611: Electromagnetic Theory II

CONTENTS

- Special relativity; Lorentz covariance of Maxwell equations
- Scalar and vector potentials, and gauge invariance
- Relativistic motion of charged particles
- Action principle for electromagnetism; energy-momentum tensor
- Electromagnetic waves; waveguides
- Fields due to moving charges
- Radiation from accelerating charges
- Antennae
- Radiation reaction
- Magnetic monopoles, duality, Yang-Mills theory

Suggested textbooks:

- J.D. Jackson, *Classical Electrodynamics*
## Contents

1 Electrodynamics and Special Relativity  
1.1 Introduction ........................................... 4  
1.2 The Lorentz Transformation .......................... 8  
1.3 An interlude on 3-vectors and suffix notation ........ 11  
1.4 4-vectors and 4-tensors ............................. 16  
1.5 Lorentz tensors ....................................... 23  
1.6 Proper time and 4-velocity .......................... 25  

2 Electrodynamics and Maxwell’s Equations .......... 27  
2.1 Natural units ........................................... 27  
2.2 Gauge potentials and gauge invariance ............... 28  
2.3 Maxwell’s equations in 4-tensor notation ............. 30  
2.4 Lorentz transformation of $\vec{E}$ and $\vec{B}$ ......... 36  
2.5 The Lorentz force .................................. 37  
2.6 Action principle for charged particles ............... 39  
2.7 Gauge invariance of the action ....................... 42  
2.8 Canonical momentum, and Hamiltonian ............... 43  

3 Particle Motion in Static Electromagnetic Fields .... 44  
3.1 Description in terms of potentials ................... 44  
3.2 Particle motion in static uniform $\vec{E}$ and $\vec{B}$ fields .. 46  

4 Action Principle for Electrodynamics ................. 51  
4.1 Invariants of the electromagnetic field ................ 51  
4.2 Action for electrodynamics .......................... 54  
4.3 Inclusion of sources ................................ 56  
4.4 Energy density and energy flux ....................... 59  
4.5 Energy-momentum tensor ............................. 61  
4.6 Energy-momentum tensor for the electromagnetic field .. 67  
4.7 Inclusion of massive charged particles ............... 71  

5 Coulomb’s Law ........................................... 73  
5.1 Potential of a point charge .......................... 73  
5.2 Electrostatic energy ................................ 75
5.3 Field of a uniformly moving charge ........................................ 76
5.4 Motion of a charge in a Coulomb potential ............................... 79

6 Electromagnetic Waves ........................................................... 84
6.1 Wave equation ........................................................................ 84
6.2 Monochromatic plane waves ..................................................... 88
6.3 Motion of a point charge in a linearly-polarised E.M. wave ......... 90
6.4 Circular and elliptical polarisation ............................................. 92
6.5 General superposition of plane waves ....................................... 94
6.6 Gauge invariance and electromagnetic fields ............................. 101
6.7 Fourier decomposition of electrostatic fields ............................. 104
6.8 Waveguides .......................................................................... 105
6.9 Resonant cavities .................................................................... 113

7 Fields Due to Accelerating Charges ............................................ 114
7.1 Retarded potentials ................................................................ 114
7.2 Lienard-Wiechert potentials ...................................................... 118
7.3 Electric and magnetic fields of a moving charge ......................... 123
7.4 Radiation by accelerated charges .............................................. 126
7.5 Applications of Larmor formula ................................................. 128
7.6 Angular distribution of the radiated power ................................. 131
7.7 Frequency distribution of radiated energy ................................. 136
7.8 Frequency spectrum for relativistic circular motion ..................... 140
7.9 Frequency spectrum for periodic motion .................................... 143
7.10 Cerenkov radiation .................................................................. 145
7.11 Thompson scattering ............................................................... 149

8 Radiating Systems ..................................................................... 151
8.1 Fields due to localised oscillating sources ................................. 151
8.2 Electric dipole radiation ............................................................ 155
8.3 Higher multipoles ..................................................................... 159
8.4 Linear antenna ......................................................................... 165

9 Electromagnetism and Quantum Mechanics .............................. 168
9.1 The Schrödinger equation and gauge transformations ................ 168
9.2 Magnetic monopoles ............................................................... 171
9.3 Dirac quantisation condition ........................................... 175

10 Local Gauge Invariance and Yang-Mills Theory .......... 176
   10.1 Relativistic quantum mechanics ................................. 176
   10.2 Yang-Mills theory .................................................. 180
1 Electrodynamics and Special Relativity

1.1 Introduction

The final form of Maxwell’s equations describing the electromagnetic field had been established by 1865. Although this was forty years before Einstein formulated his theory of special relativity, the Maxwell equations are, remