate of f and $\langle \ , \ \rangle_\tau$ represents the time average over one period ($\tau = 2\pi/\omega$), *i.e.*,

$$\langle f(t), g(t) \rangle_\tau = \frac{1}{\tau} \int_0^\tau dt f(t)g(t) \quad (14.2.7)$$

The proof is very simple. We can write $f_o(\vec{r}) = u(\vec{r}) + iv(\vec{r})$ and $g_o(\vec{r}) = p(\vec{r}) + iq(\vec{r})$ so that

$$\begin{aligned} f(\vec{r}, t) &= (u + iv)(\cos \omega t + i \sin \omega t) = (u \cos \omega t - v \sin \omega t) + i(v \cos \omega t + u \sin \omega t) \\ g(\vec{r}, t) &= (p + iq)(\cos \omega t + i \sin \omega t) = (p \cos \omega t - q \sin \omega t) + i(q \cos \omega t + p \sin \omega t) \end{aligned} \quad (14.2.8)$$

so that

$$\text{Re}(f) \cdot \text{Re}(g) = up \cos^2 \omega t - (uq + vp) \cos \omega t \sin \omega t + vq \sin^2 \omega t \quad (14.2.9)$$

The definition in (14.2.7) therefore involves three integrals: over $\cos^2 \omega t$, $\cos \omega t \sin \omega t$ and $\sin^2 \omega t$. Now we know that the second of these is identically zero and the first and last are simply 1/2. Evidently therefore,

$$\langle \text{Re}(f(\vec{r}, t)), \text{Re}(g(\vec{r}, t)) \rangle_\tau = \frac{1}{2}(up + vq) = \frac{1}{2} \text{Re}(f_o^* g_o)$$

⁴Problem: Show that if $\vec{\kappa}$ is real then (14.2.4) imply that $(\text{Re}(\vec{E}), \text{Re}(\vec{B}), \vec{\kappa})$ form a right-handed orthogonal set. Hint: expand $\vec{E} = \vec{E}^R + i\vec{E}^I$ $\vec{B} = \vec{B}^R + i\vec{B}^I$. Then noting that $\hat{\kappa}$ is real, verify that $\vec{B}^{R,I} \times \hat{\kappa} = n/c\vec{E}^{R,I}$ and $\hat{\kappa} \times \vec{E}^{R,I} = c/n\vec{B}^{R,I}$. This solves part of the problem. Argue quickly that $\vec{E}^R \times \vec{B}^R = B^{R2}\hat{\kappa}$!

■

Now the plane wave solutions in (14.2.1) are, in general, complex and the physical fields are obtained by taking the real parts. Let us assume, however, that $\vec{\kappa}$ and ω are real. The electromagnetic energy density is

$$u = \frac{1}{2} \left(\epsilon [\text{Re}(\vec{E})]^2 + \frac{1}{\mu} [\text{Re}(\vec{B})]^2 \right) = \frac{1}{2} \epsilon \left([\text{Re}(\vec{E})]^2 + \frac{c^2}{n^2} [\text{Re}(\vec{B})]^2 \right) \quad (14.2.10)$$

where we have used $1/\epsilon\mu = v^2 = c^2/n^2$. Therefore, applying the theorem we just proved

$$\langle u \rangle = \frac{1}{2} \epsilon \left(\langle [\text{Re}(\vec{E})]^2 \rangle + \frac{c^2}{n^2} \langle [\text{Re}(\vec{B})]^2 \rangle \right) = \frac{1}{4} \epsilon \left(|\vec{E}_o|^2 + \frac{c^2}{n^2} |\vec{B}_o|^2 \right) \quad (14.2.11)$$

Again, using the fact that $\vec{B}_o = n/c(\hat{\kappa} \times \vec{E}_o)$ we find

$$\langle u \rangle = \frac{1}{2} \epsilon |\vec{E}_o|^2 \quad (14.2.12)$$

We may compute the time averaged Poynting vector likewise:

$$\begin{aligned} \langle \vec{S} \rangle &= \langle \text{Re}(\vec{E}) \times \text{Re}(\vec{H}) \rangle = \frac{1}{\mu} \langle \text{Re}(\vec{E}) \times \text{Re}(\vec{B}) \rangle = \frac{1}{2\mu} \text{Re}(\vec{E}^* \times \vec{B}) \\ &= \frac{1}{2\mu} \text{Re}(\vec{E}_o^* \times \vec{B}_o) = \frac{n}{2\mu c} \text{Re}(\vec{E}_o^* \times (\hat{\kappa} \times \vec{E}_o)) \\ &= \frac{n}{2\mu c} |\vec{E}_o|^2 \hat{\kappa} = \frac{c}{n} \langle u \rangle \hat{\kappa} = \langle u \rangle v \hat{\kappa} \end{aligned} \quad (14.2.13)$$

which is a simple relationship that the momentum density bears to the energy density. It is similar to the relation $\vec{j} = \rho \vec{v}$ for the current density, where ρ is the mass density. This similarity reinforces our conceptualization of \vec{S} as an energy current density, *i.e.*, as energy density in transport at the phase velocity, $v = c/n$.

14.2.2 Polarization

Let us now take a look at another feature of our complex solutions, this one showing how the complex constants carry information about the behavior of the fields. Recall (14.2.5) and assume that $\vec{\kappa}$ and ω are real. Then only \vec{E}_o and \vec{B}_o are complex vectors. Take an arbitrary, real and *rigid* right handed basis $(\hat{e}_1, \hat{e}_2, \hat{e}_3)$, such that one of the basis vectors, say \hat{e}_3 , points in the direction of $\vec{\kappa}$, *i.e.*, $\hat{e}_3 = \hat{\kappa}$. Then \hat{e}_1 and \hat{e}_2 lie in the constant phase plane containing \vec{E} and \vec{B} . We could write

$$\vec{E}_o = E_1 \hat{e}_1 + E_2 \hat{e}_2 \quad (14.2.14)$$

where $E_{1,2}$ are clearly the components of \vec{E}_o in the directions of, respectively, \hat{e}_1 and \hat{e}_2 . $E_{1,2}$ are, in general, complex numbers, so let us write them as a magnitude times a phase, *i.e.*,

$$E_{1,2} = |E_{1,2}|e^{i\phi_{1,2}} \quad (14.2.15)$$

Although we have used *arbitrary* phases, ϕ_1 and ϕ_2 , for the components of \vec{E}_o , only the difference between the phases will be of physical relevance. Our solution, for example, for the real part of the electromagnetic field reads

$$\text{Re}(\vec{E}) = |E_1|\hat{e}_1 \cos(\vec{\kappa} \cdot \vec{r} - \omega t + \phi_1) + |E_2|\hat{e}_2 \cos(\vec{\kappa} \cdot \vec{r} - \omega t + \phi_2). \quad (14.2.16)$$

Consider this expression at some fixed spatial point, say $\vec{r} = 0$ (for convenience), then

$$\text{Re}(\vec{E}) = |E_1|\hat{e}_1 \cos(\omega t - \phi_1) + |E_2|\hat{e}_2 \cos(\omega t - \phi_2) \quad (14.2.17)$$

and let's now examine some special cases:

- Case 1: $\phi_1 = \phi_2 = 0$.

We have

$$\text{Re}(\vec{E}) = (|E_1|\hat{e}_1 + |E_2|\hat{e}_2) \cos(\omega t) \quad (14.2.18)$$

so that the electric field oscillates along the same straight line from a maximum of

$$+\sqrt{|E_1|^2 + |E_2|^2}$$

to a minimum of

$$-\sqrt{|E_1|^2 + |E_2|^2}.$$

The situation is analogous to two coupled harmonic oscillators in which the springs have the same Hooke's constant and the relative phase between them is zero (see figure 11). In such a situation, the electric field is said to be *linearly polarized*.

- Case 2: $\phi_1 = \frac{\pi}{2}$, $\phi_2 = 0$.

In this case we find the solution

$$\text{Re}(\vec{E}) = |E_1|\hat{e}_1 \sin \omega t + |E_2|\hat{e}_2 \cos \omega t = E_x \hat{e}_1 + E_y \hat{e}_2 \quad (14.2.19)$$

where we have defined $E_x = |E_1| \sin \omega t$ and $E_y = |E_2| \cos \omega t$. These are the (time-dependent) components of the \vec{E} -field in the directions of the basis vectors \hat{e}_1 and \hat{e}_2 . Now notice that

$$E_x = |E_1| \sqrt{1 - \cos^2 \omega t} = |E_1| \sqrt{1 - \frac{E_y^2}{|E_2|^2}} \quad (14.2.20)$$

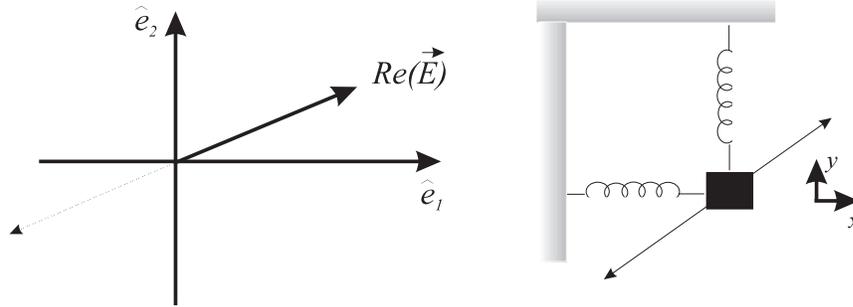


Figure 14.1: Linear Polarization

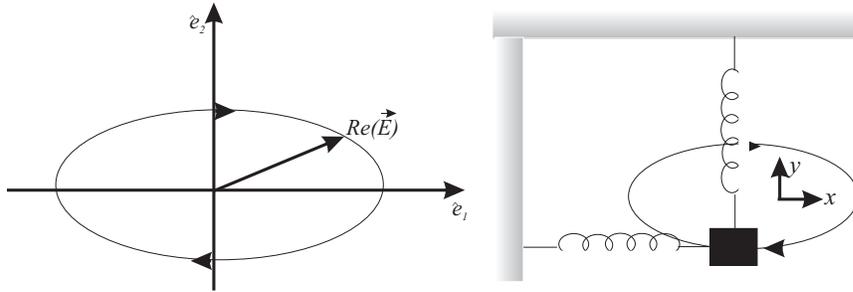


Figure 14.2: Right (negative helicity) Elliptical Polarization

so that

$$\frac{E_x^2}{|E_1|^2} + \frac{E_y^2}{|E_2|^2} = 1 \quad (14.2.21)$$

which is the equation of an ellipse. The electric field vector therefore traces out an ellipse in the (\hat{e}_1, \hat{e}_2) plane, as shown in figure 12, in a *clock-wise* direction. The electric field is said to be right (negative “helicity”) elliptically polarized. Notice that (a) if $|E_1| = |E_2|$ then the ellipse becomes a circle and the field is right (negative helicity) circularly polarized and (b) if $\phi_1 = -\pi/2$ and $\phi_2 = 0$ (so that the relative phase is the opposite) then the electric field would sweep in an *counter-clock-wise* direction and the wave would be left (positive helicity) elliptically polarized or, if $|E_1| = |E_2|$, left (positive helicity) circularly polarized.

For arbitrary phases we still find elliptical polarization, but in general with the principal axes not oriented along \hat{e}_1 and \hat{e}_2 . (As the frequencies of the two components are the same, we won’t find the analogue of Lissajou’s figures.) For the electric field at $\vec{r} = 0$, one

has

$$\operatorname{Re}(\vec{E}) = |E_1|\hat{e}_1 \cos(\omega t - \phi_1) + |E_2|\hat{e}_2 \cos(\omega t - \phi_2) \quad (14.2.22)$$

or

$$\begin{aligned} E_x &= |E_1| \cos(\omega t - \phi_1) \\ E_y &= |E_2| \cos(\omega t - \phi_2) \end{aligned} \quad (14.2.23)$$

Notice that

$$\begin{aligned} E_x &= |E_1| \cos(\omega t - \phi_2 + \delta), \quad \delta = \phi_2 - \phi_1 \\ &= |E_1|(\cos(\omega t - \phi_2) \cos \delta - \sin(\omega t - \phi_2) \sin \delta) \\ &= |E_1| \left(\frac{E_y}{|E_2|} \cos \delta - \sqrt{1 - \frac{E_y^2}{|E_2|^2}} \sin \delta \right) \end{aligned} \quad (14.2.24)$$

i.e.,

$$\left(\frac{E_x}{|E_1|} - \frac{E_y}{|E_2|} \cos \delta \right)^2 = \sin^2 \delta \left(1 - \frac{E_y^2}{|E_2|^2} \right) \quad (14.2.25)$$

Rearranging terms gives the equation

$$\frac{E_x^2}{|E_1|^2} + \frac{E_y^2}{|E_2|^2} - 2 \frac{E_x E_y}{|E_1| |E_2|} \cos \delta = \sin^2 \delta \quad (14.2.26)$$

This will be recognized as an ellipse whose principal axes make an angle, say θ , with the basis vectors (\hat{e}_1, \hat{e}_2) . To determine the angle θ , simply perform a rotation by θ and require that the equation (14.2.26), written in terms of the new components of \vec{E}_o , *viz.* E'_x and E'_y , returns to the standard form. This procedure gives,⁵

$$\tan 2\theta = \frac{2|E_1||E_2|}{|E_2|^2 - |E_1|^2} \cos \delta \quad (14.2.27)$$

Again we find the similarity with two harmonic oscillators of the same Hooke's constant. We see that only the relative phase, δ , is relevant and that $\theta = 0$ when $\delta = \frac{\pi}{2}$. Of course, when $|E_1| = |E_2|$ the ellipse degenerates to a circle and the question of rotating the basis does not arise.

⁵Problem: Show this!

An alternative description of the polarization can be given in the basis $\hat{e}_\pm = \frac{1}{\sqrt{2}}(\hat{e}_1 \pm i\hat{e}_2)$, then $\hat{e}_1 = \frac{1}{\sqrt{2}}(\hat{e}_+ + \hat{e}_-)$ and $\hat{e}_2 = \frac{i}{\sqrt{2}}(\hat{e}_+ - \hat{e}_-)$. They are orthonormal, obeying the relations

$$\hat{e}_\pm^* \cdot \hat{e}_\mp = 0, \quad \hat{e}_\pm^* \cdot \hat{e}_3 = 0, \quad \hat{e}_\pm^* \cdot \hat{e}_\pm = 1 \quad (14.2.28)$$

and we can write (14.2.14) as

$$\vec{E}_o = E_+ \hat{e}_+ + E_- \hat{e}_- \quad (14.2.29)$$

where $E_\pm = \frac{1}{\sqrt{2}}(E_1 \mp iE_2)$.

What about \vec{B} ? Because $(\hat{E}, \hat{B}, \hat{\kappa})$ form a right-handed basis, $\vec{B} = n/c(\hat{\kappa} \times \vec{E})$ and it follows that (recall that $\hat{\kappa} = \hat{e}_3$) if

$$\text{Re}(\vec{E}) = |E_1| \hat{e}_1 \cos(\omega t - \phi_1) + |E_2| \hat{e}_2 \cos(\omega t - \phi_2) \quad (14.2.30)$$

then

$$\text{Re}(\vec{B}) = \frac{n}{c} (|E_1| \hat{e}_2 \cos(\omega t - \phi_1) - |E_2| \hat{e}_1 \cos(\omega t - \phi_2)) \quad (14.2.31)$$

so \vec{B} follows \vec{E} , tracing an ellipse that is rotated 90° counter-clockwise.

The polarization state of radiation that is received in detectors can tell us a lot about the source of the radiation when direct access to the source is impossible.⁶ Thus, for example, the polarization state of radiation from some distant super-nova explosion or pulsar can tell us about the physical processes that led to the emission of the waves we eventually received, even though we have no direct access to the physical processes on the object of interest. In practice this is not so easy, however, because no natural source emits perfectly monochromatic radiation and there is always a superposition of waves of different frequencies even though the radiation may appear predominantly monochromatic.

14.3 Conducting medium

So far we have always assumed that \vec{k} and ω are real. Although this was not necessary, it made sense for non-conducting media and is compatible with the dispersion relation. We

⁶Problem: The polarization state of the plane wave is known if it can be written in one of the forms given in (14.2.14) or (14.2.29). However, confronted with a plane wave solution, \vec{E} , how can we tell its form (polarization)? There are effectively just three parameters to be determined from the electric field, *viz.* $|E_{1,2}|$ and the relative phase, δ . One can define four parameters, the Stokes parameters (not all of them will be independent), which contain this information. In the basis (\hat{e}_1, \hat{e}_2) ,

$$\begin{aligned} s_o &= |\hat{e}_1 \cdot \vec{E}|^2 + |\hat{e}_2 \cdot \vec{E}|^2 \\ s_1 &= |\hat{e}_1 \cdot \vec{E}|^2 - |\hat{e}_2 \cdot \vec{E}|^2 \\ s_2 &= 2\text{Re}[(\hat{e}_1 \cdot \vec{E})^* (\hat{e}_2 \cdot \vec{E})] \\ s_3 &= 2\text{Im}[(\hat{e}_1 \cdot \vec{E})^* (\hat{e}_2 \cdot \vec{E})] \end{aligned}$$

Compute each parameter, s_a , and convince yourself that the three parameters are indeed determined from them. Also, show that they obey the relation $s_o^2 - s_1^2 - s_2^2 - s_3^2 = 0$.

now consider the case of plane, monochromatic wave solutions in conducting media. The situation is much more complicated, because the dispersion relation is such that both ω and κ cannot be taken to be simultaneously real. A conducting medium is dissipative. This, as we will shortly see, is because in conducting media the current density is proportional to the electric field (Ohm's Law),

$$\vec{j} = g\vec{E} \quad (14.3.1)$$

where g is the electric conductivity of the medium. Physically, the medium is dissipative because the movement of charges within the conducting medium extracts energy from the electromagnetic field. Begin, as before, with Maxwell's equations and set

$$\vec{B} = \vec{\nabla} \times \vec{A}, \quad \vec{E} = -\vec{\nabla}\phi - \frac{\partial \vec{A}}{\partial t} \quad (14.3.2)$$

Substituting these into the dynamical equations we get ($\rho = 0$)

$$\begin{aligned} -\vec{\nabla}^2\phi - \frac{\partial}{\partial t}(\vec{\nabla} \cdot \vec{A}) &= 0 \\ \frac{1}{v^2} \frac{\partial^2 \vec{A}}{\partial t^2} - \vec{\nabla}^2 \vec{A} + \mu g \frac{\partial \vec{A}}{\partial t} + \vec{\nabla} \left(\vec{\nabla} \cdot \vec{A} + \frac{1}{v^2} \frac{\partial \phi}{\partial t} + \mu g \phi \right) &= 0 \end{aligned} \quad (14.3.3)$$

and examining the second equation makes it clear that the best gauge condition to choose would be⁷

$$\vec{\nabla} \cdot \vec{A} + \frac{1}{v^2} \frac{\partial \phi}{\partial t} + \mu g \phi = 0 \quad (14.3.4)$$

Then one finds the same homogeneous equation for both ϕ and \vec{A} , *i.e.*,

$$\left[\frac{1}{v^2} \frac{\partial^2}{\partial t^2} - \vec{\nabla}^2 + \mu g \frac{\partial}{\partial t} \right] \psi = 0 \quad (14.3.5)$$

which differs from the ordinary wave-equation by a first order time-derivative term. It turns out that this term makes all the difference. Let us solve this equation by a separation of variables. Assume a harmonic solution

$$\psi(\vec{r}, t) = e^{-i\omega t} \psi(\vec{r}) \quad (14.3.6)$$

and find that the equation for $\psi(\vec{r})$ is just

$$\vec{\nabla}^2 \psi + \left(\frac{\omega^2}{v^2} + i\omega\mu g \right) \psi = \vec{\nabla}^2 \psi + \frac{\omega^2}{v^2} \left(1 + i \frac{g}{\epsilon\omega} \right) \psi = 0 \quad (14.3.7)$$

The constant ϵ/g has dimensions of time and represents a characteristic relaxation time, τ , of the conductor. Notice that as g approaches infinity (perfect conductor), τ approaches

⁷Problem: Show that this is an acceptable gauge condition.

zero and vice versa. When τ is large compared to the period of the electromagnetic wave, the imaginary term becomes negligible and we return to the usual time independent wave equation for the fields. However, as τ becomes small (compared to $1/\omega$) the imaginary term dominates. In that case, the effective equation inside the conductor becomes

$$\vec{\nabla}^2 \psi + i\omega\mu g \approx 0 \quad (14.3.8)$$

which is a time independent *diffusion* equation for the fields. For most metals, $\tau \approx 10^{-14}$ s, which implies that diffusion dominates for all frequencies lower than optical frequencies. In other words, propagation is negligible for those frequencies.

A solution for $\psi(\vec{r})$ may still be written in the form

$$\psi(\vec{r}) = e^{i\vec{\kappa} \cdot \vec{r}} \quad (14.3.9)$$

but now we see that the dispersion relation reads

$$\frac{\omega^2}{v^2} + i\omega\mu g - \vec{\kappa} \cdot \vec{\kappa} = 0 \quad (14.3.10)$$

so ω and $\vec{\kappa}$ cannot both be real. If ω is real, so it is the angular frequency of the wave, then, as v is also real, $\kappa^2 = \vec{\kappa} \cdot \vec{\kappa}$ cannot be real unless $g = 0$ (which takes us back to a non-conducting medium). The solutions take the formal form

$$\begin{aligned} \vec{A} &= \vec{A}_0 e^{i(\vec{\kappa} \cdot \vec{r} - \omega t)} \\ \phi &= \phi_0 e^{i(\vec{\kappa} \cdot \vec{r} - \omega t)} \end{aligned} \quad (14.3.11)$$

but we should be careful in interpreting it because, while ω is real, $\vec{\kappa}$ is complex. As before, these equations should be supplemented with our gauge condition, which reads

$$i\vec{\kappa} \cdot \vec{A} - i \left(\frac{\omega}{v^2} + i\mu g \right) \phi = 0 \quad (14.3.12)$$

This last equation helps express ϕ in terms of \vec{A} as

$$\phi = \frac{(\vec{\kappa} \cdot \vec{A})}{\frac{\omega}{v^2} + i\mu g} \quad (14.3.13)$$

Then, taking derivatives, we find

$$\begin{aligned} \vec{E} &= -i\vec{\kappa} \frac{(\vec{\kappa} \cdot \vec{A})}{\frac{\omega}{v^2} + i\mu g} + i\omega \vec{A} \\ \vec{B} &= i\vec{\kappa} \times \vec{A} \end{aligned} \quad (14.3.14)$$

and it is a straightforward exercise to prove the following relations

$$\begin{aligned}\vec{E} \times \vec{B} &= \frac{\omega}{\vec{\kappa}^2} \vec{B}^2 \vec{\kappa} = \frac{c^2}{n^2 \omega} \vec{B}^2 \vec{\kappa} \\ \vec{B} \times \vec{\kappa} &= \frac{\vec{\kappa}^2}{\omega} \vec{E} = \frac{n^2 \omega}{c^2} \vec{E} \\ \vec{\kappa} \times \vec{E} &= \omega \vec{B}\end{aligned}\tag{14.3.15}$$

where we have defined the refractive index, $n = c\kappa/\omega$, in analogy with plane waves in non-conducting media. Formally, therefore, there is no distinction between these and the relations in (14.2.4). But, we must proceed with great care in interpreting them because the refractive index n is no longer real but complex. In general, this means that the vectors \vec{E} and \vec{B} are neither in phase nor are they necessarily perpendicular to each other. Let us note first that using the dispersion relation

$$n = \frac{c}{\omega} \sqrt{\frac{\omega^2}{v^2} + i\omega\mu g} = n^R + in^I\tag{14.3.16}$$

where $n^{R,I}$ refer respectively to the real and imaginary parts of the refractive index, n . By equating real and imaginary parts, one finds $n^{R,I}$ in terms of the real constants. Thus

$$(n^R + in^I)^2 = n^{R2} - n^{I2} + 2in^R n^I = \frac{c^2}{v^2} + i\frac{\mu g c^2}{\omega}\tag{14.3.17}$$

giving

$$\begin{aligned}n^{R2} - n^{I2} &= \frac{c^2}{v^2} \\ 2n^R n^I &= \frac{\mu g c^2}{\omega}\end{aligned}\tag{14.3.18}$$

These equations can be solved,

$$\begin{aligned}n^R &= \sqrt{\frac{1}{2} \left[\frac{c^2}{v^2} + \sqrt{\frac{c^4}{v^4} + \frac{\mu^2 g^2 c^4}{\omega^2}} \right]} \\ n^I &= \sqrt{\frac{1}{2} \left[-\frac{c^2}{v^2} + \sqrt{\frac{c^4}{v^4} + \frac{\mu^2 g^2 c^4}{\omega^2}} \right]}\end{aligned}\tag{14.3.19}$$

keeping in mind that only the real solutions are meaningful. Using $c = 1/\sqrt{\epsilon_o \mu_o}$ and $v = 1/\sqrt{\epsilon \mu}$, they may also be written as

$$n^{R,I} = c \sqrt{\frac{\mu}{2}} \sqrt{\pm \epsilon + \sqrt{\epsilon^2 + \frac{g^2}{\omega^2}}}\tag{14.3.20}$$

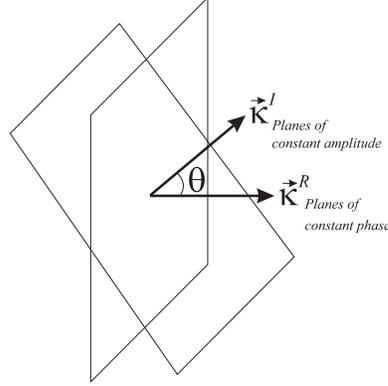


Figure 14.3: Plane of constant phase and of constant amplitude

which is a very convenient form as they are expressed exclusively in terms of the properties of the medium and the frequency of the radiation.

Likewise, because $\vec{\kappa}$ is not real, let us also write the general form of $\vec{\kappa}$ as

$$\vec{\kappa} = \vec{\kappa}^R + i\vec{\kappa}^I \quad (14.3.21)$$

where $\vec{\kappa}^{R,I}$ are real vectors and refer to the real/imaginary parts of $\vec{\kappa}$ respectively. With this we find that

$$\begin{aligned} \vec{E} &= \vec{E}_0 e^{-\vec{\kappa}^I \cdot \vec{r}} e^{i(\vec{\kappa}^R \cdot \vec{r} - \omega t)} \\ \vec{B} &= \vec{B}_0 e^{-\vec{\kappa}^I \cdot \vec{r}} e^{i(\vec{\kappa}^R \cdot \vec{r} - \omega t)} \end{aligned} \quad (14.3.22)$$

which show a piece that exponentially decays in the direction of $\vec{\kappa}^I$ (this is $e^{-\vec{\kappa}^I \cdot \vec{r}}$) and another piece which oscillates (this is $e^{i(\vec{\kappa}^R \cdot \vec{r} - \omega t)}$). The surfaces of constant phase are once again planes given by the equation

$$\vec{\kappa}^R \cdot \vec{r} = \text{const.} \quad (14.3.23)$$

and so having $\vec{\kappa}^R$ for a normal. The surfaces of constant amplitude, on the other hand, while also planes, obey the equation

$$\vec{\kappa}^I \cdot \vec{r} = \text{const.} \quad (14.3.24)$$

and are not in general parallel to the planes of constant phase. Planes of constant phase form an angle

$$\theta = \cos^{-1}(\hat{\kappa}^R \cdot \hat{\kappa}^I) \quad (14.3.25)$$

with planes of constant amplitude (see figure 13).

Now, because $\vec{\kappa}$ (or n) is complex the relations (14.3.15) are by no means simple to interpret. For example, consider the last of these:

$$\begin{aligned} (\vec{\kappa}^R + i\vec{\kappa}^I) \times (\vec{E}^R + i\vec{E}^I) &= (\vec{\kappa}^R \times \vec{E}^R - \vec{\kappa}^I \times \vec{E}^I) + i(\vec{\kappa}^I \times \vec{E}^R + \vec{\kappa}^R \times \vec{E}^I) \\ &= \omega(\vec{B}^R + i\vec{B}^I) \end{aligned} \quad (14.3.26)$$

implies that

$$\begin{aligned} (\vec{\kappa}^R \times \vec{E}^R - \vec{\kappa}^I \times \vec{E}^I) &= \omega\vec{B}^R \\ (\vec{\kappa}^I \times \vec{E}^R + \vec{\kappa}^R \times \vec{E}^I) &= \omega\vec{B}^I \end{aligned} \quad (14.3.27)$$

and so on. So we cannot say anything about the relative orientations of the vectors $\text{Re}(\vec{E})$, $\text{Re}(\vec{B})$ and $\text{Re}(\vec{\kappa})$. A greatly simplified situation is one in which both $\vec{\kappa}^R$ and $\vec{\kappa}^I$ point in the same direction. In that case, we may write

$$\vec{\kappa} = (\kappa^R + i\kappa^I)\hat{u} \quad (14.3.28)$$

where \hat{u} is a unit vector in the direction common to the real and imaginary parts of $\vec{\kappa}$. In this case we may relate the real and imaginary parts of n and $\vec{\kappa}$ in a straightforward way:

$$n^R = \frac{c\kappa^R}{\omega}, \quad n^I = \frac{c\kappa^I}{\omega} \quad (14.3.29)$$

Now we see that the second and third of (14.3.15) are

$$\begin{aligned} \vec{B} \times \hat{u} &= \frac{n}{c}\vec{E} \\ \hat{u} \times \vec{E} &= \frac{c}{n}\vec{B} \end{aligned} \quad (14.3.30)$$

We immediately deduce that $\text{Re}(\vec{E})$ and $\text{Re}(\vec{B})$ are perpendicular to \hat{u} , so the waves are certainly transverse. However, because n is complex, we cannot conclude that $\text{Re}(\vec{E})$ and $\text{Re}(\vec{B})$ are perpendicular to each other except if the relative phase between \vec{E} and \vec{B} vanishes, *i.e.*, linear polarization.⁸

The amplitude of the wave decays as

$$\exp[-\kappa^I \hat{u} \cdot \vec{r}] = \exp\left[-\frac{\omega n^I}{c} \hat{u} \cdot \vec{r}\right] = \exp\left[-\frac{\omega n^I}{c} \xi\right] \quad (14.3.31)$$

⁸Problem: Prove this!

where $\xi = \hat{u} \cdot \vec{r}$ represents the depth traveled in the medium by the wave. The amplitude therefore decays to $1/e$ of its original value when

$$\xi = \frac{1}{\kappa^I} = \frac{c}{\omega n^I} = \delta \quad (14.3.32)$$

However, the real wave-number is now $\kappa^R = \omega n^R/c = 2\pi/\lambda$, giving $c/\omega = \lambda n^R/2\pi$, thus

$$\delta = \frac{n^R}{n^I} \frac{\lambda}{2\pi} \quad (14.3.33)$$

The distance at which the wave amplitude falls to $1/e$ of its original value (δ) is called the “skin depth” of the medium. We see that for media in which $\vec{\kappa}^R$ and $\vec{\kappa}^I$ have the same direction, the skin depth is proportional to the ratio of n^R/n^I . Thus, if $n^R/n^I \approx 1$ the wave will not even penetrate one wavelength into the medium and the medium is opaque. If, on the contrary, $n^R/n^I \gg 1$ then the wave will penetrate many wavelengths into the medium without suffering an appreciable loss of amplitude. In this case the medium is transparent.

Checking back with equation (14.3.20) we see that

- if $\omega \gg |g/\epsilon|$ then $n^R \approx c/v \gg n^I$
- if $\omega \ll |g/\epsilon|$ then $n^R \approx n^I \approx c\sqrt{\mu g/2\omega}$

Chapter 15

Interfaces between media

15.1 Introduction

So far we have considered only the propagation of waves in conducting and non-conducting media but we have not considered the behavior of these waves when they encounter a boundary between two media of different electric and magnetic properties. It is known from daily experience that the waves suffer reflection and refraction at such boundaries. The phenomena of reflection and refraction divide themselves into two classes, *viz.*

1. Kinematic properties
 - (a) Geometric optics laws of reflection
 - (b) Snell's laws (of refraction)
2. Dynamical properties
 - (a) Intensities of reflected and refracted beams
 - (b) Phase changes and polarization of the beams

The kinematic properties are a consequence of the very existence of boundary conditions, which implies that the spatial and temporal variation of the fields must be the same at all points of incidence. The dynamical properties result from the boundary conditions appropriate to the media being considered and have been discussed in section 10.

15.2 Interfaces: non-conducting media

Consider, therefore, a plane monochromatic wave that is incident (at an arbitrary angle) at a *plane* boundary separating two media which we label "1" and "2" respectively. Assume

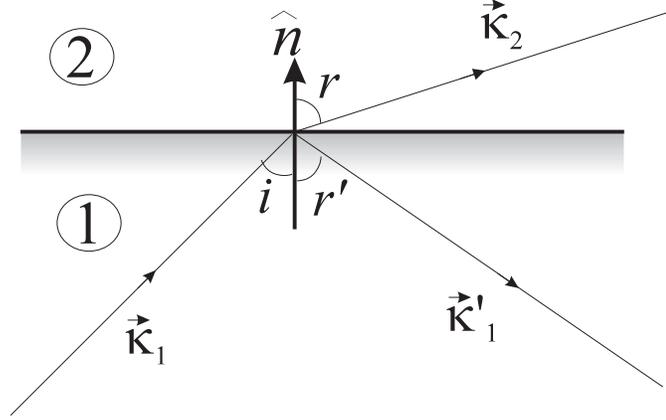


Figure 15.1: Reflection and Refraction: non-conducting media

for definiteness that the ray is incident from medium “1”. The situation is shown in figure 14. Let \hat{n} be the normal to the boundary, pointing into medium “2” as shown and let the wave-vectors in each region be given by \vec{k}_1 (incident), \vec{k}'_1 (reflected) and \vec{k}_2 (transmitted) respectively. The two media are defined by their respective permittivities and permeabilities, $(\epsilon_{1,2}, \mu_{1,2})$. The incident ray \vec{k}_1 makes an angle “ i ” with the normal, the reflected ray, \vec{k}'_1 , an angle “ r' ” with the normal and the transmitted ray, \vec{k}_2 an angle “ r ” with the normal. We can write the following expressions for the electric and magnetic fields in each case

1. Incident ray:

$$\begin{aligned}\vec{E}_1 &= \vec{E}_{1o} e^{i(\vec{k}_1 \cdot \vec{r} - \omega t)} \\ \vec{B}_1 &= \frac{n_1}{c} \hat{k}_1 \times \vec{E}_{1o} e^{i(\vec{k}_1 \cdot \vec{r} - \omega t)}\end{aligned}\quad (15.2.1)$$

2. Reflected ray:

$$\begin{aligned}\vec{E}'_1 &= \vec{E}'_{1o} e^{i(\vec{k}'_1 \cdot \vec{r} - \omega t)} \\ \vec{B}'_1 &= \frac{n_1}{c} \hat{k}'_1 \times \vec{E}'_{1o} e^{i(\vec{k}'_1 \cdot \vec{r} - \omega t)}\end{aligned}\quad (15.2.2)$$

3. Transmitted ray:

$$\begin{aligned}\vec{E}_2 &= \vec{E}_{2o} e^{i(\vec{k}_2 \cdot \vec{r} - \omega t)} \\ \vec{B}_2 &= \frac{n_2}{c} \hat{k}_2 \times \vec{E}_{2o} e^{i(\vec{k}_2 \cdot \vec{r} - \omega t)}\end{aligned}\quad (15.2.3)$$

Kinematics

Let us now apply the kinematic requirement that the spatial and temporal variations should be the same at the point of incidence; we find then that

$$\vec{\kappa}_1 \cdot \vec{r} = \vec{\kappa}'_1 \cdot \vec{r} = \vec{\kappa}_2 \cdot \vec{r} \Big|_{\text{incidence}} \quad (15.2.4)$$

Each term in the above equation defines a plane normal to the respective wave-vector, which is not surprising as surfaces of constant phase are planes for the solutions we are using. We will now show the first part of both the geometric optics “laws of reflection” and “Snell’s law”, *i.e.*, that the vectors $(\vec{\kappa}_1, \vec{\kappa}'_1, \vec{\kappa}_2, \hat{n})$ all lie in the same plane, the plane of incidence. Take the planar interface to be given by $z = 0$, so that the origin of coordinates is on the interface itself. This simplifies the considerations without loss of generality, for then \hat{n} is perpendicular to the position \vec{r} of the point of incidence. Consider the identity

$$\hat{n} \times (\hat{n} \times \vec{r}) = (\hat{n} \cdot \vec{r})\hat{n} - \hat{n}^2 \vec{r} \quad (15.2.5)$$

but, since $\hat{n}^2 = 1$ and $\hat{n} \cdot \vec{r} = 0$, we have

$$\vec{r} = -\hat{n} \times (\hat{n} \times \vec{r}) \quad (15.2.6)$$

Thus (15.2.4) reads

$$\vec{\kappa}_1 \cdot \hat{n} \times (\hat{n} \times \vec{r}) = \vec{\kappa}'_1 \cdot \hat{n} \times (\hat{n} \times \vec{r}) = \vec{\kappa}_2 \cdot \hat{n} \times (\hat{n} \times \vec{r}) \Big|_{\text{incidence}} \quad (15.2.7)$$

In the triple product, exchange the dot product for the vector product¹ and re-write the above equations as

$$(\vec{\kappa}_1 \times \hat{n}) \cdot (\hat{n} \times \vec{r}) = (\vec{\kappa}'_1 \times \hat{n}) \cdot (\hat{n} \times \vec{r}) = (\vec{\kappa}_2 \times \hat{n}) \cdot (\hat{n} \times \vec{r}) \Big|_{\text{incidence}} \quad (15.2.8)$$

But \vec{r} is arbitrary (we have in no way specified its location), so it must be true that

$$\vec{\kappa}_1 \times \hat{n} = \vec{\kappa}'_1 \times \hat{n} = \vec{\kappa}_2 \times \hat{n} \quad (15.2.9)$$

Now, $\vec{\kappa}_1 \times \hat{n}$ points in the direction of the normal to the plane of incidence, *i.e.*, (by definition) the plane containing $\vec{\kappa}_1$ and \hat{n} . Likewise $\vec{\kappa}'_1 \times \hat{n}$ and $\vec{\kappa}_2 \times \hat{n}$ define normals to the planes of reflection (plane containing $\vec{\kappa}'_1$ and \hat{n}) and refraction (plane containing $\vec{\kappa}_2$ and \hat{n}). The fact that these normals all point in the same direction imply that the planes are coincident. Thus we have proved that $(\vec{\kappa}_1, \vec{\kappa}'_1, \vec{\kappa}_2, \hat{n})$ lie in the same plane.

Next let us take a closer look at figure 14 and write out the conditions explicitly. We have from either (15.2.4) or (15.2.9)

$$\kappa_1 \sin i = \kappa'_1 \sin r' = \kappa_2 \sin r. \quad (15.2.10)$$

¹Problem: Verify that this is allowed

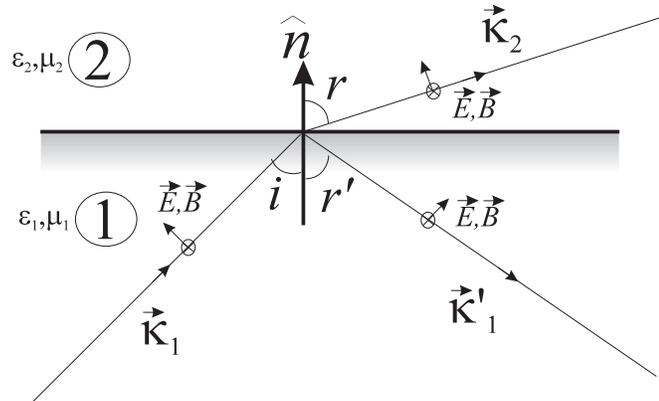


Figure 15.2: The electric field is perpendicular to the plane of incidence.

But $\kappa_2 = n_2\omega/c$ and $\kappa_1 = n_1\omega/c = \kappa'_1$, so $i = r'$, which is the law of reflection (the angle of incidence is equal to the angle of reflection) and we also have

$$\frac{\sin i}{\sin r} = \frac{\kappa_2}{\kappa_1} = \frac{n_2}{n_1} \quad (15.2.11)$$

which is Snell's law (of refraction). We have not had to use any property of the electric and magnetic fields to arrive at these laws because they are purely kinematic in nature and would hold true for waves of any kind (sound, for example). The truly dynamical properties will, however, depend heavily on the nature of the waves that are being considered. We will now consider these properties for electromagnetic waves.

Dynamics

The Dynamical properties are obviously obtained from the boundary conditions which are peculiar to the nature of the waves. In this case we must apply boundary conditions that are appropriate to the electric and magnetic fields. It is always true that

- The normal component of the magnetic field, \vec{B} , and the tangential component of the electric field, \vec{E} , are continuous across the boundary

Further, we will assume that the surface charge density of the interface is zero ($\sigma = 0$) as is the surface current density ($j_{\perp} = 0$) because we are considering non-conducting media. Then,

- The normal component of the electric displacement, \vec{D} , and the tangential component of the magnetization vector, \vec{H} , are continuous across the boundary

These two conditions translate into the following equations

$$\begin{aligned}
[\epsilon_1(\vec{E}_{1o} + \vec{E}'_{1o}) - \epsilon_2\vec{E}_{2o}] \cdot \hat{n} &= 0 \\
(\vec{E}_{1o} + \vec{E}'_{1o} - \vec{E}_{2o}) \times \hat{n} &= 0 \\
\left[\frac{n_1}{c}(\hat{\kappa}_1 \times \vec{E}_{1o} + \hat{\kappa}'_1 \times \vec{E}'_{1o}) - \frac{n_2}{c}(\hat{\kappa}_2 \times \vec{E}_{2o}) \right] \cdot \hat{n} &= 0 \\
\left[\frac{n_1}{\mu_1 c}(\hat{\kappa}_1 \times \vec{E}_{1o} + \hat{\kappa}'_1 \times \vec{E}'_{1o}) - \frac{n_2}{\mu_2 c}(\hat{\kappa}_2 \times \vec{E}_{2o}) \right] \times \hat{n} &= 0 \quad (15.2.12)
\end{aligned}$$

Let us recognize that

- the first equation arises because of the continuity of the normal component of \vec{D} ,
- the second equation arises because of the continuity of the tangential component of \vec{E} ,
- the third equation arises because of the continuity of the normal component of \vec{B} ,
- the last (fourth) equation arises because of the continuity of the tangential component of \vec{H} .

In applying the conditions we will now consider separately the two cases in which

1. the electric field is *perpendicular* to the plane of incidence (shown in figure 15)
2. the electric field is *in* to the plane of incidence (shown in figure 16)

Case 1: Electric field perpendicular to the plane of incidence (see figure 15)

In this case the first equation in (15.2.12) is empty: since all electric fields are perpendicular to the normal, \hat{n} , the inner product is automatically zero. Again, as the electric field is perpendicular to the plane of incidence and \hat{n} is a unit vector it follows from the second equation in (15.2.12) that

$$E_{1o} + E'_{1o} - E_{2o} = 0 \quad (15.2.13)$$

Now as $\hat{\kappa} \times \vec{E}$ generically points in the direction of \vec{B} it follows that

$$\begin{aligned}
(\hat{\kappa}_1 \times \vec{E}_{1o}) \cdot \hat{n} &= E_{1o} \sin i \\
(\hat{\kappa}'_1 \times \vec{E}'_{1o}) \cdot \hat{n} &= E'_{1o} \sin i \\
(\hat{\kappa}_2 \times \vec{E}_{2o}) \cdot \hat{n} &= E_{2o} \sin r \quad (15.2.14)
\end{aligned}$$

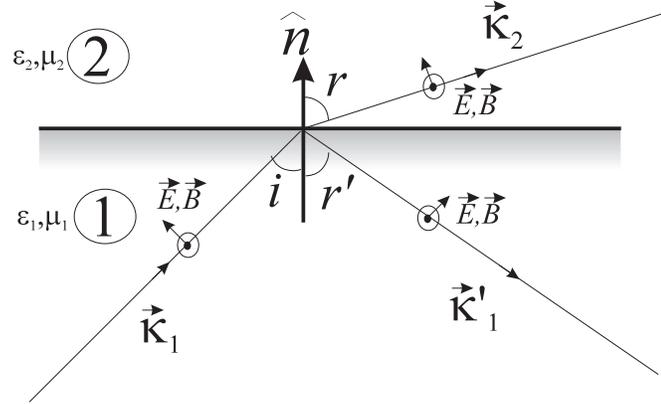


Figure 15.3: The electric field is in the plane of incidence.

and it follows that the third equation in (15.2.12) gives

$$\frac{n_1}{c}(E_{1o} + E'_{1o}) \sin i - \frac{n_2}{c}E_{2o} \sin r = 0 \quad (15.2.15)$$

But, using Snell's law we see that this equation duplicates the second equation and therefore provides no fresh information. On the other hand, the last equation in (15.2.12) gives

$$\frac{n_1}{\mu_1 c}(E_{1o} - E'_{1o}) \cos i - \frac{n_2}{\mu_2 c}E_{2o} \cos r = 0 \quad (15.2.16)$$

Thus we have ended up with two non-trivial and independent conditions: (15.2.13) and (15.2.16). Using (15.2.13) to write $E_{2o} = E_{1o} + E'_{1o}$ we get from (15.2.16),

$$\frac{E'_{1o}}{E_{1o}} = \frac{n_1 \cos i - \frac{\mu_1}{\mu_2} n_2 \cos r}{n_1 \cos i + \frac{\mu_1}{\mu_2} n_2 \cos r} \quad (15.2.17)$$

from which it also follows that

$$\frac{E_{2o}}{E_{1o}} = \frac{2n_1 \cos i}{n_1 \cos i + \frac{\mu_1}{\mu_2} n_2 \cos r} \quad (15.2.18)$$

The ratios $r_{12} = E'_{1o}/E_{1o}$ and $t_{12} = E_{2o}/E_{1o}$ are called the Fresnel coefficients, respectively for reflection and for transmission, at the interface between two non-conducting media. They give the amplitudes of the reflected ray and the transmitted ray in terms of the amplitude of the incident ray. They therefore represent the fraction of incident flux that

is reflected and refracted respectively. These expressions may be further put in terms of the angle of incidence only by using Snell's law:

$$\frac{\sin i}{\sin r} = \frac{n_2}{n_1} \rightarrow \cos r = \frac{1}{n_2} \sqrt{n_2^2 - n_1^2 \sin^2 i} \quad (15.2.19)$$

giving

$$\begin{aligned} r_{12}^\perp &= \frac{E'_{1o}}{E_{1o}} = \frac{n_1 \cos i - \frac{\mu_1}{\mu_2} \sqrt{n_2^2 - n_1^2 \sin^2 i}}{n_1 \cos i + \frac{\mu_1}{\mu_2} \sqrt{n_2^2 - n_1^2 \sin^2 i}} \\ t_{12}^\perp &= \frac{E_{2o}}{E_{1o}} = \frac{2n_1 \cos i}{n_1 \cos i + \frac{\mu_1}{\mu_2} \sqrt{n_2^2 - n_1^2 \sin^2 i}} \end{aligned} \quad (15.2.20)$$

As we said, these coefficients determine the portion of the incident flux that is reflected and refracted. We can make this notion more precise by noting that the flux is given by the Poynting vector as

$$\langle \vec{S} \rangle = \frac{n}{2\mu c} |\vec{E}_o|^2 \hat{k} \quad (15.2.21)$$

In terms of the Poynting vector in each region, one can define the “reflectance”, R , and “transmittance”, T , by the quantity of energy that reflects off or flows past the interface. This involves the projection of $\langle \vec{S} \rangle$ on the normal \hat{n} to the interface, so we define:

$$R = \frac{\hat{n} \cdot \langle \vec{S}'_1 \rangle}{\hat{n} \cdot \langle \vec{S}_1 \rangle}, \quad T = \frac{\hat{n} \cdot \langle \vec{S}_2 \rangle}{\hat{n} \cdot \langle \vec{S}_1 \rangle} \quad (15.2.22)$$

Clearly, therefore, for the case in hand,

$$\begin{aligned} R^\perp &= r_{12}^{\perp 2} \\ T^\perp &= \frac{\mu_1 n_2 \cos r}{\mu_2 n_1 \cos i} t_{12}^{\perp 2} = \frac{\mu_1 \sqrt{n_2^2 - n_1^2 \sin^2 i}}{\mu_2 n_1 \cos i} t_{12}^{\perp 2} \end{aligned} \quad (15.2.23)$$

It is not a difficult exercise to prove that $R^\perp + T^\perp = 1$, which expresses energy conservation at the interface.²

Case 2: Electric field in the plane of incidence (see figure 16)

The boundary conditions read a bit differently now that the electric field is in the plane of incidence. From (15.2.12) we have

$$\vec{E}_{1o} \cdot \hat{n} = E_{1o} \sin i$$

²Problem: Prove this.

$$\begin{aligned}
\vec{E}'_{1o} \cdot \hat{n} &= E'_{1o} \sin i \\
\vec{E}_{2o} \cdot \hat{n} &= E'_{1o} \sin r
\end{aligned} \tag{15.2.24}$$

and again

$$\begin{aligned}
\vec{E}_{1o} \times \hat{n} &= E_{1o} \cos i \\
\vec{E}'_{1o} \times \hat{n} &= -E'_{1o} \cos i \\
\vec{E}_{2o} \times \hat{n} &= E'_{2o} \cos r
\end{aligned} \tag{15.2.25}$$

Thus the first condition in (15.2.12) yields a non-trivial equation

$$n_2 \epsilon_1 (E_{1o} + E'_{1o}) - n_1 \epsilon_2 E_{2o} = 0 \tag{15.2.26}$$

which may be put in a slightly different form by noting that

$$n_1 = \frac{c}{v} \rightarrow \frac{n_1}{c} = \sqrt{\epsilon_1 \mu_1} \tag{15.2.27}$$

and likewise

$$\frac{n_2}{c} = \sqrt{\epsilon_2 \mu_2}. \tag{15.2.28}$$

First dividing by c and then by $\sqrt{\epsilon_1 \epsilon_2 \mu_1 \mu_2}$ gives

$$\sqrt{\frac{\epsilon_1}{\mu_1}} (E_{1o} + E'_{1o}) - \sqrt{\frac{\epsilon_2}{\mu_2}} E_{2o} = 0 \tag{15.2.29}$$

The second equation yields an independent condition

$$(E_{1o} - E'_{1o}) \cos i - E_{2o} \cos r = 0 \tag{15.2.30}$$

Now for the conditions on \vec{B} . As \vec{B} points out of the plane of incidence, the third equation in (15.2.12) is empty. The fourth gives

$$\frac{n_1}{\mu_1 c} (E_{1o} + E'_{1o}) - \frac{n_2}{\mu_2 c} E_{2o} = 0 \tag{15.2.31}$$

Again, using the expressions above for $n_{1,2}$, (15.2.31) reads

$$\sqrt{\frac{\epsilon_1}{\mu_1}} (E_{1o} + E'_{1o}) - \sqrt{\frac{\epsilon_2}{\mu_2}} E_{2o} = 0 \tag{15.2.32}$$

which is precisely (15.2.26), so this equation does not provide any new information. We have just two conditions (as we did when the electric field was perpendicular to the plane of incidence), *viz.*,

$$\begin{aligned} n_2 \epsilon_1 (E_{1o} + E'_{1o}) - n_1 \epsilon_2 E_{2o} &= 0 \\ (E_{1o} - E'_{1o}) \cos i - E_{2o} \cos r &= 0 \end{aligned} \quad (15.2.33)$$

from which we find, eliminating first E_{2o} , as before,

$$r_{12}^{\parallel} = \frac{E'_{1o}}{E_{1o}} = \frac{\epsilon_2 n_1 \cos i - \epsilon_1 n_2 \cos r}{\epsilon_2 n_1 \cos i + \epsilon_1 n_2 \cos r} \quad (15.2.34)$$

and then

$$t_{12}^{\parallel} = \frac{E_{2o}}{E_{1o}} = \frac{2\epsilon_1 n_2 \cos i}{\epsilon_2 n_1 \cos i + \epsilon_1 n_2 \cos r} \quad (15.2.35)$$

Eliminating $\epsilon_{1,2}$ from the above expressions by using their relationship to $n_{1,2}$ and further using Snell's law to express $\cos r$ in terms of $\sin i$, *i.e.*,

$$\cos r = \frac{1}{n_2} \sqrt{n_2^2 - n_1^2 \sin^2 i} \quad (15.2.36)$$

we find

$$\begin{aligned} r_{12}^{\parallel} &= \frac{E'_{1o}}{E_{1o}} = \frac{\frac{\mu_1}{\mu_2} n_2^2 \cos i - n_1 \sqrt{n_2^2 - n_1^2 \sin^2 i}}{\frac{\mu_1}{\mu_2} n_2^2 \cos i + n_1 \sqrt{n_2^2 - n_1^2 \sin^2 i}} \\ t_{12}^{\parallel} &= \frac{E_{2o}}{E_{1o}} = \frac{2n_1 n_2 \cos i}{\frac{\mu_1}{\mu_2} n_2^2 \cos i + n_1 \sqrt{n_2^2 - n_1^2 \sin^2 i}} \end{aligned} \quad (15.2.37)$$

The reflectance and transmittance for this case follows from their respective definitions

$$\begin{aligned} R^{\parallel} &= r_{12}^{\parallel 2} \\ T^{\parallel} &= \frac{\mu_1 n_2 \cos r}{\mu_2 n_1 \cos i} t_{12}^{\parallel 2} = \frac{\mu_1 \sqrt{n_2^2 - n_1^2 \sin^2 i}}{\mu_2 n_1 \cos i} t_{12}^{\parallel 2} \end{aligned} \quad (15.2.38)$$

Of course $R^{\parallel} + T^{\parallel} = 1$ as before.³ We have considered only linearly polarized waves and said nothing of the most general case, that of elliptical polarization. However, we know that the latter can be obtained as a linear combination of the two types of linear polarization (\perp and \parallel) that we have considered and so the Fresnel coefficients in the general case may be deduced from the above formulæ following the methods outlined before.

³Problem: Verify this.

Physical Implications

The expressions for the transmittance and reflectance obtained in the two cases treated above imply that there are two particularly interesting angles of incidence. We ask: is there any angle for which, in each case, the reflectance is (a) unity (perfect reflectance) and (b) zero (perfect transmittance).

Perfect reflectance:

If we want to have $R^\perp = 1$ then from (15.2.23)

$$(n_1 \cos i - \frac{\mu_1}{\mu_2} \sqrt{n_2^2 - n_1^2 \sin^2 i}) = \pm (n_1 \cos i + \frac{\mu_1}{\mu_2} \sqrt{n_2^2 - n_1^2 \sin^2 i}) \quad (15.2.39)$$

so two situations arise.

- Choosing the positive sign we see that the term inside the square-root must be zero, so the solution is

$$\sin i_c = \frac{n_2}{n_1} \quad (15.2.40)$$

The angle i_c is called the “critical” angle. The phenomenon is called “total internal reflection” – all of the incident energy is reflected and there is no transmission. Evidently for it to be a real angle $n_2 < n_1$, so total internal reflection occurs when the ray propagates from a medium with a higher refractive index to one with a lower refractive index (glass to air, say, or water to air or glass to water) and not vice-versa. For angles of incidence that are greater than i_c , Snell’s law would give

$$\sin r = \frac{n_1 \sin i}{n_2} > \frac{n_1 \sin i_c}{n_2} = 1 \quad (15.2.41)$$

so there is no real refracted ray (r is not a real angle as $\sin r > 1$) and the ray continues to be totally reflected off the surface.

- The second solution is one that would be obtained by taking the negative sign in (15.2.39). Then we would simply get $\cos i_g = 0$ or $i_g = \pi/2$. This is “grazing” incidence. We notice that as the incidence angle approaches grazing incidence the reflectance increases rapidly toward unity. This is why a calm lake would look very like a mirror when viewed at grazing angles.

The same conclusions are arrived at by looking at $R^\parallel = 1$. Thus both polarizations obey the same phenomena.

Perfect transmittance

To have $R^\perp = 0$ (so that $T^\perp = 1$) we require

$$n_1 \cos i = \frac{\mu_1}{\mu_2} \sqrt{n_2^2 - n_1^2 \sin^2 i} \quad (15.2.42)$$

Squaring both sides and doing a wee bit of algebra we find quite simply

$$\cos i = \frac{\mu_1}{n_1} \sqrt{\frac{n_2^2 - n_1^2}{\mu_2^2 - \mu_1^2}}, \quad \mu_1 \neq \mu_2 \quad (15.2.43)$$

and no solution if $\mu_1 = \mu_2$ *i.e.*, there is always reflectance of this polarization when the two media have the same magnetic properties. On the other hand, to have $R^\parallel = 0$ ($T^\parallel = 1$) we must satisfy the equation

$$\frac{\mu_1}{\mu_2} n_2^2 \cos i = n_1 \sqrt{n_2^2 - n_1^2 \sin^2 i}. \quad (15.2.44)$$

It gives, after a bit of algebra,

$$\cos i = \sqrt{\frac{n_1^2(n_2^2 - n_1^2)}{\frac{\mu_1^2}{\mu_2^2} n_2^4 - n_1^4}} \quad (15.2.45)$$

If the two media have the same magnetic properties ($\mu_1 = \mu_2$) then $i = i_B$ takes a particularly simple form:

$$\cos i_B = \frac{n_1}{\sqrt{n_1^2 + n_2^2}} \rightarrow \tan i_B = \frac{n_2}{n_1} \quad (15.2.46)$$

The angle i_B is called Brewster's angle. For this angle of incidence, all of the polarization in the plane of incidence will be transmitted. But we have just seen that to have $R^\perp = 0$ ($T^\perp = 1$) the two media must have different magnetic properties (*i.e.*, $\mu_1 \neq \mu_2$) or there is no solution for the incidence angle. Thus for media that have the same magnetic permeability, only the perpendicular polarization is reflected at the incident angle i_B , *i.e.*, the reflected ray is polarized perpendicular to the plane of incidence in this particular case.

15.3 Interfaces: non-conducting/conducting medium

The fundamental difference between a non-conducting medium and a conducting medium is in the dispersion relation: the wave-vector \vec{k} in a conducting medium must be complex whereas it can be chosen to be real in a conducting medium. This has certain consequences some of which we shall now examine.

Kinematics

Let medium “1” be a non-conducting medium and medium “2” be conducting. Consider the condition of no spatial and temporal variation of the phase at the incidence point. It implies that the frequencies in both regions are the same and further that

$$\vec{\kappa}_1 \cdot \vec{r} = \kappa_1' \cdot \vec{r} = \vec{\kappa}_2 \cdot \vec{r} = \kappa_2^R \cdot \vec{r} + i\kappa_2^I \cdot r \quad (15.3.1)$$

By the same arguments we have applied before, we could re-write this equation as

$$\vec{\kappa}_1 \times \hat{n} = \kappa_1' \times \hat{n} = \kappa_2^R \times \hat{n} + i\kappa_2^I \times \hat{n} \quad (15.3.2)$$

But, because $\vec{\kappa}_1$ and κ_1' are both real (medium “1” is non-conducting) it follows that

$$\kappa_2^I \times \hat{n} = 0 \quad (15.3.3)$$

and, assuming that $\kappa_2^I \neq 0$, it follows that κ_2^I must point in the direction of the normal to the surface. The amplitude of the wave in the conducting medium therefore decays as

$$\exp[-\kappa_2^I \hat{n} \cdot \vec{r}] \quad (15.3.4)$$

so that the skin depth (*perpendicular to the interface*) is simply

$$\delta = \frac{1}{\kappa_2^I}. \quad (15.3.5)$$

Furthermore we derive the same laws of reflection and refraction: (a) the incident ray, the refracted ray, the reflected ray and the normal are in the same plane (same argument as before) and (b) the sine law

$$\kappa_1 \sin i = \kappa_1' \sin r' = \kappa_2 \sin r \quad (15.3.6)$$

which implies that $i = r'$ (because $\kappa_1 = \kappa_1'$) and

$$\frac{\sin i}{\sin r} = \frac{\kappa_2^R}{\kappa_1} \quad (15.3.7)$$

These expressions simplify considerably for normal incidence ($i = 0$). The last equation says that $r = 0$ and therefore $\vec{\kappa}_2 = (\kappa_2^R + \kappa_2^I)\hat{n}$. Because, in this particular case (only), $n_2^{R,I} = c\kappa_2^{R,I}/\omega$, the skin depth can be written as

$$\delta = \frac{n^R}{n^I} \frac{\lambda}{2\pi} \quad (15.3.8)$$

precisely as we had before.

Let us briefly consider the case of oblique incidence. Treating this case is algebraically more complicated than it is for non-conducting media, but it is conceptually the same. Fundamentally, we would like to express Snell's law in terms of the real and imaginary parts of the refractive index as we had it in section (14). However, because the real and imaginary parts of $\vec{\kappa}_2$ now point in different directions, the relationship between them and $n^{R,I}$ is considerably more complex. Recall the relations (g is the conductivity of medium "2")

$$n_2 = \frac{c\kappa_2}{\omega} = \frac{c}{\omega} \sqrt{\frac{\omega^2}{v_2^2} + i\omega\mu_2g} \quad (15.3.9)$$

and

$$\vec{\kappa}_2^2 = (\vec{\kappa}_2^R + i\vec{\kappa}_2^I) \cdot (\vec{\kappa}_2^R + i\vec{\kappa}_2^I) = \kappa_2^{R2} - \kappa_2^{I2} + 2i\vec{\kappa}_2^R \cdot \vec{\kappa}_2^I = \frac{\omega^2}{v_2^2} + i\omega\mu_2g \quad (15.3.10)$$

The first of these led to our expressions for the real and imaginary components of the refractive index of the conducting medium, *viz.*,

$$n^{R,I} = c\sqrt{\frac{\mu_2}{2}} \sqrt{\pm\epsilon_2 + \sqrt{\epsilon_2^2 + \frac{g^2}{\omega^2}}} \quad (15.3.11)$$

We could similarly use the second to obtain the real and imaginary parts of $\vec{\kappa}_2$ in terms of the conducting medium's fundamental parameters, (ϵ_2, μ_2, g) . However, this would implicitly involve the incident angle via the imaginary part:

$$\begin{aligned} \kappa_2^{R2} - \kappa_2^{I2} &= \frac{\omega^2}{v_2^2} \\ 2\vec{\kappa}_2^R \cdot \vec{\kappa}_2^I &= 2\kappa_2^R \kappa_2^I \cos r = 2\kappa_2^I \sqrt{\kappa_2^{R2} - \kappa_1^2 \sin^2 i} = \omega\mu_2g \end{aligned} \quad (15.3.12)$$

where we have used the fact that $\hat{\kappa}_2^I = \hat{n}$ and

$$\sin r = \frac{\kappa_1}{\kappa_2^R} \sin i \rightarrow \cos r = \frac{1}{\kappa_2^R} \sqrt{(\kappa_2^R)^2 - \kappa_1^2 \sin^2 i} \quad (15.3.13)$$

From the second equation in (15.3.12) we find κ_2^I in terms of the other variables ($v_2 = 1/\sqrt{\epsilon_2\mu_2}$):

$$\kappa_2^I = \frac{\omega\mu_2g}{2\sqrt{\kappa_2^{R2} - \kappa_1^2 \sin^2 i}} \quad (15.3.14)$$

When this is inserted into the first equation we end up with an equation for κ_2^R ,

$$\kappa_2^{R2} - \frac{\omega^2\mu_2^2g^2}{4(\kappa_2^{R2} - \kappa_1^2 \sin^2 i)} = \epsilon_2\mu_2\omega^2 \quad (15.3.15)$$

which may be solved to recover κ_2^R in terms of (ϵ_2, μ_2, g) and the angle of incidence, i . It gives a quartic equation in κ_2^R (more precisely, a quadratic equation in κ_2^{R2})

$$\kappa_2^{R4} - \kappa_2^{R2} [\kappa_1^2 \sin^2 i + \epsilon_2 \mu_2 \omega^2] - \frac{\omega^2}{4} [g^2 \mu_2^2 - 4\epsilon_2 \mu_2 \kappa_1^2 \sin^2 i] = 0 \quad (15.3.16)$$

Thus we find

$$\kappa_2^R = \sqrt{\frac{1}{2} [\epsilon_2 \mu_2 \omega^2 + \kappa_1^2 \sin^2 i] + \frac{1}{2} \sqrt{[\epsilon_2 \mu_2 \omega^2 + \kappa_1^2 \sin^2 i]^2 + \omega^2 [g^2 \mu_2^2 - 4\epsilon_2 \mu_2 \kappa_1^2 \sin^2 i]}} \quad (15.3.17)$$

where we have chosen the solution with the positive sign to ensure the positivity of κ_2^R . Knowing κ_2^R we could easily extract κ_2^I using (15.3.12). This complicated expression may be slightly simplified by replacing

$$\kappa_1 = \frac{\omega}{v_1} = \omega \sqrt{\epsilon_1 \mu_1} \quad (15.3.18)$$

(κ_1 is the wave-number in medium “1” which is non-conducting). We find the following final expression:⁴

$$\kappa_2^R = \omega \sqrt{\frac{1}{2} \sqrt{(\epsilon_2 \mu_2 + \epsilon_1 \mu_1 \sin^2 i) + \sqrt{(\epsilon_2 \mu_2 - \epsilon_1 \mu_1 \sin^2 i)^2 + \frac{g^2 \mu_2^2}{\omega^2}}}} \quad (15.3.19)$$

(We could, of course, also express κ_2^R in terms of $n^{R,I}$, using $n_2 = c\kappa_2/\omega$, although this will not eliminate the dependence of the result on the incident angle.) One sees that Snell’s law is deceptively simple looking: while it reads

$$\frac{\sin i}{\sin r} = \frac{\kappa_2^R}{\kappa_1} \quad (15.3.20)$$

we have seen that κ_2^R depends on $\sin i$. Thus the “angle of refraction”, r , is really quite a complicated function of the angle of incidence. However suppose that the conducting medium approaches a perfect conductor ($g \rightarrow \infty$) then κ_2^R is approximately independent of the incident angle, being dominated by the term containing g ,

$$\kappa_2^R \approx \sqrt{\frac{g\mu_2\omega}{2}} \quad (15.3.21)$$

⁴Problem: To verify that this complicated expression makes sense, check the two limits: $i = 0$ and $g = 0$. In the first limit we know that the real and imaginary parts of $\vec{\kappa}_2$ point in the same direction *i.e.*, the direction of the normal, \hat{n} , to the interface. Thus $\vec{\kappa}_2 = (\kappa_2^R + \kappa_2^I)\hat{n}$ and therefore, in particular, $n_2^R = c\kappa_2^R/\omega$. But we know the general result (see section (14)) for n_2^R . Make sure that the $i = 0$ limit yields precisely this expression. The second limit should yield a non-conductor. Check that this is indeed so.

but as $g \rightarrow \infty$, Snell's law gives $r \approx 0$. Thus, for *every* angle of incidence the propagation in the conductor is almost directly into it, *i.e.*, almost in the direction of \hat{n} . In this case, the expression for the skin-depth, derived before for normal incidence, holds for *any* angle of incidence. However, the attenuation in the conducting medium in this case is extremely large (κ_2^I is then proportional to g) and the skin-depth is very small.

Can $r = \frac{\pi}{2}$? This would mean that there is some angle of incidence for which the refracted ray is grazing. When $r = \frac{\pi}{2}$ we have, by Snell's law

$$\sin i = \sqrt{\frac{1}{2\epsilon_1\mu_1}} \sqrt{(\epsilon_2\mu_2 + \epsilon_1\mu_1 \sin^2 i) + \sqrt{(\epsilon_2\mu_2 - \epsilon_1\mu_1 \sin^2 i)^2 + \frac{g^2\mu_2^2}{\omega^2}}} \quad (15.3.22)$$

By squaring both sides this equation can be simplified to read

$$(\epsilon_2\mu_2 - \epsilon_1\mu_1 \sin^2 i) = -\sqrt{(\epsilon_2\mu_2 - \epsilon_1\mu_1 \sin^2 i)^2 + \frac{g^2\mu_2^2}{\omega^2}} \quad (15.3.23)$$

The equation has no solution unless $g = 0$, in which case we return to the results for two non-conducting media. Alternatively, we could have obtained this directly from (15.3.12) for then $\vec{\kappa}_2^R \cdot \vec{\kappa}_2^I = 0$ which implies that $g = 0$.

The dynamics are quite difficult to describe, owing to the fact that κ_2 is complex. This complicates the matching conditions and we will not consider the dynamics in this course.

Chapter 16

Wave Guides and Resonant Cavities

16.1 Introduction

We will now consider another situation of practical importance: the propagation of electromagnetic waves inside conducting shells that are either hollow or filled with a non-conducting medium. If the ends of the conducting shell are sealed, also by conducting surfaces, then the objects are called Cavities, otherwise they are called Wave Guides. We consider Wave Guides first (see figure 17): they are used in the transmission of electromagnetic power.

Our equation in the dielectric medium inside the hollow shell (this could also be a

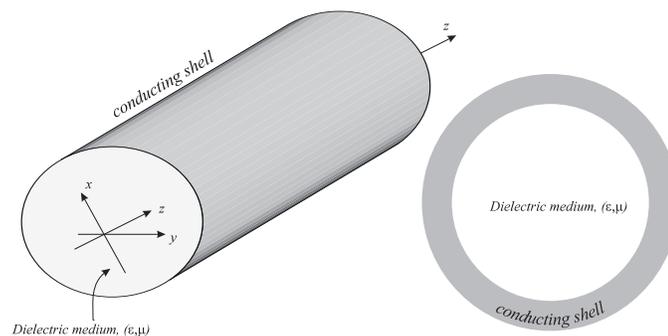


Figure 16.1: A cylindrical Wave Guide

vacuum) are

$$\left[\frac{1}{v^2} \frac{\partial^2}{\partial t^2} - \vec{\nabla}^2 \right] \begin{Bmatrix} \phi \\ \vec{A} \end{Bmatrix} = 0 \quad (16.1.1)$$

and we want solutions that respect the symmetry of the problem. Let z be the (open) axis of the shell, so that (x, y) are the transverse directions as shown in the figure. We look for harmonic solutions.

16.2 Wave Guides

As there are no end surfaces, the z -direction is unbounded and we expect there to exist traveling waves in this direction. Let us then search for solutions that possess the following z -dependence

$$\begin{Bmatrix} \phi \\ \vec{A} \end{Bmatrix} = \begin{Bmatrix} \phi(x, y) \\ \vec{A}(x, y) \end{Bmatrix} e^{i(\kappa z - \omega t)} \quad (16.2.1)$$

Of course, this implies that the solution for \vec{E} and \vec{B} will have the same form as for the potentials,

$$\begin{Bmatrix} \vec{E} \\ \vec{B} \end{Bmatrix} = \begin{Bmatrix} \vec{E}(x, y) \\ \vec{B}(x, y) \end{Bmatrix} e^{i(\kappa z - \omega t)} \quad (16.2.2)$$

and, in the absence of sources, they also obey the same equations of motion, *i.e.*,

$$\left[\frac{1}{v^2} \frac{\partial^2}{\partial t^2} - \vec{\nabla}^2 \right] \begin{Bmatrix} \vec{E} \\ \vec{B} \end{Bmatrix} = 0 \quad (16.2.3)$$

We do not expect the solutions to have the simple form of a plane wave in the transverse direction because of the presence of the conducting boundary. The symmetry is such that it is convenient to separate the gradient operator into longitudinal (along z) and transverse (in the (x, y) plane) parts. Let \hat{e}_3 point in the z -direction and write

$$\vec{\nabla} = \hat{e}_3 \partial_z + \vec{\nabla}_t. \quad (16.2.4)$$

Then, we may write the three dimensional Laplace operator as $\vec{\nabla}^2 = \partial_z^2 + \vec{\nabla}_t^2$ and our equation for the electric and the magnetic fields can be put in the form

$$\left[\frac{1}{v^2} \frac{\partial^2}{\partial t^2} - \vec{\nabla}_t^2 - \frac{\partial^2}{\partial z^2} \right] \begin{Bmatrix} \vec{E} \\ \vec{B} \end{Bmatrix} = 0, \quad (16.2.5)$$

whereupon, using the dependence on (t, z) that we have imposed (this dependence implies that $\partial_t \rightarrow -i\omega$ and $\partial_z \rightarrow i\kappa$), we find

$$\left[\vec{\nabla}_t^2 + \left(\frac{\omega^2}{v^2} - \kappa^2 \right) \right] \begin{Bmatrix} \vec{E}(x, y) \\ \vec{B}(x, y) \end{Bmatrix} = 0, \quad (16.2.6)$$

Let us also separate \vec{E} (and \vec{B}) into longitudinal and transverse parts:

$$\vec{E} = E_z \hat{e}_3 + \vec{E}_t, \quad \vec{B} = B_z \hat{e}_3 + \vec{B}_t \quad (16.2.7)$$

where

$$\vec{E}_t = \hat{e}_3 \times (\vec{E} \times \hat{e}_3), \quad \vec{B}_t = \hat{e}_3 \times (\vec{B} \times \hat{e}_3) \quad (16.2.8)$$

and consider Maxwell's equations, written explicitly in terms of the longitudinal and transverse components. The two scalar equations are quite trivial to expand

$$\begin{aligned} \vec{\nabla} \cdot \vec{D} = \rho = 0 &\rightarrow \partial_z E_z + \vec{\nabla}_t \cdot \vec{E}_t = 0 \\ \vec{\nabla} \cdot \vec{B} = 0 &\rightarrow \partial_z B_z + \vec{\nabla}_t \cdot \vec{B}_t = 0. \end{aligned} \quad (16.2.9)$$

Separating the vector equations requires a bit more work, but it can be done by projecting along the z axis and transverse to it as we show now. Take the Bianchi identity

$$\vec{\nabla} \times \vec{E} + \frac{\partial \vec{B}}{\partial t} = 0 \quad (16.2.10)$$

and consider its decomposition that is obtained first by taking the scalar product with \hat{e}_3 and then by taking the vector product with the same. We can write the two equations

$$\begin{aligned} \hat{e}_3 \cdot \left(\vec{\nabla} \times \vec{E} + \frac{\partial \vec{B}}{\partial t} \right) &= 0, & \text{longitudinal part} \\ \hat{e}_3 \times \left(\vec{\nabla} \times \vec{E} + \frac{\partial \vec{B}}{\partial t} \right) &= 0 & \text{transverse part} \end{aligned} \quad (16.2.11)$$

Now

$$\hat{e}_3 \cdot (\vec{\nabla} \times \vec{E}) = \epsilon^{ijk} \hat{e}_{3i} \partial_j E_k, \quad (16.2.12)$$

and because ϵ^{ijk} is totally antisymmetric in its indices while \hat{e}_3 points in the direction of z , we find that the indices j, k can only be transverse: the longitudinal part of the equation thus reads

$$\hat{e}_3 \cdot (\vec{\nabla}_t \times \vec{E}_t) + \frac{\partial B_z}{\partial t} = 0 \quad (16.2.13)$$

The transverse part may be simplified if we note that¹

$$\hat{e}_3 \times (\vec{\nabla} \times \vec{E}) = \vec{\nabla}_t E_z - \frac{\partial \vec{E}_t}{\partial z} \quad (16.2.14)$$

¹Problem: Convince yourself that this is true. Use the necessary vector identities, etc.

which gives

$$\vec{\nabla}_t E_z - \frac{\partial \vec{E}_t}{\partial z} + \hat{e}_3 \times \frac{\partial \vec{B}}{\partial t} = 0 \quad (16.2.15)$$

This Bianchi identity therefore yields the two equations. Likewise the same decomposition of the last of Maxwell's equations,

$$\vec{\nabla} \times \vec{H} - \frac{\partial \vec{D}}{\partial t} = \vec{j} = 0 \quad (16.2.16)$$

yields (by duality)

$$\begin{aligned} \hat{e}_3 \cdot (\vec{\nabla}_t \times \vec{B}_t) - \frac{1}{v^2} \frac{\partial E_z}{\partial t} &= 0 \\ \vec{\nabla}_t B_z - \frac{\partial \vec{B}_t}{\partial z} - \frac{1}{v^2} \hat{e}_3 \times \frac{\partial \vec{E}_t}{\partial t} &= 0 \end{aligned} \quad (16.2.17)$$

The six equations above constitute Maxwell's equations decomposed into their longitudinal and transverse parts. Using $\partial_t \rightarrow -i\omega$ and $\partial_z \rightarrow i\kappa$, and cancelling a common phase factor, we may summarize them as follows for future use

$$\begin{aligned} \vec{\nabla}_t \cdot E_t(x, y) &= -i\kappa E_z(x, y) \\ \vec{\nabla}_t \cdot B_t(x, y) &= -i\kappa B_z(x, y) \\ \hat{e}_3 \cdot (\vec{\nabla}_t \times \vec{E}_t(x, y)) &= i\omega B_z(x, y) \\ \vec{\nabla}_t E_z(x, y) &= i\kappa \vec{E}_t(x, y) + i\omega \hat{e}_3 \times \vec{B}_t(x, y) \\ \hat{e}_3 \cdot (\vec{\nabla}_t \times \vec{B}_t(x, y)) &= -i \frac{\omega}{v^2} E_z(x, y) \\ \vec{\nabla}_t B_z(x, y) &= i\kappa \vec{B}_t(x, y) - i \frac{\omega}{v^2} \hat{e}_3 \times \vec{E}_t(x, y) \end{aligned} \quad (16.2.18)$$

These equations must obviously be supplemented with appropriate boundary conditions, namely that the tangential component of the electric field and the normal component of the magnetic field vanish at the perfectly conducting boundary. If \hat{n} is the unit normal to the surface, then

$$\hat{n} \times \vec{E} \Big|_C = 0, \quad \hat{n} \cdot \vec{B} \Big|_C = 0 \quad (16.2.19)$$

is the correct mathematical way to express these conditions, where C represents the (perfectly) conducting surface. Let us now consider some special solutions.

Transverse Electromagnetic (TEM) Waves

One particular solution is obtained by letting $E_z = 0 = B_z$. From the third equation in (16.2.18) and the first equation in (16.2.18) it then follows that

$$\vec{\nabla}_t \times \vec{E}_t(x, y) = 0, \quad \vec{\nabla}_t \cdot \vec{E}_t(x, y) = 0 \quad (16.2.20)$$

i.e., the electric field is a solution of the *electrostatic* problem. It follows that

$$\vec{\nabla}_t \times (\vec{\nabla}_t \times \vec{E}(x, y)) = 0 = \vec{\nabla}_t (\vec{\nabla}_t \cdot \vec{E}_t(x, y)) - \vec{\nabla}_t^2 \vec{E}(x, y) = -\vec{\nabla}_t^2 \vec{E}(x, y) \quad (16.2.21)$$

and therefore, from (16.2.6) that $\kappa = \omega/v$. Furthermore, the sixth equation in (16.2.18) then gives

$$\vec{B}_t(x, y) = \frac{1}{v} \hat{e}_3 \times \vec{E}_t(x, y) \quad (16.2.22)$$

Notice that the connection between \vec{E}_t and \vec{B}_t is just the same as for a plane wave in an infinite medium. Such a solution is called a "Transverse Electro-Magnetic" or TEM wave. We only have to find a solution to the corresponding electrostatic problem in two dimensions. The fact that $\vec{E}_t(x, y)$ is irrotational implies that $\vec{E}_t(x, y) = -\vec{\nabla}_t \Phi(x, y)$ (Φ is the electrostatic potential) and so we ask for a solution of $\vec{\nabla}_t^2 \Phi = 0$.

As an example, consider the cylindrical wave guide shown in figure 17. The general solution to the electrostatic problem with cylindrical symmetry was obtained in the first part of this subject: in terms of the radial coordinate, ρ , and the azimuthal angle, φ ,

$$\begin{aligned} \rho &= \sqrt{x^2 + y^2} \\ \varphi &= \tan^{-1} \left(\frac{y}{x} \right) \end{aligned} \quad (16.2.23)$$

it is

$$\Phi(\rho, \varphi) = a_o + b_o \ln \left(\frac{\rho}{\rho_o} \right) + \sum_{n=1}^{\infty} \rho^n (a_n \cos n\varphi + b_n \sin n\varphi) + \sum_{n=1}^{\infty} \rho^{-n} (c_n \cos n\varphi + d_n \sin n\varphi) \quad (16.2.24)$$

If the boundary is a perfect cylindrical conductor of radius R then it cannot support an electric field, which means that Φ should be constant at $\rho = R$. This is only possible if Φ is independent of φ , *i.e.*, if all the coefficients a_n , b_n , c_n and d_n vanish for $n \geq 1$. We're left with

$$\Phi(\rho) = a_o + b_o \ln \left(\frac{\rho}{\rho_o} \right), \quad (16.2.25)$$

but this potential is typical of a line charge if $b_o \neq 0$. As there is no line charge in this problem we should also set $b_o = 0$ which means that Φ is a constant *everywhere*, whose value is equal to its value on the conducting boundary! Thus a TEM wave cannot exist in a single, perfectly conducting, cylindrical wave guide. Of course we could have two

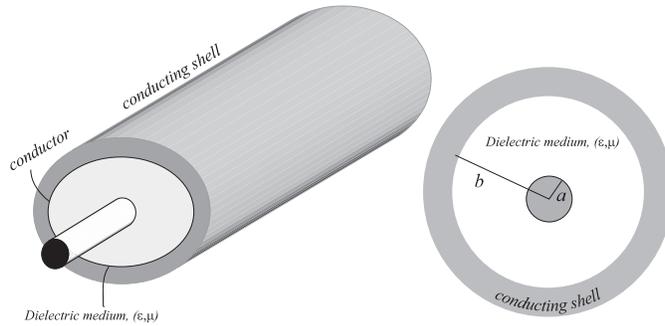


Figure 16.2: A co-axial cable can support TEM waves

or more cylindrical surfaces as, for example, in figure 18. In that case TEM waves are supported.²

Transverse Electric (TE) Waves

A more general solution is obtained by requiring only that $E_z = 0$. This yields the “Transverse Electric” (TE) wave. If $E_z = 0$, the fourth equation in (16.2.18) implies that

$$\vec{E}_t(x, y) = -\frac{\omega}{\kappa} \hat{e}_3 \times \vec{B}_t(x, y) \rightarrow \hat{e}_3 \times \vec{E}_t(x, y) = \frac{\omega}{\kappa} \vec{B}_t(x, y) \quad (16.2.26)$$

and the fifth equation in (16.2.18) gives

$$\vec{\nabla}_t \times \vec{B}_t(x, y) = 0 \rightarrow \vec{B}_t(x, y) = -\vec{\nabla}_t \psi(x, y) \quad (16.2.27)$$

Thus, from the second equation we find

$$\vec{\nabla}_t^2 \psi(x, y) = i\kappa B_z(x, y) \quad (16.2.28)$$

Again, using (16.2.26) in the sixth equation in (16.2.18) gives

$$\begin{aligned} \vec{\nabla}_t B_z(x, y) &= i \left(\kappa - \frac{\omega^2}{\kappa v^2} \right) \vec{B}_t(x, y) = -i \left(\kappa - \frac{\omega^2}{\kappa v^2} \right) \vec{\nabla}_t \psi(x, y) \\ &\rightarrow B_z = -i \left(\kappa - \frac{\omega^2}{\kappa v^2} \right) \psi(x, y) \end{aligned} \quad (16.2.29)$$

which finally implies that ψ obeys the equation

$$(\vec{\nabla}_t^2 + \gamma^2) \psi(x, y) = 0, \quad \gamma^2 = \frac{\omega^2}{v^2} - \kappa^2 \quad (16.2.30)$$

²Problem: Find the TEM fields in this case. Assume the situation depicted in figure 18 and that the potentials on the inner and outer conductors are respectively Φ_A and Φ_B .

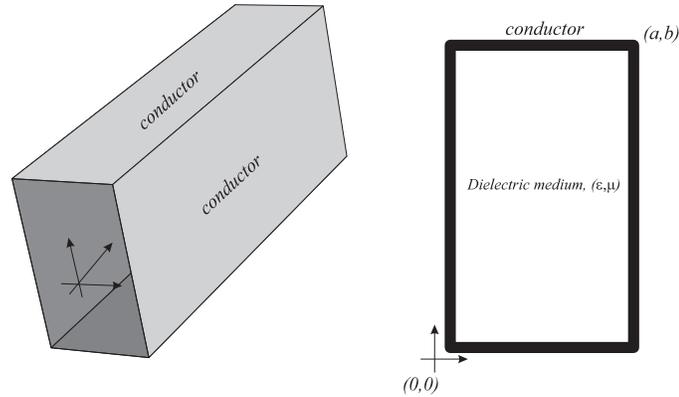


Figure 16.3: A rectangular Wave Guide

The boundary conditions are obviously given by (16.2.19), but how do they translate into conditions on ψ for the case of TE waves in the wave guide? Notice that $E_z = 0$ so the first condition ($\hat{n} \times \vec{E}|_C = 0$) is empty. The second condition reads

$$\hat{n} \cdot \vec{B}|_C = 0 \rightarrow \hat{n} \cdot \vec{\nabla}_t \psi(x, y)|_C := \frac{\partial \psi}{\partial n}|_C = 0 \quad (16.2.31)$$

or, in words, the normal derivative (directional derivative normal to the surface) of ψ should vanish on the boundary. Together the two equations, *viz.*, (16.2.30) and (16.2.31), completely and uniquely define an eigenvalue problem describing the system.

This time let us apply the above considerations to the case of a rectangular wave-guide as shown in figure 19. In this case, we consider the equation

$$\left[\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \gamma^2 \right] \psi(x, y) = 0 \quad (16.2.32)$$

and separate variables, *i.e.*, let $\psi(x, y) = \psi_1(x)\psi_2(y)$. This procedure results in two equations:

$$\begin{aligned} \frac{\partial^2 \psi_1}{\partial x^2} &= -\lambda^2 \psi_1 \\ \frac{\partial^2 \psi_2}{\partial y^2} &= -(\gamma^2 - \lambda^2) \psi_2 \end{aligned} \quad (16.2.33)$$

where λ^2 is an arbitrary constant, which we take, to be positive to ensure oscillatory solutions. The general solution is therefore

$$\psi(x, y) = \psi_o \cos(\lambda x + \phi_1) \cos(\sqrt{\gamma^2 - \lambda^2} y + \phi_2) \quad (16.2.34)$$

where $\phi_{1,2}$ are arbitrary constants, but we must satisfy the boundary conditions. The normal to the surfaces of the guide are $\hat{n} = (\mp 1, 0, 0)$ for the surfaces $x = 0, a$ (see figure 19) and $\hat{n} = (0, \mp 1, 0)$ for the surfaces $y = 0, b$. Thus, we must have

$$\left. \frac{\partial \psi_1}{\partial x} \right|_{x=0,a} = 0, \quad \left. \frac{\partial \psi_2}{\partial y} \right|_{y=0,b} = 0 \quad (16.2.35)$$

It is easy to see that these conditions imply that $\phi_1 = \phi_2 = 0$ and

$$\lambda = \frac{n\pi}{a}, \quad \gamma = \pi \left[\frac{n^2}{a^2} + \frac{m^2}{b^2} \right]^{\frac{1}{2}} \quad (16.2.36)$$

for arbitrary integers, n, m . The solution for $\psi(x, y)$ is therefore

$$\psi(x, y) = \psi_o \cos\left(\frac{n\pi x}{a}\right) \cos\left(\frac{m\pi y}{b}\right) \quad (16.2.37)$$

Let us observe that they are standing waves in the transverse directions, with wave-numbers, respectively $\kappa_x = n\pi/a$ and $\kappa_y = m\pi/b$. That the solution should have this form in the transverse directions should be intuitively obvious. It depends on two integers. Every pair of integers (n, m) is referred to as a mode. If we define the wave-vector $\vec{\kappa} = (\kappa_x, \kappa_y, \kappa)$, then the dispersion relation reads

$$\frac{\omega^2}{v^2} - \vec{\kappa}^2 = 0. \quad (16.2.38)$$

If we write out the solution for κ (the component describing propagation in the z - direction):

$$\kappa^2 = \frac{\omega^2}{v^2} - \pi^2 \left[\frac{n^2}{a^2} + \frac{m^2}{b^2} \right] \quad (16.2.39)$$

then we find that frequencies that yield a *real* value of κ must obey the condition

$$\omega_{nm} \geq \frac{\pi}{\sqrt{\mu\epsilon}} \left[\frac{n^2}{a^2} + \frac{m^2}{b^2} \right]^{\frac{1}{2}} \quad (16.2.40)$$

If, for example, $a > b$ then the *minimum* frequency allowed is given by the mode $n = 1, m = 0$ (the TE_{1,0} mode), or

$$\omega_{10} = \frac{\pi}{a\sqrt{\mu\epsilon}} \quad (16.2.41)$$

Below this frequency, κ is imaginary and the wave is attenuated in the z - direction. The problem of TE waves in a rectangular wave guide is thus completely solved in principle.³

³Problem: Solve the problem of TE waves in a cylindrical wave guide and determine the modes.

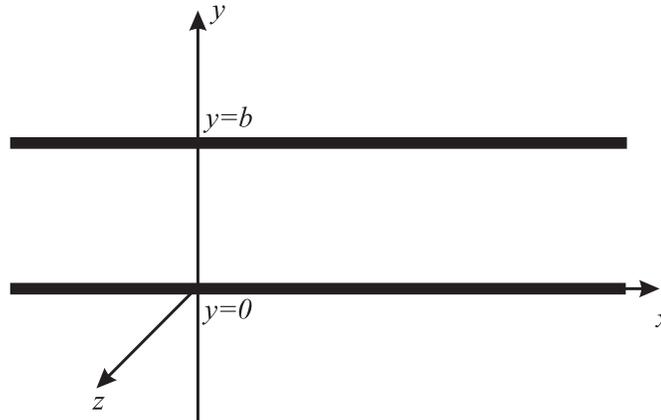


Figure 16.4: Infinite, parallel conducting planes

Transverse Magnetic (TM) waves

Another situation is obtained by letting $B_z = 0$. These yield the Transverse Magnetic (TM) waves. It should be clear, however, that the dynamical equations will have the same form (by duality, of course, because these are “vacuum” solutions!). Indeed, all that is required is the interchange $\vec{E} \rightarrow \vec{B}$ and $\vec{B} \rightarrow -\vec{E}/v^2$. However, the boundary conditions are different and this must be taken into account.⁴ The solution, subject to the boundary conditions appropriate to the rectangular wave guide treated in the previous section is given by

$$\psi(x, y) = \psi_o \sin\left(\frac{n\pi x}{a}\right) \sin\left(\frac{m\pi y}{b}\right) \quad (16.2.42)$$

where γ is still given by (16.2.36). The lowest (non-vanishing) mode is given by $n = 1 = m$ (the $\text{TM}_{1,1}$ mode), so its cut-off frequency is greater than that of the $\text{TE}_{1,0}$ mode.⁵

Let us consider in a bit more detail the problem of TM waves between two infinite, parallel conducting planes situated, say, at $y = 0, b$ (see figure 20). Let the waves propagate only in the $y - z$ plane, for convenience (one can always orient the axis so that this is so). By arrangement, then, we seek solutions representing traveling waves in the z - direction:

$$\vec{E} = \vec{E}(x, y)e^{i(\kappa z - \omega t)}$$

⁴Problem: Show that the appropriate boundary condition to be applied to $\psi(x, y)$ (in this case, $\vec{E}_t(x, y) = -\vec{\nabla}_t \psi(x, y)$) for TM waves is $\psi|_C = 0$.

⁵Problem: Determine the factor by which the cut-off frequency of the $\text{TM}_{1,1}$ mode is greater than that of the $\text{TE}_{1,0}$ mode.

$$\vec{B} = \vec{B}(x, y)e^{i(\kappa z - \omega t)} \quad (16.2.43)$$

and having $B_z = 0$. Consider the set of equations in (16.2.18). From the third we conclude that

$$\vec{\nabla} \times \vec{E}_t(x, y) = 0 \rightarrow \vec{E}_t(x, y) = -\vec{\nabla}_t \psi(x, y) \quad (16.2.44)$$

and therefore the first gives

$$\vec{\nabla}_t^2 \psi(x, y) = i\kappa E_z(x, y) \quad (16.2.45)$$

Now the last equation in (16.2.18) implies that

$$\vec{B}_t(x, y) = \frac{\omega}{\kappa v^2} \hat{e}_3 \times \vec{E}_t(x, y) \rightarrow \hat{e}_3 \times \vec{B}_t = -\frac{\omega}{\kappa v^2} \vec{E}_t(x, y) \quad (16.2.46)$$

which, when inserted into the fourth equation gives

$$\vec{\nabla}_t E_z(x, y) = -i\kappa \vec{\nabla}_t \psi + \frac{i\omega^2}{\kappa v^2} \vec{\nabla}_t \psi \quad (16.2.47)$$

or

$$E_z(x, y) = -i \left(\kappa - \frac{\omega^2}{\kappa v^2} \right) \psi(x, y) \quad (16.2.48)$$

This means that the equation for $\psi(x, y)$ is simply

$$(\vec{\nabla}_t^2 + \gamma^2)\psi(x, y) = 0, \quad \gamma^2 = \frac{\omega^2}{v^2} - \kappa^2 \quad (16.2.49)$$

Knowing $\psi(x, y)$ we determine \vec{E} and knowing \vec{E} we determine \vec{B} . We see, as promised earlier, that the dynamical equations are related to those for TE waves by duality. But what of the boundary conditions? We must require $\hat{n} \times \vec{E}|_C = 0$ or simply that $\psi|_C = 0$. Also (and this is important) we cannot allow traveling waves or standing waves in the x -direction.⁶ The general solution (before applying the boundary conditions) is then just

$$\psi = \psi_o \cos(\gamma y + \phi) \quad (16.2.50)$$

where ϕ is an arbitrary phase. Applying the boundary condition $\psi|_{0,b} = 0$ then gives $\phi = \frac{\pi}{2}$ and $\gamma = n\pi/b$. The solution turns out to be

$$\psi(y) = \psi_o \sin\left(\frac{n\pi y}{b}\right) \quad (16.2.51)$$

from which we may extract the electric and magnetic fields by taking suitable derivatives and using (16.2.46).⁷ The condition

$$\gamma^2 = \frac{\omega^2}{v^2} - \kappa^2 = \frac{n^2\pi^2}{b^2} \rightarrow \kappa^2 = \frac{\omega^2}{v^2} - \frac{n^2\pi^2}{b^2} \quad (16.2.52)$$

⁶Question: Why?

⁷Problem: Do it!

implies, again, that there is a minimum frequency for traveling waves (in the z - direction):

$$\omega_1 = \frac{\pi}{b\sqrt{\epsilon\mu}} \quad (16.2.53)$$

below which κ is imaginary.

16.3 Resonant Cavities

We now turn to another problem, one that is closely related to wave guides, that of resonant cavities (or cavity resonators). Resonant cavities consist of a finite region of space filled with a non-conducting medium and bounded by a conducting medium. They are used in the storage of electromagnetic energy. In practice one never has perfect (lossless) dielectric media or perfect conductors, although this is the case we will consider here. Thus there are practical losses of energy, but these are small compared with other means of storage as, for example, the standard LC circuit.

In its ideal form the cavity resonator will be made up of a lossless dielectric bounded by a perfect conductor. As the region is finite, it does not necessarily have an axial symmetry that can be exploited and it will not be convenient to use the decomposition of Maxwell's equations into "longitudinal" and transverse components as was done in the treatment of wave guides. Instead we will use the equations given in (16.2.3) for the electric and magnetic fields. If we want harmonic solutions, then let

$$\vec{E} = \vec{E}(\vec{r})e^{-i\omega t}, \quad \vec{B} = \vec{B}(\vec{r})e^{-i\omega t} \quad (16.3.1)$$

Our dynamical equation, for example for the electric field, becomes

$$(\vec{\nabla}^2 + \omega^2)\vec{E}(\vec{r}) = 0. \quad (16.3.2)$$

We will consider a special kind of resonant cavity: one that has the shape of a rectangular parallelepiped of sides (a, b, c) as shown in figure 21. Other geometries are possible of course, but in practice the rectangular parallelepiped and the cylindrical cavity are most common.⁸ The geometry of the rectangular parallelepiped implies that we can work in Cartesian coordinates. Thus we have

$$\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} + \omega^2 \right) \vec{E}(x, y, z) = 0 \quad (16.3.3)$$

⁸We will not be treating the cylindrical resonant cavity in this course for a good reason: the coordinate system that respects cylindrical symmetry are the cylindrical coordinates, but we have not learned how to apply the Laplacian on vector fields in curvilinear coordinates! This is not as simple as applying the Laplacian on a scalar field, because if curvilinear coordinates are used one has to account for the basis vectors that appear in the expansion (eg. $\vec{E} = E_i \hat{e}_i$) of the vector field. These are not rigid, as they are for Cartesian coordinates, but change from point to point.

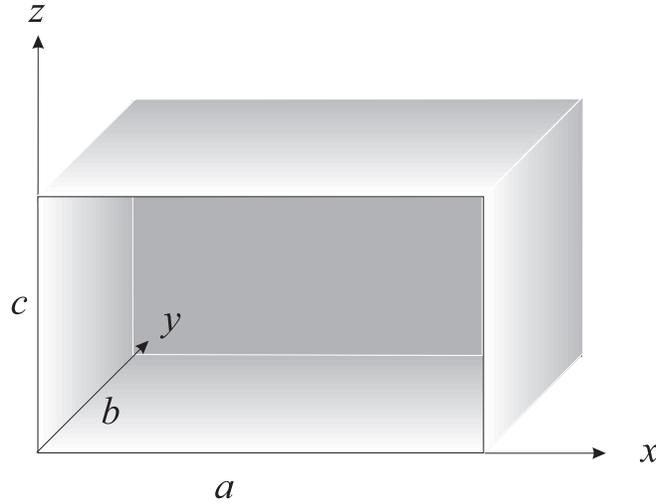


Figure 16.5: A Rectangular Resonant Cavity

and, for each component of \vec{E} we may solve the equation by the separation of variables. This procedure is, by now familiar to all. For each component of \vec{E} , we will end up with three equations, one each in x, y and z , two arbitrary constants which we choose in such a way as to have standing waves and three arbitrary phases, all of which are to be determined from the boundary conditions. The boundary conditions imply that the tangential component of the electric field on the walls of the conducting boundary should vanish,

$$\hat{n} \times \vec{E} \Big|_C = 0 \quad (16.3.4)$$

where \hat{n} represents the normal to the bounding surfaces.

The general solution, for example for the x -component of the \vec{E} field, is found to be

$$E_x = E_x^{(0)} \cos(\sqrt{\omega^2 - \lambda_x^2 - \sigma_x^2} x + \phi_x^{(1)}) \cos(\lambda_x y + \phi_x^{(2)}) \cos(\sigma_x z + \phi_x^{(3)}) \quad (16.3.5)$$

and the boundary condition requires that

$$E_x|_{y=0,b} = E_x|_{z=0,c} = 0 \quad (16.3.6)$$

The conditions at $y = 0 = z$ imply that $\phi_x^{(2)} = \phi_x^{(3)} = \frac{\pi}{2}$, while the conditions at $y = b, z = c$ require

$$\lambda_x = \frac{n\pi}{b}, \quad \sigma_x = \frac{m\pi}{c} \quad (16.3.7)$$

for $n, m \in \mathbb{N}$. Thus we have

$$E_x = E_x^{(0)} \cos\left(\sqrt{\omega^2 - \frac{n^2\pi^2}{b^2} - \frac{m^2\pi^2}{c^2}} x + \phi_x^{(1)}\right) \sin\left(\frac{n\pi y}{b}\right) \sin\left(\frac{m\pi z}{c}\right) \quad (16.3.8)$$

In a similar fashion we may write the solutions for the other two components:

$$\begin{aligned} E_y &= E_y^{(0)} \sin\left(\frac{l\pi x}{a}\right) \cos\left(\sqrt{\omega^2 - \frac{l^2\pi^2}{a^2} - \frac{m^2\pi^2}{c^2}} y + \phi_y^{(2)}\right) \sin\left(\frac{m\pi z}{c}\right) \\ E_z &= E_z^{(0)} \sin\left(\frac{l\pi x}{a}\right) \sin\left(\frac{n\pi y}{b}\right) \cos\left(\sqrt{\omega^2 - \frac{l^2\pi^2}{a^2} - \frac{n^2\pi^2}{b^2}} z + \phi_z^{(3)}\right) \end{aligned} \quad (16.3.9)$$

where $l, m, n \in \mathbb{N}$ are arbitrary integers. Again, the divergence equation, $\vec{\nabla} \cdot \vec{E} = 0$, is solved by taking

$$\omega^2 = \frac{l^2\pi^2}{a^2} + \frac{n^2\pi^2}{b^2} + \frac{m^2\pi^2}{c^2} \quad (16.3.10)$$

and $\phi_x^{(1)} = \phi_y^{(2)} = \phi_z^{(3)} = 0$. The final form of the solution is then

$$\begin{aligned} E_x &= E_x^{(0)} \cos\left(\frac{l\pi x}{a}\right) \sin\left(\frac{n\pi y}{b}\right) \sin\left(\frac{m\pi z}{c}\right) e^{-i\omega t} \\ E_y &= E_y^{(0)} \sin\left(\frac{l\pi x}{a}\right) \cos\left(\frac{n\pi y}{b}\right) \sin\left(\frac{m\pi z}{c}\right) e^{-i\omega t} \\ E_z &= E_z^{(0)} \sin\left(\frac{l\pi x}{a}\right) \sin\left(\frac{n\pi y}{b}\right) \cos\left(\frac{m\pi z}{c}\right) e^{-i\omega t} \end{aligned} \quad (16.3.11)$$

and the condition $\vec{\nabla} \cdot \vec{E} = 0$ translates simply into the condition $\vec{\kappa} \cdot \vec{E} = 0$, *i.e.*, that the wave-vector $\vec{\kappa} = \pi(l/a, n/b, m/c)$ is perpendicular to \vec{E} . An analogous treatment of the cylindrical resonant cavity would give a solution in terms of Bessel functions instead of the trigonometric functions we have found for the rectangular case. We will not develop these solutions in this course.

Chapter 17

Time Varying Charge Distributions

17.1 General Considerations

So far we have said very little about how the waves treated in the previous chapters are actually generated. Naturally we expect that radiation is produced by charges in motion or, more precisely, by the inhomogeneous part,

$$A_\mu(x) = \int d^4x' G_R(x, x') j_\mu(x') \quad (17.1.1)$$

of our general solutions in Part I of these notes. In Part III we will concentrate on the integral on the r.h.s. of the above equation and its physical interpretation. Let us imagine that we have a localized system of charges and currents occupying a region, \mathcal{R} , of space, whose characteristic dimension is \mathfrak{d} . Specify the current 4-vector describing the distribution by $j_\mu(\vec{r}, t)$. In studying processes that radiate electromagnetic waves, we are particularly interested in the so-called “spectral distribution” of the radiation, *i.e.*, in the intensity of the radiation that is emitted at very large distances (compared to the typical size of the distribution itself) as a function of the frequency and the solid angle of emission. The spectral distribution is best examined by working with the temporal Fourier transforms of the fields and currents. Consider, therefore, the temporal Fourier transform of the current

$$j_\mu(\vec{r}, \omega) = \int_{-\infty}^{\infty} \frac{dt}{\sqrt{2\pi}} e^{i\omega t} j_\mu(\vec{r}, t). \quad (17.1.2)$$

The functions $j_\mu(\vec{r}, \omega)$ are the Fourier components of $\vec{j}(\vec{r}, t)$, ω is the frequency and the inverse transformation is

$$j_\mu(\vec{r}, t) = \int_{-\infty}^{\infty} \frac{d\omega}{\sqrt{2\pi}} e^{-i\omega t} j_\mu(\vec{r}, \omega) \quad (17.1.3)$$

Now we know that the (retarded) solution for the vector potential, $A_\mu(x)$, in the absence of external incoming or outgoing radiation takes the form

$$A_\mu(x) = \int d^4x' G_R(x, x') j_\mu(x') = \frac{c}{4\pi} \int d^3\vec{r}' \int dt' \frac{j_\mu(\vec{r}', t')}{|\vec{r} - \vec{r}'|} \delta(ct' + |\vec{r} - \vec{r}'| - ct) \quad (17.1.4)$$

(notice that $d^4x \equiv c dt d^3\vec{r}$) and inserting (17.1.3) into the r.h.s. of this equation yields

$$A_\mu(\vec{r}, t) = \frac{c}{4\pi} \int_{-\infty}^{\infty} \frac{d\omega}{\sqrt{2\pi}} \int d^3\vec{r}' \int dt' \frac{j_\mu(\vec{r}', \omega) e^{-i\omega t'}}{|\vec{r} - \vec{r}'|} \delta(ct' + |\vec{r} - \vec{r}'| - ct) \quad (17.1.5)$$

But the r.h.s. of this equation can be greatly simplified by exploiting the δ -function and integrating over t' . Recall that

$$\delta(f(x)) = \sum_{\alpha} \frac{\delta(x - x_\alpha)}{|f'(x)|_{x_\alpha}} \quad (17.1.6)$$

where x_α are the roots of the equation $f(x) = 0$ and $f'(x)$ is the derivative of $f(x)$ w.r.t. x . Using this we have

$$A_\mu(\vec{r}, t) = \frac{1}{4\pi} \int_{-\infty}^{\infty} \frac{d\omega}{\sqrt{2\pi}} e^{-i\omega t} \int d^3\vec{r}' \frac{j_\mu(\vec{r}', \omega)}{|\vec{r} - \vec{r}'|} e^{ik|\vec{r} - \vec{r}'|} \quad (17.1.7)$$

where $k = \omega/c$. Again, if we note that the vector potential $A_\mu(x)$ also has a temporal Fourier expansion just like $j_\mu(x)$, *i.e.*,

$$A_\mu(\vec{r}, t) = \int_{-\infty}^{\infty} \frac{d\omega}{\sqrt{2\pi}} e^{-i\omega t} A_\mu(\vec{r}, \omega) \quad (17.1.8)$$

and compare the two expressions, we find that

$$A_\mu(\vec{r}, \omega) = \frac{1}{4\pi} \int d^3\vec{r}' \frac{j_\mu(\vec{r}', \omega) e^{ik|\vec{r} - \vec{r}'|}}{|\vec{r} - \vec{r}'|} \quad (17.1.9)$$

All relevant information (the fields \vec{E} and \vec{B}) can be computed if the potentials are known, *i.e.*, if we are able to solve the integrals that appear in the above formula for $A_\mu(\vec{r}, \omega)$. However, in practice this is impossible to do except in some very simple cases. Therefore let us establish some general properties of the solution under some special conditions.

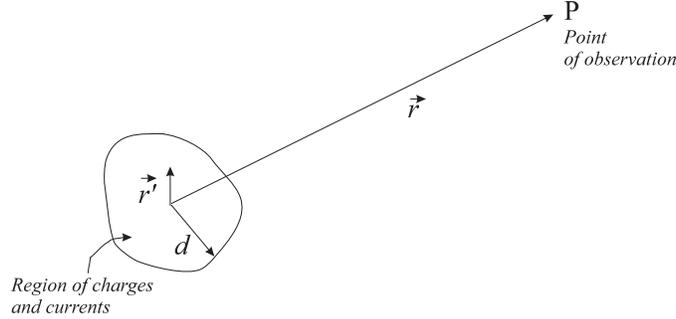


Figure 17.1: Localized charge distributions

When necessary we will assume that the relevant frequencies, ω , are such that the corresponding wavelengths $\lambda = 2\pi c/\omega$ are very large compared with the characteristic size, \mathfrak{d} , of the distribution, *i.e.*, $\lambda \gg \mathfrak{d}$ (see figure 22). (What this means is that the Fourier coefficients in (17.1.2) are appreciable only when $\lambda \gg \mathfrak{d}$.) This is called the long wavelength approximation. Further we will consider the following observational “zones”:

- The “near” (static) zone ($|\vec{r}'| \ll \lambda$): This is the region in which the observation point is distant from the source but much less than one wavelength away from it. In this case the exponent in the integral on the r.h.s. of (17.1.9) is much less than unity and the exponential can be safely approximated as “one”,

$$A_\mu(\vec{r}, \omega) \approx \frac{1}{4\pi} \int d^3\vec{r}' \frac{j_\mu(\vec{r}', \omega)}{|\vec{r} - \vec{r}'|} \quad (17.1.10)$$

The solutions are quasi-stationary, *i.e.*, harmonic but otherwise static.

- The “far” (radiation) zone ($|\vec{r}'| \gg \lambda$): In the far zone, $k|\vec{r} - \vec{r}'| \gg 1$. Let us consider the expansion

$$|\vec{r} - \vec{r}'| = |\vec{r}'| \sqrt{1 + \frac{\vec{r}'^2}{r'^2} - 2\frac{\hat{n} \cdot \vec{r}'}{|\vec{r}'|}} \approx |\vec{r}'| - \hat{n} \cdot \vec{r}' \quad (17.1.11)$$

where \hat{n} is the unit vector in the direction of \vec{r} . Therefore, in the far zone we have

$$A_\mu(\vec{r}, \omega) = \frac{e^{ikr}}{4\pi r} \int d^3\vec{r}' j_\mu(\vec{r}', \omega) e^{-ik\hat{n} \cdot \vec{r}'} \quad (17.1.12)$$

where $r = |\vec{r}'|$. The appearance of the term e^{ikr}/r signals that asymptotically the fields correspond to outgoing spherical waves. Again,

$$e^{ik\hat{n} \cdot \vec{r}'} = \sum_{n=0}^{\infty} \frac{(-ik)^n}{n!} (\hat{n} \cdot \vec{r}')^n \quad (17.1.13)$$

but, as $k|\vec{r}'| \ll 1$ we see that the successive terms in the series will fall off very rapidly so that the main contributions to the integral come from the lowest non-vanishing terms.

- The “intermediate” (induction) zone ($|\vec{r}'| \approx \lambda$) The intermediate zone is more difficult because neither of the approximations made in the preceding two cases will succeed. In this case the exact expressions in (17.1.9) and (17.1.8) should be used.

Away from charges, the electric and magnetic fields are extracted from our solutions by taking appropriate derivatives:

$$\vec{B} = \vec{\nabla} \times \vec{A} \quad (17.1.14)$$

and

$$\begin{aligned} \vec{\nabla} \times \vec{B} &= \frac{1}{c^2} \frac{\partial \vec{E}}{\partial t} \quad \rightarrow \quad -i \frac{k}{c} \vec{E}(\vec{r}, \omega) = \vec{\nabla} \times \vec{B}(\vec{r}, \omega) \\ \rightarrow \quad \vec{E}(\vec{r}, \omega) &= \frac{ic}{k} \vec{\nabla} \times (\vec{\nabla} \times \vec{A}(\vec{r}, \omega)) \end{aligned} \quad (17.1.15)$$

where, taking into account the definition of $j_\mu(x)$, the three vector potential is

$$\vec{A}(\vec{r}, \omega) = \frac{\mu_o}{4\pi} \int d^3\vec{r}' \frac{\vec{j}(\vec{r}', \omega) e^{ik|\vec{r}-\vec{r}'|}}{|\vec{r}-\vec{r}'|} \quad (17.1.16)$$

17.2 Current densities

Before we go on to discussing the fields produced by some interesting charge distributions, we should say a few words about the currents. Recall that in Part I we discovered that the current density $j_\mu(x)$ is defined by

$$j^\mu = (j^0, j^i); \quad j^0 = \mu_0 \rho, \quad j^i = \mu_0 j_i \quad (17.2.1)$$

and must transform as a contravariant vector. Consider a single particle of charge e . We can later construct more complex densities by putting together individual charges. Now we know that in the non-relativistic limit, at least, the current three-vector should take the form

$$\vec{j} = e\vec{v}\delta^{(3)}(\vec{r} - \mathbf{r}(t)), \quad (17.2.2)$$

where $\mathbf{r}(t)$ is the trajectory of the particle. In the same spirit, a natural definition of the charge density would be

$$\rho = e\delta^{(3)}(\vec{r} - \mathbf{r}(t)) \quad (17.2.3)$$

The question is this: is the vector defined by (17.2.1), with ρ and \vec{j} as given in (17.2.2) and (17.2.3) a four-vector density as required, *i.e.*, does it transform appropriately under Lorentz transformations?

Naïvely it certainly doesn't *look* like a four-vector density, because we know that the three vector velocity \vec{v} does not transform as a vector in special relativity (and, again, the δ -function, $\delta^{(3)}(\vec{r}-\mathbf{r}(t))$, appears to make matters worse!). We can unify the description of the current and density by considering the appropriate generalization of the three velocity, *i.e.*, the four-vector velocity which is defined by

$$U^\mu = \frac{dx^\mu}{d\tau} \quad (17.2.4)$$

where $d\tau = ds/c$ is the *proper time* (and so it is a scalar under Lorentz transformations). U^μ is manifestly a four-vector and its spatial components are

$$U^i = \frac{dx^i}{d\tau} = \frac{dx^i}{dt} \frac{dt}{d\tau} = \frac{v^i}{\sqrt{1 - \vec{v}^2/c^2}}, \quad U^0 = \frac{1}{\sqrt{1 - \vec{v}^2/c^2}} \quad (17.2.5)$$

with the following non-relativistic limits

$$\lim_{c \rightarrow \infty} U^i = v^i, \quad \lim_{c \rightarrow \infty} U^0 = 1. \quad (17.2.6)$$

It should therefore be clear that the quantity

$$j^\mu = e\mu_o \int d\tau U^\mu(\tau) \delta^{(4)}(x - x(\tau)) = e\mu_o \int d\tau \frac{dx^\mu}{d\tau} \delta^{(4)}(x - x(\tau)) \quad (17.2.7)$$

is a four-vector density under Lorentz transformations. However, we can replace the integration variable by the coordinate time parameter, t , writing

$$j^\mu = e\mu_o \int dt \frac{dx^\mu}{dt} \delta^{(3)}(\vec{r} - \mathbf{r}(t)) \delta(t - t(\tau)) \quad (17.2.8)$$

and perform the integration over t with the help of the δ -function. We should find just

$$j^\mu = e\mu_o \frac{dx^\mu(t)}{dt} \delta^{(3)}(\vec{r} - \mathbf{r}(t)). \quad (17.2.9)$$

Its components are precisely those given in (17.2.1), (17.2.2) and (17.2.3) above, so these definitions were sound to begin with. If we had N charges, e_n , with trajectories $\vec{r}_n(t)$, the four-vector current density representing the distribution would be simply

$$j^\mu = \mu_o \sum_{n=1}^N e_n \delta(\vec{r} - \vec{r}_n(t)) \frac{dx_n^\mu(t)}{dt} \quad (17.2.10)$$

as we had used in Chapter 6.

Let's also verify that it does obey the conservation condition. We see that

$$\begin{aligned}
 \vec{\nabla} \cdot \vec{j}(\vec{r}, t) &= \sum_n e_n \frac{\partial}{\partial x^i} \delta^{(3)}(\vec{r} - \vec{r}_n) \frac{dx_n^i(t)}{dt} \\
 &= - \sum_n e_n \frac{\partial}{\partial x_n^i} \delta^{(3)}(\vec{r} - \vec{r}_n) \frac{dx_n^i}{dt} \\
 &= - \sum_n e_n \frac{\partial}{\partial t} \delta^{(3)}(\vec{r} - \vec{r}_n) = - \frac{\partial}{\partial t} \rho(\vec{r}, t)
 \end{aligned} \tag{17.2.11}$$

or, in our covariant language,

$$\partial_\mu j^\mu = 0 \tag{17.2.12}$$

in which form, manifest covariance is obvious.

17.3 Fields in the far zone: Spectral Distribution

The electric and magnetic fields in the far zone are radiation fields and merit further attention. Using equations (17.1.15) and (17.1.12) we find,

$$\vec{B}(\vec{r}, \omega) \approx \frac{\mu_o}{4\pi} \left(\vec{\nabla} \frac{e^{ikr}}{r} \right) \times \int d^3\vec{r}' \vec{j}(\vec{r}', \omega) e^{-ik\hat{n}\cdot\vec{r}'}. \tag{17.3.1}$$

Now it is an easy exercise to check that

$$\vec{\nabla} \frac{e^{ikr}}{r} = \frac{e^{ikr}}{r} \left[ik - \frac{1}{r} \right] \hat{n}, \tag{17.3.2}$$

where \hat{n} , as before, is the unit vector pointing in the direction of \vec{r} . Therefore, neglecting terms of order r^{-2} we have

$$\vec{B}(\vec{r}, \omega) \approx \frac{ik\mu_o}{4\pi} \frac{e^{ikr}}{r} \int d^3\vec{r}' [\hat{n} \times \vec{j}(\vec{r}', \omega)] e^{-ik\hat{n}\cdot\vec{r}'} \tag{17.3.3}$$

To compute the electric field we need to take one more rotation of the above. Again, neglecting terms of order r^{-2} we come up with

$$\vec{E}(\vec{r}, \omega) \approx -\frac{ick\mu_o}{4\pi} \frac{e^{ikr}}{r} \int d^3\vec{r}' [\hat{n} \times (\hat{n} \times \vec{j}(\vec{r}', \omega))] e^{-ik\hat{n}\cdot\vec{r}'} \tag{17.3.4}$$

Thus we find that

$$\vec{E}(\vec{r}, \omega) = c\vec{B}(\vec{r}, \omega) \times \hat{n} \tag{17.3.5}$$

and

$$\vec{B}(\vec{r}, \omega) = \frac{1}{c} \hat{n} \times \vec{E}(\vec{r}, \omega), \quad (17.3.6)$$

so the electric and magnetic fields in the far zone are perpendicular to each other and to the direction \hat{n} , which is the direction of propagation of the outgoing spherical waves.

Next, consider the flux of energy across any surface. It is given by the Poynting vector

$$\vec{S} = \frac{1}{\mu_o} \text{Re}(\vec{E}) \times \text{Re}(\vec{B}) = \frac{1}{c\mu_o} \text{Re}(\vec{E}) \times [\hat{n} \times \text{Re}(\vec{E})] = c\epsilon_o [\text{Re}(\vec{E})]^2 \hat{n} \quad (17.3.7)$$

Alternatively,

$$\vec{S} = \frac{c}{\mu_o} (\text{Re}(\vec{B}) \times \hat{n}) \times \text{Re}(\vec{B}) = \frac{c}{\mu_o} [\text{Re}(\vec{B})]^2 \hat{n} \quad (17.3.8)$$

Now if we call $\vec{k} = k\hat{n} = \hat{n}\omega/c$ and

$$\vec{j}(\vec{k}, \omega) = \int d^3\vec{r}' \vec{j}(\vec{r}', \omega) e^{-i\vec{k}\cdot\vec{r}'} = \int dt \int d^3\vec{r}' \vec{j}(\vec{r}', t) e^{-i(\vec{k}\cdot\vec{r}' - \omega t)} \quad (17.3.9)$$

then

$$\vec{B}(\vec{r}, \omega) \approx \frac{ik\mu_o}{4\pi} \frac{e^{ikr}}{r} [\hat{n} \times \vec{j}(\vec{k}, \omega)] \quad (17.3.10)$$

and, for the electric field,

$$\vec{E}(\vec{r}, \omega) \approx -\frac{ick\mu_o}{4\pi} \frac{e^{ikr}}{r} \hat{n} \times [\hat{n} \times \vec{j}(\vec{k}, \omega)] \quad (17.3.11)$$

we obviously find the analogous expression. Note that the reality of $\vec{j}(\vec{r}, t)$ ensures the relation

$$\vec{j}^*(\vec{k}, \omega) = \vec{j}(-\vec{k}, -\omega) \quad (17.3.12)$$

between the Fourier components of \vec{j} and their complex conjugates.

Let us evaluate the *time averaged* power radiated, at a fixed frequency, across the surface of a large sphere of radius r , with element $d\vec{\sigma} = r^2\hat{n}d\Omega$ (where Ω is the solid angle). This is given by

$$\begin{aligned} \mathfrak{I}(\omega) &= 4\pi \int_{\Sigma} \langle \vec{S} \rangle \cdot d\vec{\sigma} = \frac{4\pi r^2 c}{\mu_o} \int_{\Omega} d\Omega \langle [\text{Re}(\vec{B}(\vec{r}, \omega) e^{-i\omega t})]^2 \rangle \\ &= \frac{2\pi r^2 c}{\mu_o} \int_{\Omega} d\Omega \vec{B}^*(\vec{r}, \omega) \cdot \vec{B}(\vec{r}, \omega) = \frac{2\pi r^2 c}{\mu_o} \int_{\Omega} d\Omega |\vec{B}(\vec{r}, \omega)|^2 \end{aligned} \quad (17.3.13)$$

using theorem (14.2.6). With (17.3.10), the above expression allows us to write the average power radiated in the frequency interval $(\omega, \omega + d\omega)$ within the solid angle $d\Omega$ as

$$d^2\mathfrak{I}(\omega) = \frac{\mu_o\omega^2}{2c} |\hat{n} \times \vec{j}(\vec{k}, \omega)|^2 \frac{d\Omega d\omega}{4\pi} = \frac{\mu_o\omega^2}{2c} [\vec{j} \cdot \vec{j}^* - (\hat{n} \cdot \vec{j})(\hat{n} \cdot \vec{j}^*)] \frac{d\Omega d\omega}{4\pi} \quad (17.3.14)$$

This is called a spectral distribution. The central problem in understanding the radiation fields in the far zone is in the determination of $\vec{j}(\vec{k}, \omega)$. It is most often not easy to do and we must make approximations. These approximations will be the subject of future sections, but let us first have a look at some cases that are important and in which $\vec{j}(\vec{k}, \omega)$ can be determined exactly.

17.4 Examples

17.4.1 Point Charge in Uniform motion

As our first and particularly simple application of the formalism above, we'll now compute the fields seen by a distant observer due to a single point charge in uniform motion. All we need is the Fourier transform of the current density for a point charge in uniform motion. The current density is

$$\vec{j}(\vec{r}, t) = e\vec{v}\delta^{(3)}(\vec{r} - \mathbf{r}(t)) \quad (17.4.1)$$

where $\mathbf{r}(t)$ is the trajectory of the particle. Its Fourier transform may be calculated as

$$\begin{aligned} \vec{j}(\vec{k}, \omega) &= e\vec{v} \int_{-\infty}^{\infty} dt \int d^3\vec{r}' \delta(\vec{r}' - \mathbf{r}(t)) e^{-i(\vec{k}\cdot\vec{r}' - \omega t)} \\ &= e\vec{v} \int_{-\infty}^{\infty} dt e^{i\omega(t - \hat{n}\cdot\mathbf{r}(t)/c)} = e\vec{v} \int_{-\infty}^{\infty} dt e^{i\omega t(1 - \hat{n}\cdot\vec{v}/c)} \\ &= \frac{e\vec{v}}{\omega} \delta\left(1 - \frac{\hat{n}\cdot\vec{v}}{c}\right) \end{aligned} \quad (17.4.2)$$

Because $|\vec{v}| < c$, this is *always* vanishing for a particle in the vacuum and we have just proved that a charge in uniform motion does not radiate.

Note: This is true only in a vacuum. In a medium the speed of light is less than c and a uniformly moving particle can radiate. This is called Cerenkov radiation. In a medium we should replace c by the velocity of light in the medium: c/n where n is the refractive index of the medium. Then the δ -function is supported when

$$1 = \frac{n}{c} \hat{n} \cdot \vec{v} = \frac{nv \cos \theta}{c} \rightarrow \theta = \cos^{-1}\left(\frac{c}{nv}\right) \quad (17.4.3)$$

which angle defines the opening angle of a cone along which the radiation will be emitted.

17.4.2 Point Charge in uniform circular motion: Synchrotron Radiation

A modestly more complicated example is presented by a charge in uniform circular motion. Consider the following (simplified) treatment of a charge in instantaneous circular motion,

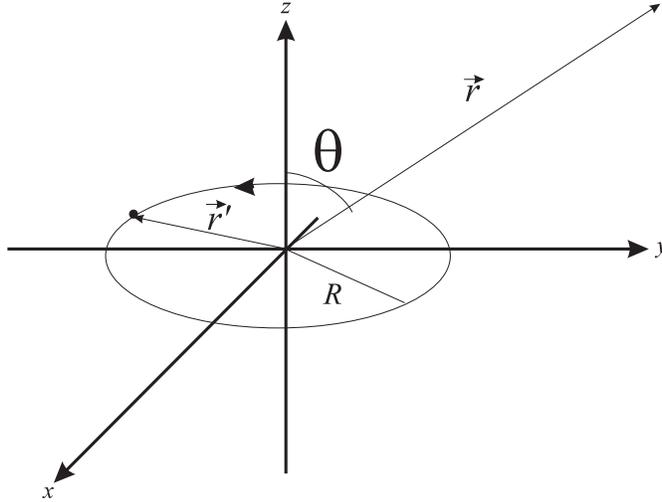


Figure 17.2: Charge in uniform circular motion

assuming that the particle has a circular motion in the $x - y$ plane (see figure 23). It is convenient to go to cylindrical coordinates and because the motion is circular only the φ component of the current vector will be non-vanishing. Let ω_o be the angular velocity of the charge (ω_o is constant) and let R be the radius of the circle, then it is easy to see that

$$\vec{j}(\vec{r}', t') = \vec{v}' \delta^{(3)}(\vec{r}' - \mathbf{r}(t')) = e\omega_o \delta(z') \delta(r' - R) \delta(\varphi' - \omega_o t') \hat{\varphi}'(t) \quad (17.4.4)$$

and all other components vanish identically. We must compute

$$\vec{j}(\vec{k}, \omega) = \int dt' \int d^3\vec{r}' \vec{j}(\vec{r}', t') e^{-i(\vec{k} \cdot \vec{r}' - \omega t')} \quad (17.4.5)$$

so let us first evaluate $\vec{k} \cdot \vec{r}'$. We have (in Cartesian coordinates)

$$\begin{aligned} \hat{n} &= \sin \theta \cos \varphi \hat{x} + \sin \theta \sin \varphi \hat{y} + \cos \theta \hat{z} \\ \vec{r}' &= r' \cos \varphi' \hat{x} + r' \sin \varphi' \hat{y} \end{aligned} \quad (17.4.6)$$

giving

$$\vec{k} \cdot \vec{r}' = \frac{\omega}{c} \hat{n} \cdot \vec{r}' = \frac{\omega}{c} r' \sin \theta \cos(\varphi - \varphi') \quad (17.4.7)$$

Moreover, because $\hat{\varphi}' \equiv \hat{\varphi}'(t')$ depends on t' and we will eventually be integrating over t' , it is convenient to make the time dependence explicit by expressing $\vec{j}(\vec{r}', t')$ in terms of

the (rigid) basis vectors \hat{x} , \hat{y} and \hat{z} :

$$\vec{j}(\vec{r}', t') = e\omega_o\delta(z')\delta(r' - R)\delta(\varphi' - \omega_o t')[-\sin\varphi'\hat{x} + \cos\varphi'\hat{y}] \quad (17.4.8)$$

Therefore

$$\begin{aligned} \vec{j}(\vec{k}, \omega) &= e \int dt' e^{i\omega t'} \int r' dr' d\varphi' dz' \vec{j}(\vec{r}', t') e^{-i\frac{\omega}{c}r' \sin\theta \cos(\varphi - \varphi')} \\ &= e\omega_o \int dt' e^{i\omega t'} \int r' dr' d\varphi' dz' \delta(z')\delta(r' - R)\delta(\varphi' - \omega_o t') \times \\ &\quad [-\sin\varphi'\hat{x} + \cos\varphi'\hat{y}] e^{-i\frac{\omega}{c}r' \sin\theta \cos(\varphi - \varphi')} \end{aligned}$$

Performing the integrals over the δ -functions is easy and leads to

$$\vec{j}(\vec{k}, \omega) = e\omega_o R \int dt' e^{i\omega(t' - \frac{1}{c}R \sin\theta \cos(\varphi - \omega_o t'))} [-\sin\omega_o t' \hat{x} + \cos\omega_o t' \hat{y}] \quad (17.4.9)$$

Now recall that the speed of the particle, v , is given by $\omega_o R$, so we could replace $R = v/\omega_o$. Again, because the motion is periodic, what is really needed is a discrete Fourier transform. We could obtain this without starting from the beginning by letting $\omega/\omega_o = n \in \mathbb{N}$ and restricting the integration to the interval $t' \in [-\pi/\omega_o, +\pi/\omega_o]$. Call $\omega_o t' = u$ then

$$\vec{j}(\vec{k}, \omega) = eR \int_{-\pi}^{\pi} du e^{i(nu - n\frac{v}{c} \sin\theta \cos(\varphi - u))} [-\sin u \hat{x} + \cos u \hat{y}] \quad (17.4.10)$$

Of course, the situation is cylindrically symmetric and let us exploit this fact. Cylindrical (axial) symmetry allows us to choose any convenient value of φ . The choice $\varphi = \pi/2$ brings the integral above to standard form:

$$\vec{j}(\vec{k}, \omega) = eR \int_{-\pi}^{\pi} du [-\sin u \hat{x} + \cos u \hat{y}] e^{i(nu - n\frac{v}{c} \sin\theta \sin u)} \quad (17.4.11)$$

It is convenient to compute the components separately. Thus, consider the integral

$$\begin{aligned} - \int_{-\pi}^{\pi} du \sin u e^{i(nu - n\frac{v}{c} \sin\theta \sin u)} \hat{x} &= \\ &= -\frac{1}{2i} \int_{-\pi}^{\pi} du [e^{i([n+1]u - n\frac{v}{c} \sin\theta \sin u)} - e^{i([n-1]u - n\frac{v}{c} \sin\theta \sin u)}] \\ &= \frac{i}{2} [J_{n+1}(z) - J_{n-1}(z)] \end{aligned} \quad (17.4.12)$$

where $z = n\frac{v}{c} \sin\theta$ and where $J_n(z)$ is a Bessel function of order $n \in \mathbb{N}$. Likewise, the second term

$$- \int_{-\pi}^{\pi} du \cos u e^{i(nu - n\frac{v}{c} \sin\theta \sin u)} \hat{y} =$$

$$\begin{aligned} & \frac{1}{2} \int_{-\pi}^{\pi} du [e^{i([n+1]u - n\frac{v}{c} \sin \theta \sin u)} + e^{i([n-1]u - n\frac{v}{c} \sin \theta \sin u)}] \\ &= \frac{1}{2} [J_{n+1}(z) + J_{n-1}(z)] \end{aligned} \quad (17.4.13)$$

Combining the two we find

$$\vec{j}(\vec{k}, \omega) = \frac{eR}{2} [J_{n+1}(z)(i\hat{x} + \hat{y}) + J_{n-1}(z)(-i\hat{x} + \hat{y})] \quad (17.4.14)$$

The spectral distribution is obtained directly from the expression we had earlier. We easily compute (noting that, with $\varphi = \pi/2$ we have $\hat{n} = \hat{y} \sin \theta + \hat{z} \cos \theta$)

$$\begin{aligned} \vec{j} \cdot \vec{j}^* &= \frac{e^2 R^2}{2} [J_{n+1}^2(z) + J_{n-1}^2(z)] \\ (\hat{n} \cdot \vec{j})(\hat{n} \cdot \vec{j}^*) &= \frac{e^2 R^2}{4} [J_{n+1}(z) + J_{n-1}(z)]^2 \sin^2 \theta \end{aligned} \quad (17.4.15)$$

which gives

$$d\mathfrak{J}_n = \frac{\mu_o e^2 n^2 v^2}{8c} [(J_{n+1}(z) + J_{n-1}(z))^2 + (J_{n+1}(z) - J_{n-1}(z))^2 \cos^2 \theta] \frac{d\Omega}{4\pi} \quad (17.4.16)$$

where we've replaced ωR by $v\omega/\omega_o = nv$. The Bessel function is highly peaked about $\theta = \pi/2$, which is in the plane of the orbit and vanishes along the z -axis.¹ Thus there is no power radiated along the axis of rotation and most of it is radiated in the plane of the orbit. Of course, we also know the electric and magnetic fields, if we know $\vec{j}(\vec{k}, \omega)$. In the far zone, they follow directly from (17.3.10) and (17.3.11).²

The formula above can be simplified in special cases. First, let us note the following identities obeyed by the Bessel functions

$$\begin{aligned} J_{n+1}(z) + J_{n-1}(z) &= \frac{2n}{z} J_n(z) \\ J_{n+1}(z) - J_{n-1}(z) &= -2 \frac{d}{dz} J_n(z) \end{aligned} \quad (17.4.17)$$

Using them, we might re-express the formula (17.4.16) in terms of $J_n(z)$ as follows

$$d\mathfrak{J}_n = \frac{\mu_o e^2 n^2 v^2}{2c} \left[\frac{n^2}{z^2} J_n^2(z) + \left(\frac{d}{dz} J_n(z) \right)^2 \cos^2 \theta \right] \frac{d\Omega}{4\pi} \quad (17.4.18)$$

¹It is a good exercise to plot these angular distributions!

²Problem: Compute the electric and magnetic fields to leading order in $1/r$, for uniform circular motion.

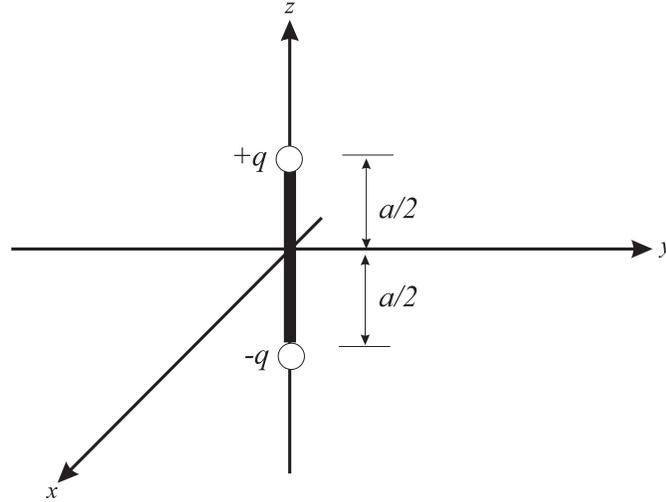


Figure 17.3: An oscillating Dipole

17.4.3 Radiation from an oscillating Dipole

Consider the radiation from an oscillating dipole. The physical setup is shown in figure 24. The dipole consists of two charged spheres of opposite charges, $\pm q(t)$ situated respectively at $z = \pm a/2$. The two spheres are connected by an infinitesimally thin wire of length a . The current in the wire must be precisely dq/dt so as to conserve charge. The current density is in the \hat{z} direction and given by

$$\vec{j}(\vec{r}', t') = \dot{q}(t')\delta(x')\delta(y')\Theta\left(\frac{a}{2} - |z'|\right)\hat{z} = I(t')\delta(x')\delta(y')\Theta\left(\frac{a}{2} - |z'|\right)\hat{z} \quad (17.4.19)$$

where $I(t)$ is the current in the wire and $\Theta(x)$ above represents the Heaviside θ -function. (We have used the capital Θ only to distinguish the Heaviside function from the polar angle, θ .) Its function is to ensure that the current vanishes outside the assembly. Note, first of all, that $I(t)$ is a function only of t and not of z ; this uniform current can be achieved in practice only if the length of the wire is small compared with the wavelength of the radiation. Let

$$I(t') = I_0 e^{-i\tilde{\omega}t'} \quad (17.4.20)$$

be a harmonic function of time (so the current oscillates at a definite frequency), and calculate

$$\vec{j}(\vec{k}, \omega) = \hat{z} I_0 \delta(\omega - \tilde{\omega}) \int dx' dy' dz' \delta(x') \delta(y') \Theta\left(\frac{a}{2} - |z'|\right) e^{-i\frac{\omega}{c}\hat{n}\cdot\vec{r}'}$$

$$\begin{aligned}
&= \hat{z} I_o \delta(\omega - \tilde{\omega}) \int_{-a/2}^{a/2} dz' e^{-i\frac{\omega}{c} z' \cos \theta} \\
&= -\frac{2icI_o}{\omega \cos \theta} \delta(\omega - \tilde{\omega}) \sin\left(\frac{\omega}{c} a \cos \theta\right) \hat{z}.
\end{aligned} \tag{17.4.21}$$

The δ -function formally says that the frequency is fixed to be $\tilde{\omega}$. We are now able to compute the fields in the far zone: they are given by³

$$\begin{aligned}
\vec{B}(\vec{r}, \tilde{\omega}) &= \frac{i\tilde{k}\mu_o}{4\pi} \frac{e^{i\tilde{k}r}}{r} [\hat{n} \times \vec{j}(\vec{k}, \tilde{\omega})] = \frac{2I_o}{4\pi\epsilon_o c^2} \frac{e^{i\tilde{k}r}}{r} \tan \theta \sin\left(\frac{\tilde{\omega}}{c} a \cos \theta\right) \hat{\varphi} \\
\vec{E}(\vec{r}, \tilde{\omega}) &= -\frac{ick\mu_o}{4\pi} \frac{e^{i\tilde{k}r}}{r} \hat{n} \times [\hat{n} \times \vec{j}(\vec{k}, \tilde{\omega})] \\
&= -\frac{2I_o}{4\pi\epsilon_o c} \frac{e^{i\tilde{k}r}}{r} \tan \theta \sin\left(\frac{\tilde{\omega}}{c} a \cos \theta\right) \hat{\theta}
\end{aligned} \tag{17.4.22}$$

to leading order in $1/r$. Higher order corrections ($\mathcal{O}(r^{-2})$ and $\mathcal{O}(r^{-3})$) may likewise be calculated.

The angular distribution of the (monochromatic) radiation, on the other hand, is similarly found from our master equations:

$$d\mathcal{J}(\tilde{\omega}) = \frac{\mu_o \tilde{\omega}^2}{2c} [\vec{j} \cdot \vec{j}^* - (\hat{n} \cdot \vec{j})(\hat{n} \cdot \vec{j}^*)] \frac{d\Omega}{4\pi} = \frac{2I_o^2}{\epsilon_o c} \tan^2 \theta \sin^2\left(\frac{\tilde{\omega}}{c} a \cos \theta\right) \frac{d\Omega}{4\pi} \tag{17.4.23}$$

This time we find that the angular dependence is more complicated but it can be simplified if we remember that we are considering the long wavelength limit, *i.e.*, $\mathfrak{d} = a \ll \lambda$. Put $\tilde{\omega} = 2\pi c/\lambda$ and approximate

$$\sin^2\left(\frac{\tilde{\omega}}{c} a \cos \theta\right) = \sin^2\left(\frac{2\pi a}{\lambda} \cos \theta\right) \approx \frac{4\pi^2 a^2}{\lambda^2} \cos^2 \theta \tag{17.4.24}$$

Then

$$d\mathcal{J}(\lambda) = \frac{2\pi I_o^2}{\epsilon_o c} \frac{a^2}{\lambda^2} \sin^2 \theta d\Omega \tag{17.4.25}$$

The radiation vanishes along the axis of the dipole, achieving its maximum value in the $x - y$ plane.

³Problem: Compute the corrections of order r^{-2} and r^{-3} to these formulæ.

17.4.4 The Half-Wave Antenna

In the previous example, we had restricted ourselves to a wire that is small compared to one wavelength, so the variation of the current along the z -axis was ignored. Let us now consider the same example with $a = \lambda/2$ and a current that behaves as

$$\vec{j}(\vec{r}', t') = I_o e^{-i\tilde{\omega}t'} \delta(x') \delta(y') \Theta\left(\frac{\lambda}{4} - |z'|\right) \cos\left(\frac{2\pi z'}{\lambda}\right) \hat{z} \quad (17.4.26)$$

where $\tilde{\omega} = 2\pi c/\lambda$. This is called a half-wave antenna for an obvious reason: its length is precisely one half the wavelength of the radiation. The amplitude of the current oscillates harmonically at the frequency $\tilde{\omega}$. As usual we need to compute $\vec{j}(\vec{k}, \omega)$:

$$\begin{aligned} \vec{j}(\vec{k}, \omega) &= I_o \delta(\omega - \tilde{\omega}) \int_{-\lambda/4}^{\lambda/4} dz' \cos\left(\frac{2\pi z'}{\lambda}\right) e^{-i\frac{2\pi z'}{\lambda} \cos\theta} \\ &= \frac{\lambda I_o}{\pi \sin^2 \theta} \delta(\omega - \tilde{\omega}) \cos\left(\frac{\pi}{2} \cos\theta\right) \hat{z} \end{aligned} \quad (17.4.27)$$

and our master equations give us the asymptotic fields

$$\begin{aligned} \vec{B}(\vec{r}, \tilde{\omega}) &= \frac{ik\mu_o}{4\pi} \frac{e^{i\vec{k}\cdot\vec{r}}}{r} [\hat{n} \times \vec{j}(\vec{k}, \tilde{\omega})] = \frac{i\mu_o I_o}{2\pi \sin\theta} \frac{e^{i\vec{k}\cdot\vec{r}}}{r} \cos\left(\frac{\pi}{2} \cos\theta\right) \hat{\varphi} \\ \vec{E}(\vec{r}, \tilde{\omega}) &= -\frac{ick\mu_o}{4\pi} \frac{e^{i\vec{k}\cdot\vec{r}}}{r} \hat{n} \times [\hat{n} \times \vec{j}(\vec{k}, \tilde{\omega})] \\ &= -\frac{ic\mu_o I_o}{2\pi \sin\theta} \frac{e^{i\vec{k}\cdot\vec{r}}}{r} \cos\left(\frac{\pi}{2} \cos\theta\right) \hat{\theta} \end{aligned} \quad (17.4.28)$$

as well as the angular distribution of the radiation for this case

$$\begin{aligned} d\mathfrak{J}(\tilde{\omega}) &= \frac{\mu_o \omega^2}{2c} [j \cdot j^* - (\hat{n} \cdot j)(\hat{n} \cdot j^*)] \frac{d\Omega}{4\pi} \\ &= \frac{c\mu_o I_o^2}{2 \sin^2 \theta} \cos^2\left(\frac{\pi}{2} \cos\theta\right) \frac{d\Omega}{4\pi} \end{aligned} \quad (17.4.29)$$

17.5 More General Distributions: Further Approximations

We now consider some approximations in the near and far zones, assuming always the long wavelength approximation. We are always interested in

$$\vec{j}(\vec{k}, \omega) = \int_{-\infty}^{\infty} dt' \int d^3\vec{r}' \vec{j}(\vec{r}', t') e^{-i(\vec{k}\cdot\vec{r}' - \omega t')} \quad (17.5.1)$$

Now the exponent $e^{-i\vec{k}\cdot\vec{r}'}$ may be expanded in a Taylor series,

$$e^{-i\vec{k}\cdot\vec{r}'} = \sum_n \frac{(-i\vec{k}\cdot\vec{r}')^n}{n!} \quad (17.5.2)$$

so the expression for $\vec{j}(\vec{k}, \omega)$ becomes

$$\vec{j}(\vec{k}, \omega) = \sum_n \frac{(-ik)^n}{n!} \int_{-\infty}^{\infty} dt' \int d^3\vec{r}' (\hat{n}\cdot\vec{r}')^n \vec{j}(\vec{r}', t') e^{i\omega t'} \quad (17.5.3)$$

In the long wave-length approximation,

$$\vec{k}\cdot\vec{r}' = \frac{2\pi}{\lambda} \hat{n}\cdot\vec{r}' \ll 1 \quad (17.5.4)$$

the expansion above can be thought of as a perturbation series in \mathfrak{d}/λ , where \mathfrak{d} is the characteristic size of the distribution.

17.5.1 The Electric Dipole Approximation: Fields and Radiation

Retaining only the $n = 0$ term (neglecting all higher order corrections) we would have,

$$\vec{j}(\vec{k}, \omega) \approx \int_{-\infty}^{\infty} dt' \int d^3\vec{r}' \vec{j}(\vec{r}', t') e^{i\omega t'} \quad (17.5.5)$$

This is called the ‘‘Electric Dipole’’ approximation for reasons we will see below. Let us note the identity

$$\int_V d^3\vec{r}' \vec{j} = \oint_{\Sigma} \vec{r}' (\vec{j}\cdot\hat{n}') d\sigma' - \int_V d^3\vec{r}' \vec{r}' (\vec{\nabla}'\cdot\vec{j}) \quad (17.5.6)$$

where Σ is some bounding surface enclosing the volume V . If the bounding surface is taken to be at infinity and assuming that the charge distribution is localized, there is no flux across the surface Σ , so that

$$\int_V d^3\vec{r}' \vec{j} = - \int_V d^3\vec{r}' \vec{r}' (\vec{\nabla}'\cdot\vec{j}) \quad (17.5.7)$$

Now, the continuity equation for an arbitrary distribution of charges, means that

$$\vec{\nabla}'\cdot\vec{j}(\vec{r}', t') = - \frac{\partial\rho(\vec{r}', t')}{\partial t'} \quad (17.5.8)$$

where $\rho(\vec{r}', t')$ is the charge density. Inserting this into our expression for $\vec{j}(\vec{k}, \omega)$ we find, to leading order,

$$\vec{j}(\vec{k}, \omega) \approx \int_{-\infty}^{\infty} dt' e^{i\omega t'} \int d^3\vec{r}' \vec{j}(\vec{r}', t') = \int_{-\infty}^{\infty} dt' e^{i\omega t'} \int d^3\vec{r}' \vec{r}' \frac{\partial\rho(\vec{r}', t')}{\partial t'}$$

$$= -i\omega \int_{-\infty}^{\infty} dt' e^{i\omega t'} \int d^3\vec{r}' \vec{r}' \rho(\vec{r}', t') \quad (17.5.9)$$

where we have simply integrated by parts. But now notice that

$$\vec{d}(t') = \int d^3\vec{r}' \vec{r}' \rho(\vec{r}', t') \quad (17.5.10)$$

represents the electric dipole moment of the charge distribution and the integral on the r.h.s. of (17.5.9) represents the Fourier transform of the dipole moment (this is where the approximation gets its name from), which we call $\vec{d}(\omega)$. Thus we have

$$\vec{j}_d(\vec{k}, \omega) \approx -i\omega \vec{d}(\omega). \quad (17.5.11)$$

and the magnetic and electric fields in the far zone are given by (17.3.10) and (17.3.11) of section 20. They are

$$\begin{aligned} \vec{B}_d(\vec{r}, \omega) &\approx \frac{\omega^2 \mu_o}{4\pi c} \frac{e^{ikr}}{r} [\hat{n} \times \vec{d}(\omega)] \\ \vec{E}_d(\vec{r}, \omega) &\approx -\frac{\omega^2 \mu_o}{4\pi} \frac{e^{ikr}}{r} [\hat{n} \times (\hat{n} \times \vec{d}(\omega))] \end{aligned} \quad (17.5.12)$$

to order r^{-1} , showing leading order behavior that is typical of radiation fields.

We may write the spectral distribution due to the time-varying dipole moment of the distribution as

$$d^2\mathcal{J}(\omega) = \frac{\mu_o \omega^4}{2c} |\vec{d}(\omega)|^2 \sin^2 \theta \frac{d\Omega d\omega}{4\pi} \quad (17.5.13)$$

using (17.3.14).

17.5.2 The Magnetic Dipole and Electric Quadrupole Approximation

The next order in the Taylor expansion is quite a bit more difficult to work with. It involves three terms, one of which is interpretable as the radiation from a magnetic dipole and the other two as the electric “quadrupole” radiation. Consider the term with $n = 1$. We should have

$$\vec{j}(\vec{k}, \omega) = -i\frac{\omega}{c} \int_{-\infty}^{\infty} dt' e^{i\omega t'} \int d^3\vec{r}' (\hat{n} \cdot \vec{r}') \vec{j}(\vec{r}', t') \quad (17.5.14)$$

The vector potential can be written as a sum of two terms, one of which gives a transverse magnetic induction and the other of which gives a transverse electric field. To see this, write the integrand above as the sum of two parts as follows

$$(\hat{n} \cdot \vec{r}') \vec{j} = \frac{1}{2} [(\hat{n} \cdot \vec{r}') \vec{j} + (\hat{n} \cdot \vec{j}) \vec{r}'] + \frac{1}{2} (\vec{r}' \times \vec{j}) \times \hat{n} \quad (17.5.15)$$

The second (antisymmetric) part is related to the magnetization due to \vec{j} ,

$$\vec{M} = \frac{1}{2}(\vec{r} \times \vec{j}) \quad (17.5.16)$$

and leads to the “magnetic dipole” contribution to this approximation. The second term is the “electric quadrupole” contribution.

Consider, first, only the magnetic dipole contribution: for the far-zone fields we need the Fourier transform of the current,

$$\begin{aligned} \vec{j}_m(\vec{k}, \omega) &= -i\frac{\omega}{2c} \int_{-\infty}^{\infty} dt' e^{i\omega t'} \int d^3\vec{r}' [\vec{r}' \times \vec{j}(\vec{r}', t')] \times \hat{n} \\ &= -i\frac{\omega}{c} \int_{-\infty}^{\infty} dt' e^{i\omega t'} \vec{m}(t') \times \hat{n} \\ &= -i\frac{\omega}{c} \vec{m}(\omega) \times \hat{n} \end{aligned} \quad (17.5.17)$$

where we have defined the magnetic dipole moment,

$$\vec{m}(t') = \frac{1}{2} \int d^3\vec{r}' [\vec{r}' \times \vec{j}(\vec{r}', t')] = \int d^3\vec{r}' M(\vec{r}', t') \quad (17.5.18)$$

and $\vec{m}(\omega)$ is its Fourier transform. It is now a trivial matter to give the electric and magnetic fields. To leading order in $1/r$ we find

$$\begin{aligned} \vec{E}_m(\vec{r}, \omega) &\approx -\frac{ck^2\mu_o}{4\pi} \frac{e^{ikr}}{r} \hat{n} \times [\hat{n} \times (\vec{m} \times \hat{n})] = \frac{\omega^2\mu_o}{4\pi c} \frac{e^{ikr}}{r} \vec{m}(\omega) \times \hat{n} \\ \vec{B}_m(\vec{r}, \omega) &\approx \frac{\omega^2\mu_o}{4\pi c^2} \frac{e^{ikr}}{r} \hat{n} \times [\vec{m}(\omega) \times \hat{n}] \end{aligned} \quad (17.5.19)$$

We see that the behavior of the magnetic dipole fields is related to the behavior of the electric dipole fields by the correspondence $\vec{d} \rightarrow \vec{m}$, $\vec{B}_d \rightarrow -\vec{E}_m$ and $\vec{E}_d \rightarrow \vec{B}_m$.

Next we consider the quadrupole contribution, but this is a bit more complicated. We must develop the integral

$$j_q(\vec{k}, \omega) = -\frac{i\omega}{2c} \int_{-\infty}^{\infty} dt' e^{i\omega t'} \int d^3\vec{r}' [(\hat{n} \cdot \vec{r}') \vec{j} + (\hat{n} \cdot \vec{j}) \vec{r}']. \quad (17.5.20)$$

It can be written as

$$j_q(\vec{k}, \omega) = -\frac{\omega^2}{2c} \int d^3\vec{r}' [\vec{r}' (\hat{n} \cdot \vec{r}') \rho(\vec{r}', \omega)] \quad (17.5.21)$$

after integrating by parts, using the continuity equation and letting $\partial_t \rightarrow -i\omega$. We can now write the fields (to leading order in $1/r$) as

$$\vec{B}_q(\vec{r}, \omega) \approx -\frac{i\omega^3\mu_o}{8\pi c^2} \frac{e^{ikr}}{r} \int d^3\vec{r}' (\hat{n} \times \vec{r}') (\hat{n} \cdot \vec{r}') \rho(\vec{r}', \omega)$$

$$\vec{E}_q(\vec{r}, \omega) \approx \frac{i\omega^3 \mu_o}{8\pi c} \frac{e^{ikr}}{r} \int d^3\vec{r}' [\hat{n} \times (\hat{n} \times \vec{r}')](\hat{n} \cdot \vec{r}') \rho(\vec{r}', \omega) \quad (17.5.22)$$

However, from the definition of the quadrupole moment tensor:

$$Q_{ij} = \int d^3\vec{r}' [3r'_i r'_j - r'^2 \delta_{ij}] \rho(\vec{r}') \quad (17.5.23)$$

we see that, if we define the vector \vec{Q} by (summation over repeated indices)

$$Q_i = \frac{1}{3} (n_j Q_{ij}) = \int d^3\vec{r}' [r'_i (\hat{n} \cdot \vec{r}') - \frac{1}{3} r'^2 n_i] \rho(\vec{r}'), \quad (17.5.24)$$

then

$$\hat{n} \times \vec{Q} = \int d^3\vec{r}' (\hat{n} \times \vec{r}') (\hat{n} \cdot \vec{r}') \rho(\vec{r}'). \quad (17.5.25)$$

Thus we may express the fields in terms of the quadrupole moment tensor as

$$\begin{aligned} \vec{B}_q(\vec{r}, \omega) &\approx -\frac{i\omega^3 \mu_o}{8\pi c^2} \frac{e^{ikr}}{r} [\hat{n} \times \vec{Q}] \\ \vec{E}_q(\vec{r}, \omega) &\approx \frac{i\omega^3 \mu_o}{8\pi c} \frac{e^{ikr}}{r} [\hat{n} \times (\hat{n} \times \vec{Q})] \end{aligned} \quad (17.5.26)$$

The fields to this order are a sum of the magnetic dipole and the electric quadrupole fields given above. The spectral distribution is straightforward to write out, but the angular distribution is quite complicated.

As one can imagine, evaluating the higher order corrections (beyond the electric quadrupole approximation) is prohibitively complicated in this technique. A more systematic development of the multipole expansion requires some fairly sophisticated mathematical tools.

17.6 The Multipole Expansion

Chapter 18

The Liénard Wiechert Potentials

We turn now to some exact solutions of the inhomogeneous equation. Naturally such solutions can be given and interpreted only when the charge distribution is exceedingly simple, *i.e.*, involving a small number of charges. We consider a single point charge in arbitrary motion.

18.1 Potentials for a single point charge

A single point charge in arbitrary motion (we have already seen that a point charge in *uniform* motion does not radiate) is described by the 4-vector current

$$j^\mu(x) = e\mu_o \int d\tau U^\mu(\tau) \delta^{(4)}(x - x(\tau)) \quad (18.1.1)$$

where $U^\mu = dx^\mu/d\tau$ is time-like (see figure 25). We have chosen to use the 4-vector form of the current here for reasons that will become clear soon. Note that time-like velocities obey the relation $U^\mu U_\mu = U^2 = -c^2$. This follows from the definition of the proper time:

$$d\tau^2 = -\frac{1}{c^2} \eta_{\mu\nu} dx^\mu dx^\nu \quad \rightarrow \quad -c^2 = \eta_{\mu\nu} \frac{dx^\mu}{d\tau} \frac{dx^\nu}{d\tau} = U^\mu U_\mu = U^2 \quad (18.1.2)$$

We are interested, as before, in evaluating the integral

$$A_\mu(x) = \int d^4x' G_R(x, x') j_\mu(x'), \quad (18.1.3)$$

but this time we want an *exact* and *manifestly covariant* solution. Therefore, insert the 4-vector expression for j_μ given above into the integral on the r.h.s. and write

$$A_\mu(\vec{r}, t) = \frac{e\mu_o c}{4\pi} \int d\tau dt' d^3\vec{r}' \frac{\Theta(t - t') U_\mu(\tau)}{|\vec{r} - \vec{r}'|} \delta(t' - \mathfrak{t}(\tau)) \delta^{(3)}(\vec{r}' - \mathfrak{r}(\tau)) \delta(|\vec{r} - \vec{r}'| - c(t - t')) \quad (18.1.4)$$

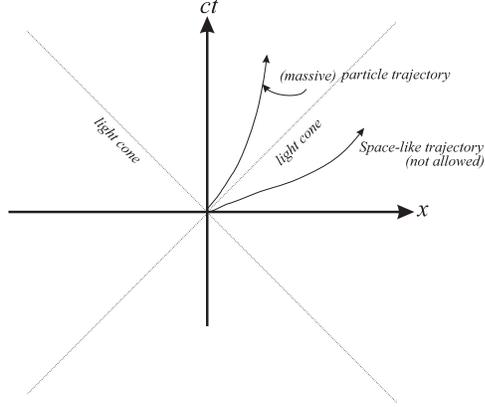


Figure 18.1: Time-like trajectory of a (massive) particle

An integration over \vec{r}' immediately yields

$$A_\mu(\vec{r}, t) = \frac{e\mu_0 c}{4\pi} \int d\tau dt' \frac{\Theta(t - t') U_\mu(\tau)}{|\vec{r} - \mathbf{r}(\tau)|} \delta(t' - \mathbf{t}(\tau)) \delta(|\vec{r} - \mathbf{r}(\tau)| - c(t - t')) \quad (18.1.5)$$

and a further integration over t' may now be performed to give

$$A_\mu(\vec{r}, t) = \frac{e\mu_0 c}{4\pi} \int d\tau \frac{\Theta(t - \mathbf{t}(\tau)) U_\mu(\tau)}{|\vec{r} - \mathbf{r}(\tau)|} \delta(|\vec{r} - \mathbf{r}(\tau)| - c(t - \mathbf{t}(\tau))) \quad (18.1.6)$$

We are left with an integral over the particle proper time. Let

$$f(\tau) = |\vec{r} - \mathbf{r}(\tau)| - c(t - \mathbf{t}(\tau)) \quad (18.1.7)$$

and assume that $f(\tau)$ has only one root, $\tau_o(\vec{r}, t)$, such that $t > \mathbf{t}(\tau_o)$. This makes sense because a time-like trajectory can be expected to intersect the light cone of the observation point in two points, one in the past and one in the future (see figure 26), then

$$\delta(f(\tau)) = \frac{\delta(\tau - \tau_o)}{|\dot{f}(\tau)|_{\tau_o}} \quad (18.1.8)$$

where $\dot{f}(\tau) = df(\tau)/d\tau$. Finally, performing the integral over τ , we get

$$A_\mu(\vec{r}, t) = \frac{e}{4\pi c \epsilon_0} \frac{U_\mu(\tau_o)}{|\dot{f}(\tau_o)| |\vec{r} - \mathbf{r}(\tau_o)|} \quad (18.1.9)$$

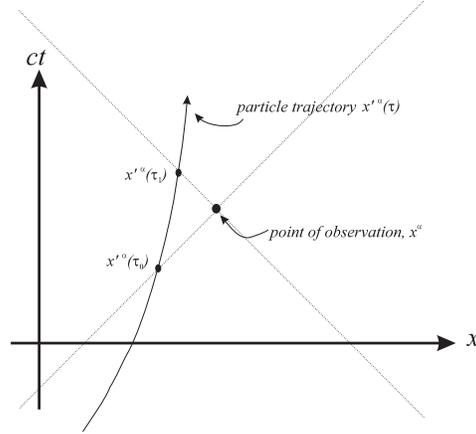


Figure 18.2: Meaning of the retarded potentials

To complete the calculation, we must evaluate $\dot{f}(\tau)$ and demonstrate manifest covariance. Take the derivative of $f(\tau)$:

$$\frac{df}{d\tau} = \frac{d}{d\tau} |\vec{r} - \mathbf{r}(\tau)| + c \frac{dt(\tau)}{d\tau} \quad (18.1.10)$$

and notice that,

$$\begin{aligned} \frac{d}{d\tau} |\vec{r} - \mathbf{r}(\tau)| &= \frac{d}{d\tau} \sqrt{\vec{r}^2 + \mathbf{r}^2(\tau) - 2\vec{r} \cdot \mathbf{r}(\tau)} \\ &= -\frac{\vec{r} - \mathbf{r}(\tau)}{|\vec{r} - \mathbf{r}(\tau)|} \cdot \frac{d\mathbf{r}(\tau)}{d\tau} = -\hat{n} \cdot \frac{d\mathbf{r}(\tau)}{d\tau} \end{aligned} \quad (18.1.11)$$

where \hat{n} is the unit vector that points in the direction of $\vec{r} - \mathbf{r}(\tau)$. Therefore,

$$\frac{df}{d\tau} = c \frac{dt(\tau)}{d\tau} - \hat{n} \cdot \frac{d\mathbf{r}(\tau)}{d\tau} \quad (18.1.12)$$

and

$$\dot{f}(\tau) |\vec{r} - \mathbf{r}(\tau)| = c |\vec{r} - \mathbf{r}(\tau)| \frac{dt(\tau)}{d\tau} - (\vec{r} - \mathbf{r}(\tau)) \cdot \frac{d\mathbf{r}(\tau)}{d\tau} \quad (18.1.13)$$

We know that $U^0 = dt(\tau)/d\tau$ and $U^i = dx^i/d\tau$ define the particle velocity 4-vector, so, if we define the (null, $R^\mu R_\mu = 0$) vector R^μ according to

$$R^0(\vec{r}, \tau) = \frac{1}{c} |\vec{r} - \mathbf{r}(\tau)|, \quad \vec{R}(\vec{r}, \tau) = \vec{r} - \mathbf{r}(\tau) \quad (18.1.14)$$

then

$$|\dot{f}(\tau)| |\vec{r} - \mathbf{r}(\tau)| = -U(\tau) \cdot R(\vec{r}, \tau) \quad (18.1.15)$$

and (18.1.9) can be put in the form

$$A_\mu(\vec{r}, t) = -\frac{e}{4\pi c\epsilon_o} \frac{U_\mu(\tau_o)}{U(\tau_o) \cdot R(\vec{r}, \tau_o)} \quad (18.1.16)$$

where $U \cdot R = U^\alpha R_\alpha$ is the 4-d scalar product between the velocity 4-vector and $R^\mu(\vec{r}, \tau)$. $R^\mu(\vec{r}, \tau_o)$ represents the segment of the observation point's light cone between its intersection with the particle trajectory in the past and the point itself. It is directed from the intersection point to the observation point.

Whereas the formula in (18.1.16) is manifestly covariant, it is subtle and not as simple to interpret as it might appear. Remember that $\tau_o(\vec{r}, t)$ is a root of the equation $f(\tau) = 0$, which is verified only on the light-cone of the observation point, while the Θ -function requires us to select the root that lies in the past of the observation point. The situation is shown schematically in figure 26. The particle trajectory intersects the light cone of the observation point in two points, one in the past (at τ_o) and the other in the future (at τ_1). Only the past intersection contributes to the potentials. Therefore the potentials $A_\mu(\vec{r}, t)$ above are called “retarded potentials” (in this case, of a single charge in arbitrary motion). The solution is formally exact.

Let us now try to uncover the subtleties of (18.1.16). Begin by determining the potential in a particularly trivial case: that of a point charge situated at the origin of coordinates. Thus $\mathbf{r} = 0$ and $\mathbf{t} = \tau$, giving $\dot{f}(\tau) = c$. Also, $U_0 = -c^2$, $U_i = 0$ and $U \cdot R = -c|\vec{r}|$, so clearly $A_i(\vec{r}) = 0$ and

$$-A_0(\vec{r}) = \phi(\vec{r}) = \frac{e}{4\pi\epsilon_o|\vec{r}|} \quad (18.1.17)$$

which we recognize as the electrostatic potential of a single point charge. A boost will give the potentials seen by an observer whose velocity is \vec{v} relative to the charge and we can calculate the fields by simply taking derivatives of these.¹ In the general case, (18.1.16) can be put in a more recognizable, but non-covariant form. Use

$$U_0 = -\frac{c^2}{\sqrt{1 - \vec{v}^2/c^2}}, \quad \vec{U} = \frac{\vec{v}}{\sqrt{1 - \vec{v}^2/c^2}} \quad (18.1.18)$$

then

$$A_\mu(\vec{r}, t) = \frac{e}{4\pi c\epsilon_o} \frac{v_\mu(\tau_o)}{v(\tau_o) \cdot R(\vec{r}, \tau_o)} \quad (18.1.19)$$

¹Problem: Show that one recovers the transformations for \vec{E} and \vec{B} derived in Part I, noting that in this case $\vec{B} = 0$ because the charge is static in this frame.

where $v^\mu = (1, \vec{v}) = dx^\mu/dt$. Separating the time and space components we get

$$\begin{aligned}
 -A_0(\vec{r}, t) = \phi(\vec{r}, t) &= \frac{e}{4\pi\epsilon_o} \frac{1}{(1 - \frac{\vec{v}(\tau_o)}{c} \cdot \hat{n})|\vec{r} - \mathbf{r}(\tau_o)|} \\
 &= \left[\frac{e}{4\pi\epsilon_o(1 - \frac{\vec{v}}{c} \cdot \hat{n})R} \right]_{\text{ret}} \\
 \vec{A}(\vec{r}, t) &= \frac{e}{4\pi\epsilon_o} \frac{\vec{v}(\tau_o)/c^2}{(1 - \frac{\vec{v}(\tau_o)}{c} \cdot \hat{n})|\vec{r} - \mathbf{r}(\tau_o)|} \\
 &= \left[\frac{e\vec{v}/c^2}{4\pi\epsilon_o(1 - \frac{\vec{v}}{c} \cdot \hat{n})R} \right]_{\text{ret}} \tag{18.1.20}
 \end{aligned}$$

where $R = |\vec{r} - \mathbf{r}(\tau)|$ and the square brackets with the subscript “ret” means that the quantity within the brackets is to be evaluated at the retarded time, τ_o .

18.2 The Field Strength Tensor

We are interested in computing the Maxwell field strength tensor $F_{\alpha\beta} = \partial_{[\alpha}A_{\beta]}$ (the square brackets indicate antisymmetrization). One could compute this directly from the expressions in (18.1.16), but it is less exhausting to do so from the integral expression in (18.1.6)

$$A_\mu(\vec{r}, t) = \frac{e}{4\pi c\epsilon_o} \int d\tau \frac{\Theta(t - \mathbf{t}(\tau))U_\mu(\tau)}{|\vec{r} - \mathbf{r}(\tau)|} \delta(|\vec{r} - \mathbf{r}(\tau)| - c(t - \mathbf{t}(\tau))) \tag{18.2.1}$$

We have seen that the integral on the right, although not manifestly a 4-vector, is indeed so (it does indeed give us a 4-vector as (18.1.16) shows). Before we proceed, let's put the integral in a form that explicitly shows this. Let $\tilde{x} = (\mathbf{t}(\tau), \mathbf{r}(\tau))$ and note that

$$\begin{aligned}
 \delta([x - \tilde{x}]^2) &= \delta(-c^2(t - \mathbf{t})^2 + |\vec{r} - \mathbf{r}|^2) = \delta([\vec{r} - \mathbf{r} + c(t - \mathbf{t})][\vec{r} - \mathbf{r} - c(t - \mathbf{t})]) \\
 &= \frac{1}{2c|\vec{r} - \mathbf{r}|} [\delta(|\vec{r} - \mathbf{r} + c(t - \mathbf{t})|) + \delta(|\vec{r} - \mathbf{r} - c(t - \mathbf{t})|)] \tag{18.2.2}
 \end{aligned}$$

With the Θ -function in the integral, only the second δ -function above is selected, so the expression for $A_\mu(\vec{r}, t)$ is equivalent to

$$A_\mu(\vec{r}, t) = \frac{e}{2\pi c\epsilon_o} \int d\tau \Theta(t - \mathbf{t}(\tau))U_\mu(\tau)\delta([x - \tilde{x}(\tau)]^2) \tag{18.2.3}$$

which is manifestly covariant as desired. Now let us take the required derivatives to obtain $F_{\mu\nu}$. A derivative w.r.t. x^ν will act on two terms in the integrand, *viz.*, the Θ -function

and the δ -function. Only the time derivative will act on the Θ -function. Now we know that $\Theta'(x) = \delta(x)$, so the (time) derivative, acting on Θ will constrain $t = \mathbf{t}(\tau)$, which would restrain the delta function to $\delta(|\vec{r} - \mathbf{r}(\tau)|^2)$. This, in turn, is supported only when the particle trajectory intersects the point of observation. Exclude this possibility from consideration, then we might ignore derivatives of $\Theta(t - \mathbf{t}(\tau))$. We need to consider only one term: the derivative of the δ -function. Thus we arrive at

$$F_{\mu\nu} = \partial_{[\mu} A_{\nu]}(\vec{r}, t) = \frac{e}{2\pi\epsilon_0} \int d\tau \Theta(t - \mathbf{t}(\tau)) \partial_{[\mu} \delta([x - \tilde{x}(\tau)]^2) U_{\nu]}(\tau), \quad (18.2.4)$$

but what is the derivative of the δ -function? We have ($g = [x - \tilde{x}(\tau)]^2$)

$$\begin{aligned} \partial_{\mu} \delta([x - \tilde{x}(\tau)]^2) &= \partial_{\mu} g \frac{d}{dg} \delta(g) = \partial_{\mu} g \frac{d\tau}{dg} \frac{d}{d\tau} \delta(g) = \frac{\partial_{\mu} g}{\dot{g}} \frac{d}{d\tau} \delta(g) \\ &= -\frac{[x_{\mu} - \tilde{x}_{\mu}(\tau)]}{U \cdot [x - \tilde{x}(\tau)]} \frac{d}{d\tau} \delta([x - \tilde{x}(\tau)]^2) \end{aligned} \quad (18.2.5)$$

so we get

$$F_{\mu\nu} = \partial_{[\mu} A_{\nu]}(\vec{r}, t) = -\frac{e}{2\pi\epsilon_0} \int d\tau \Theta(t - \mathbf{t}(\tau)) \frac{d}{d\tau} \delta([x - \tilde{x}(\tau)]^2) \frac{[x - \tilde{x}(\tau)]_{[\mu} U_{\nu]}(\tau)}{U \cdot [x - \tilde{x}(\tau)]}. \quad (18.2.6)$$

Perform an integration by parts to get

$$F_{\mu\nu} = \partial_{[\mu} A_{\nu]}(\vec{r}, t) = \frac{e}{2\pi\epsilon_0} \int d\tau \frac{d}{d\tau} \left[\frac{[x - \tilde{x}(\tau)]_{[\mu} U_{\nu]}(\tau)}{U \cdot [x - \tilde{x}(\tau)]} \right] \Theta(t - \mathbf{t}(\tau)) \delta([x - \tilde{x}(\tau)]^2). \quad (18.2.7)$$

The last integral is carried out in exactly the same manner as before: simply replace

$$\Theta(t - \mathbf{t}(\tau)) \delta([x - \tilde{x}(\tau)]^2) \rightarrow \Theta(t - \mathbf{t}(\tau)) \frac{1}{2c} \delta(|\vec{r} - \mathbf{r}(\tau)| - c(t - \mathbf{t}(\tau))) \quad (18.2.8)$$

and repeat the steps performed earlier to get

$$F_{\mu\nu} = -\frac{e}{4\pi\epsilon_0 c} \left[\frac{1}{U \cdot R} \frac{d}{d\tau} \left[\frac{R_{[\mu} U_{\nu]}}{U \cdot R} \right] \right]_{\text{ret}}. \quad (18.2.9)$$

where we have used the fact that $R^{\mu} = x^{\mu} - \tilde{x}^{\mu}(\tau)$ in the end.²

²This equality follows from the fact that the δ -function requires $(x - \tilde{x})^0 = t - \mathbf{t}(\tau) = \frac{1}{c} |\vec{r} - \mathbf{r}(\tau)| = R^0$ and $(x - \tilde{x})^i = (r - \mathbf{r}(\tau))^i = R^i$.

18.3 The Fields

Finally, from Maxwell's field strength tensor we want to extract the physical fields. To do so we must, of course, evaluate the derivative w.r.t. the particle proper time, τ . Thus,

$$\frac{d}{d\tau} \left[\frac{R_{[\mu} U_{\nu]}}{U \cdot R} \right] = \frac{(U \cdot R)(\dot{R}_{[\mu} U_{\nu]} + R_{[\mu} \dot{U}_{\nu]}) - R_{[\mu} U_{\nu]}(\dot{U} \cdot R + U \cdot \dot{R})}{(U \cdot R)^2} \quad (18.3.1)$$

Now $\dot{R}_\mu = -U_\mu$, $U^2 = -c^2$ and \dot{U}_μ is the acceleration. Moreover it should be clear that $U_{[\mu} U_{\nu]} = 0$. The expression for $F_{\mu\nu}$ therefore simplifies to

$$F_{\mu\nu} = -\frac{e}{4\pi\epsilon_0 c} \left[\frac{1}{U \cdot R} \frac{(U \cdot R)(R_{[\mu} \dot{U}_{\nu]}) - R_{[\mu} U_{\nu]}(\dot{U} \cdot R + c^2)}{(U \cdot R)^2} \right]_{\text{ret}} \quad (18.3.2)$$

Let us note that there is one term above that behaves as R^{-2} and two terms that behave as R^{-1} . As we have seen before, behavior of the fields as R^{-1} indicates a flux of energy across a large sphere and so is typical of radiation fields. Not coincidentally, the two terms that behave as R^{-1} are linearly proportional to \dot{U} , *i.e.*, to the acceleration. The important lesson here is that there is no energy flux at infinity (in other words, there are no radiation fields) without acceleration.

In the absence of acceleration the expression reduces to (these are sometimes referred to as the "velocity fields")

$$F_{\mu\nu} = \frac{ec}{4\pi\epsilon_0} \left[\frac{R_{[\mu} U_{\nu]}}{(U \cdot R)^3} \right]_{\text{ret}} \quad (18.3.3)$$

giving (no acceleration)

$$\begin{aligned} E_i &= F_{i0} = \frac{ec}{4\pi\epsilon_0} \left[\frac{R_{[i} U_{0]}}{(U \cdot R)^3} \right]_{\text{ret}} = \frac{ec}{4\pi\epsilon_0} \left[\frac{R_i U_0 - R_0 U_i}{(U \cdot R)^3} \right]_{\text{ret}} \\ \vec{E} &= \frac{e}{4\pi\epsilon_0 \gamma^2} \left[\frac{\hat{n} - \vec{v}/c}{\left(1 - \frac{\vec{v}}{c} \cdot \hat{n}\right)^3} \frac{1}{R^2} \right]_{\text{ret}} \end{aligned} \quad (18.3.4)$$

and

$$\vec{B} = [\hat{n} \times \vec{E}]_{\text{ret}} \quad (18.3.5)$$

The acceleration (radiation) fields may likewise be computed (the calculation is more laborious, but quite straightforward):

$$E_i = F_{i0} = -\frac{e}{4\pi\epsilon_0 c} \left[\frac{(R_{[i} \dot{U}_{0]})}{(U \cdot R)^2} - \frac{R_{[i} U_{0]}(\dot{U} \cdot R)}{(U \cdot R)^3} \right]_{\text{ret}}$$

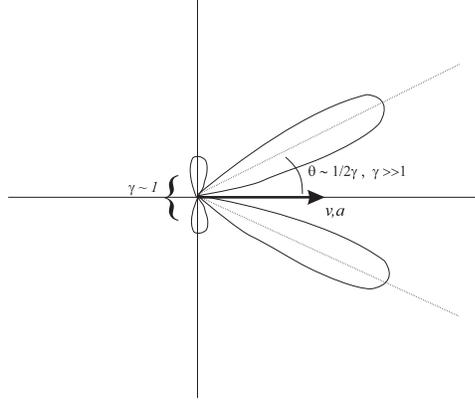


Figure 18.3: Emission from a relativistic accelerated particle

$$\begin{aligned}
 &= -\frac{e}{4\pi\epsilon_0 c} \left[\frac{(R_i \dot{U}_0 - R_0 \dot{U}_i)}{(U \cdot R)^2} - \frac{(R_i U_0 - R_0 U_i)(\dot{U} \cdot R)}{(U \cdot R)^3} \right]_{\text{ret}} \\
 \vec{E} &= \frac{e}{4\pi\epsilon_0 c^2} \left[\frac{\hat{n} \times \{(\hat{n} - \vec{v}/c) \times \vec{a}\}}{\left(1 - \frac{\vec{v}}{c} \cdot \hat{n}\right)^3 R} \right]_{\text{ret}} \quad (18.3.6)
 \end{aligned}$$

where $\vec{a} = \dot{\vec{v}} = d\vec{v}/dt$ is the acceleration (note, **not** the proper acceleration) and once again

$$\vec{B} = [\hat{n} \times \vec{E}]_{\text{ret}}, \quad (18.3.7)$$

which completes the expressions for the fields. Remember that these expressions **are exact**.

18.3.1 Example: Larmor's formula

We conclude this section by considering a simple example: the radiation from a charge whose velocity is small compared with the velocity of light in the frame in which the measurement is being performed. We will be concerned only with the radiation term and set $\vec{v}/c \approx 0$. In this limit the acceleration term is

$$\vec{E}_a \approx \frac{e}{4\pi\epsilon_0 c^2} \left[\frac{[\hat{n} \times (\hat{n} \times \vec{a})]}{R} \right]_{\text{ret}}. \quad (18.3.8)$$

From the expression for the instantaneous flux

$$c\epsilon_0 |\vec{E}_a|^2 \hat{n} \quad (18.3.9)$$

we might compute the power radiated at infinity per solid angle, using the formulas of the previous sections, to be

$$d\mathcal{J} = \frac{e^2}{4\pi\epsilon_0 c^3} [\hat{n} \times (\hat{n} \times \vec{a})]^2 d\Omega \quad (18.3.10)$$

Let θ be the angle between \hat{n} and \vec{a} , so that $\hat{n} \cdot \vec{a} = a \cos \theta$. Then

$$[\hat{n} \times (\hat{n} \times \vec{a})]^2 = (a \cos \theta \hat{n} - \vec{a})^2 = a^2 \sin^2 \theta \quad (18.3.11)$$

where $a = |\vec{a}|$, so the angular dependence of the power radiated is $\sin^2 \theta$ and the total power radiated is the integral of the power radiated per solid angle over the entire solid angle

$$\mathcal{J} = \frac{e^2}{4\pi\epsilon_0 c^3} a^2 \int_0^{2\pi} d\phi \int_0^\pi d\theta \sin^3 \theta = \frac{2}{3} \frac{e^2 a^2}{\epsilon_0 c^3} \quad (18.3.12)$$

This is the Larmor formula for the *total* power radiated by a non-relativistic, accelerated charge. Because we treated only the non-relativistic limit ($\gamma \approx 0$), the angular dependence of the radiation was simple: just $\sin^2 \theta$, where θ is measured relative to the acceleration direction, as we have seen. This is maximum in the plane perpendicular to the acceleration. For relativistic particles the angular dependence is significantly different as the velocity dependence of the acceleration fields cannot be ignored. One finds instead (see figure 27), for example for a particle accelerated in the direction of its velocity, that the peak intensity is radiated at an angle $\approx \pm \frac{1}{2\gamma}$, which defines the opening angle of a cone whose axis lies along the velocity vector and whose value can be quite small for large values of γ (extreme relativistic motion).

Appendix A

The δ -function

A.1 Introduction

Consider the electrostatic potential ϕ , which obeys

$$\vec{\nabla}^2 \phi(\vec{r}) = -\frac{\rho(\vec{r})}{\epsilon_o} \quad (\text{A.1.1})$$

in the Coulomb gauge ($\vec{\nabla} \cdot \vec{A} = 0$). In the case of a single point charge, q , we know that in the vacuum outside the source the electrostatic potential is given by

$$\phi(\vec{r}) = \frac{1}{4\pi\epsilon_o} \frac{q}{|\vec{r} - \vec{r}_0|} \quad (\text{A.1.2})$$

where \vec{r}_0 represents the position of the point charge, and $\vec{r} \neq \vec{r}_0$. Likewise, for any number, N , of discrete point charges

$$\phi(\vec{r}) = \frac{1}{4\pi\epsilon_o} \sum_{j=1}^N \frac{q_j}{|\vec{r} - \vec{r}_j|} \quad (\text{A.1.3})$$

is the potential. Of course, this is a vacuum solution, so it obeys

$$\vec{\nabla}^2 \phi(\vec{r}) = 0 \quad (\text{A.1.4})$$

Now if we had a continuous distribution instead of a discrete one, we could generalize (A.1.3) to something like

$$\phi(\vec{r}) = \int d^3\vec{r}' G(\vec{r}', \vec{r}) \rho(\vec{r}') \quad (\text{A.1.5})$$

where the integration is over all of space. Taking the three dimensional Laplacian of the r.h.s. gives

$$\vec{\nabla}^2 \phi(\vec{r}) = \int d^3\vec{r}' [\vec{\nabla}^2 G(\vec{r}', \vec{r})] \rho(\vec{r}') \quad (\text{A.1.6})$$

so that, to satisfy (A.1.1), the function $\vec{\nabla}^2 G(\vec{r}', \vec{r})$ should be a very strange object: it must have support (it must be non-vanishing) only for $\vec{r}' = \vec{r}$ and must have no support when $\vec{r}' \neq \vec{r}$. Moreover, for the integral to give a non-zero result, $\vec{\nabla}^2 G(\vec{r}', \vec{r})$ should be very large (infinite!) where it has support. Such an object we will formally call a δ -function. It is not really a well defined function but the limit of a sequence of functions as we will shortly see. We will write

$$\vec{\nabla}^2 G(\vec{r}', \vec{r}) = -\frac{1}{\epsilon_o} \delta^{(3)}(\vec{r}' - \vec{r}) \quad (\text{A.1.7})$$

where $\delta^{(3)}(\vec{r}' - \vec{r})$ is such that

$$\begin{aligned} \int d^3 \vec{r}' \delta^{(3)}(\vec{r}' - \vec{r}) &= 1 \\ \int d^3 \vec{r}' \delta^{(3)}(\vec{r}' - \vec{r}) f(\vec{r}') &= f(\vec{r}) \end{aligned} \quad (\text{A.1.8})$$

for all well-behaved functions $f(\vec{r})$.

A.2 The δ -function as a distribution

It is easier to work in one dimension and, in any case, the three dimensional δ -function can be thought of as a product of three one dimensional δ -functions:

$$\delta^{(3)}(\vec{r}' - \vec{r}) = \delta(x' - x)\delta(y' - y)\delta(z' - z). \quad (\text{A.2.1})$$

Let $x \in (a, b)$ and define the object $\delta(x' - x)$ by

$$\begin{aligned} \int_a^b dx' \delta(x' - x) &= \begin{cases} 0 & x \notin (a, b) \\ 1 & x \in (a, b) \end{cases} \\ \int_a^b dx' f(x') \delta(x' - x) &= \begin{cases} 0 & x \notin (a, b) \\ f(x) & x \in (a, b) \end{cases} \end{aligned} \quad (\text{A.2.2})$$

Normally, we'll be concerned with all of space, so the limits will be from $-\infty$ to ∞ . Let us consider some examples.

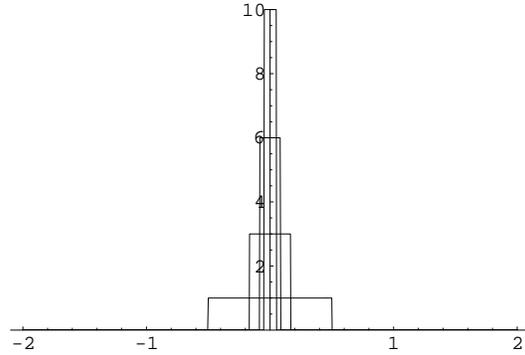


Figure A.1: The sequence of functions in (A.2.3)

A.2.1 An example

Consider the sequence of functions (see figure A1)

$$\delta_n(x) = \begin{cases} 0 & x < -\frac{1}{2n} \\ n & x \in \left(-\frac{1}{2n}, \frac{1}{2n}\right) \\ 0 & x > \frac{1}{2n} \end{cases} \quad (\text{A.2.3})$$

and note that

$$\int_{-\infty}^{\infty} \delta_n(x) dx = \int_{-\frac{1}{2n}}^{\frac{1}{2n}} n dx = 1, \quad \forall n \in \mathbb{N} \quad (\text{A.2.4})$$

while

$$\int_{-\infty}^{\infty} dx \delta_n(x) f(x) = \int_{-\frac{1}{2n}}^{\frac{1}{2n}} n f(x) dx = \frac{[g(1/2n) - g(-1/2n)]}{1/n} \quad (\text{A.2.5})$$

where $g(x)$ is the primitive of $f(x)$. Calling $\epsilon = \frac{1}{n}$ and taking the limit as $n \rightarrow \infty$ ($\epsilon \rightarrow 0$) we have

$$\lim_{n \rightarrow \infty} \int_{-\infty}^{\infty} dx \delta_n(x) f(x) = \lim_{\epsilon \rightarrow 0} \frac{[g(\epsilon/2) - g(-\epsilon/2)]}{\epsilon} = g'(0) = f(0) \quad (\text{A.2.6})$$

Thus we may define the δ -function as

$$\delta(x) = \lim_{n \rightarrow \infty} \delta_n(x) \quad (\text{A.2.7})$$

because in this limit both conditions in (A.2.2) are obeyed.

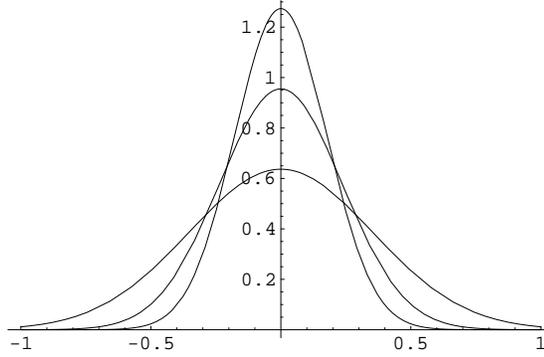


Figure A.2: The sequence of functions in (A.2.8)

A.2.2 Another example

Consider a slightly more complicated sequence of functions

$$\delta_n(x) = \frac{n}{\sqrt{\pi}} e^{-n^2 x^2} \quad (\text{A.2.8})$$

and again note that

$$\int_{-\infty}^{\infty} dx \delta_n(x) = 1, \quad \forall n \in \mathbb{N} \quad (\text{A.2.9})$$

Now let us consider

$$\int_{-\infty}^{\infty} dx \delta_n(x) f(x) = \frac{n}{\sqrt{\pi}} \int_{-\infty}^{\infty} dx e^{-n^2 x^2} f(x). \quad (\text{A.2.10})$$

Expanding $f(x)$ about the origin in a Taylor series

$$f(x) = \sum_{j=0}^{\infty} \frac{f^{(j)} x^j}{j!} \quad (\text{A.2.11})$$

our integral becomes

$$\frac{n}{\sqrt{\pi}} \sum_j \frac{f^{(j)}}{j!} \int_{-\infty}^{\infty} dx x^j e^{-n^2 x^2}. \quad (\text{A.2.12})$$

where $f^{(j)}$ is the j^{th} derivative of $f(x)$ at $x = 0$. Clearly, the only non-vanishing contributions come from even j .

$$\int_{-\infty}^{\infty} dx \delta_n(x) f(x) = \frac{n}{\sqrt{\pi}} \sum_j \frac{f^{(2j)}}{(2j)!} \int_{-\infty}^{\infty} dx x^{2j} e^{-n^2 x^2}$$

$$= \sum_j \frac{f^{(2j)}}{(2j)!} \frac{\Gamma(j + \frac{1}{2})}{\sqrt{\pi}} n^{-2j}. \quad (\text{A.2.13})$$

We see that the limit as $n \rightarrow \infty$ of the r.h.s. is just $f(0)$. Thus we also have

$$\lim_{n \rightarrow \infty} \int_{-\infty}^{\infty} dx \delta_n(x) f(x) = f(0) \quad (\text{A.2.14})$$

and we could define

$$\delta(x) = \lim_{n \rightarrow \infty} \delta_n(x), \quad \delta_n(x) = \frac{n}{\sqrt{\pi}} e^{-n^2 x^2}, \quad (\text{A.2.15})$$

thereby getting an alternative representation for the δ -function.

A.2.3 Properties

We may likewise analyze expressions such as

$$\begin{aligned} \delta(x) &= \lim_{n \rightarrow \infty} \delta_n(x), & \delta_n(x) &= \frac{n}{\pi} \frac{1}{1 + n^2 x^2} \\ \delta(x) &= \lim_{n \rightarrow \infty} \delta_n(x), & \delta_n(x) &= \frac{\sin nx}{\pi x} = \frac{1}{2\pi} \int_{-n}^n dt e^{ixt} \\ \delta(x) &= \lim_{n \rightarrow \infty} \delta_n(x), & \delta_n(x) &= \frac{\sin^2 nx}{n\pi x^2} \end{aligned} \quad (\text{A.2.16})$$

with the same results.

Here is a list of some of the more interesting properties of the δ -function. They can be proved by simply applying the defining equations.

1. $\delta(cx) = \frac{1}{|c|} \delta(x)$ (therefore, $\delta(-x) = \delta(x)$), or more generally,

$$\delta(g(x)) = \sum_j \frac{\delta(x - x_j)}{|g'(x_j)|} \quad (\text{A.2.17})$$

where x_j is a simple zero of the function $g(x)$, *i.e.*, $g(x_j) = 0$ and $g'(x_j) \neq 0$,

2. $g(x)\delta(x - x_o) = g(x_o)\delta(x - x_o)$,
3. $\int_{-\infty}^{\infty} dx \delta(x - y)\delta(x - z) = \delta(y - z)$ and
4. $\Theta'(x) = \delta(x)$, where $\Theta(x)$ is the Heaviside Θ -function.

Note, however, that the limits of the defining sequences themselves do not exist on the real line, *i.e.*, the δ -function has no meaning independently. The only meaning that can be given the object is via its defining integrals. It is a distribution.

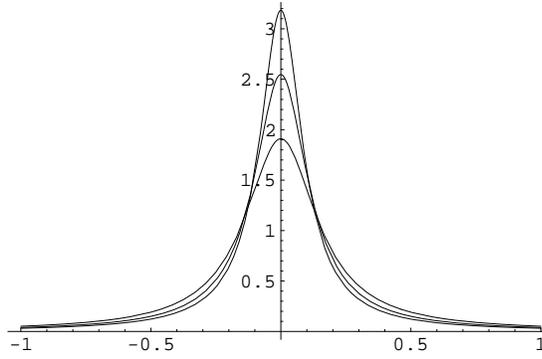


Figure A.3: $\delta_n(x) = \frac{n}{\pi} \frac{1}{1+n^2x^2}$

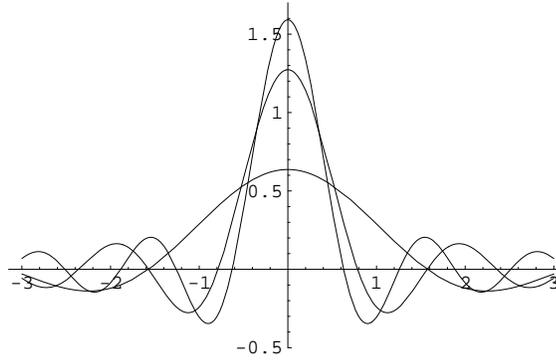


Figure A.4: $\delta_n(x) = \frac{\sin nx}{\pi x}$

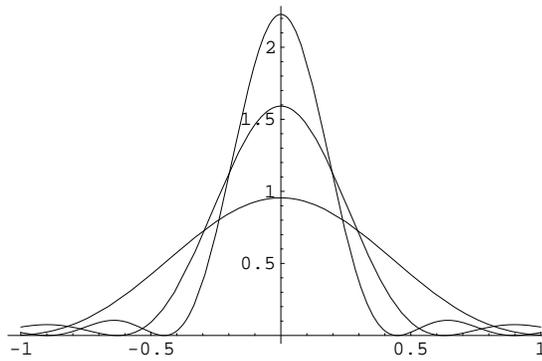


Figure A.5: $\delta_n(x) = \frac{\sin^2 nx}{n\pi x^2}$

A.3 The δ -function in curvilinear coordinates

Beginning with the δ -function in a cartesian system, we can deduce its form in a general curvilinear coordinate system by using its defining properties. By definition

$$\begin{aligned} \int d^n x' \delta^n(x' - x) &= 1 = \int d^n \xi' \left\| \frac{\partial x'}{\partial \xi'} \right\| \delta^n(x' - x) \\ \int d^n x' \delta^n(x' - x) f(x') &= f(x) = \int d^n \xi' \left\| \frac{\partial x'}{\partial \xi'} \right\| \delta^n(x' - x) f(x') \end{aligned} \quad (\text{A.3.1})$$

where $\| \cdot \|$ is the Jacobian of the transformation, f is a scalar function and, on the r.h.s., $x = x(\xi)$. But we have seen in Chapter 3 that the Jacobian can be replaced by the determinant of the metric, *i.e.*,

$$1 = \int d^n \xi' \left\| \frac{\partial x'}{\partial \xi'} \right\| \delta^n(x' - x) = \frac{1}{c} \int d^n \xi' \sqrt{-g(\xi')} \delta^n(x' - x). \quad (\text{A.3.2})$$

Now we *define* the n dimensional δ -function in a general coordinate system in the same way as we had before:

$$\begin{aligned} \int d^n \xi' \delta^n(\xi' - \xi) &= 1 \\ \int d^n \xi' f(\xi') \delta^n(\xi' - \xi) &= f(\xi). \end{aligned} \quad (\text{A.3.3})$$

Then, simply comparing the expressions above, we find

$$\delta^n(x' - x) = \frac{c}{\sqrt{-g(\xi)}} \delta^n(\xi' - \xi) \quad (\text{A.3.4})$$

Thus $cg^{-\frac{1}{2}}(\xi) \delta^n(\xi' - \xi)$ is a scalar under general coordinate transformations and $\delta^n(\xi' - \xi)$ is a scalar *density*. As examples let's write the three dimensional δ -function in spherical coordinates (in three dimensions, the factor of c does not appear of course):

$$\delta^{(3)}(\vec{r}' - \vec{r}) = \frac{1}{r^2 \sin \theta} \delta(r' - r) \delta(\theta' - \theta) \delta(\varphi' - \varphi) \quad (\text{A.3.5})$$

and in cylindrical coordinates:

$$\delta^{(3)}(\vec{r}' - \vec{r}) = \frac{1}{\rho} \delta(\rho' - \rho) \delta(\varphi' - \varphi) \delta(z' - z). \quad (\text{A.3.6})$$

Notice the density factors in each case.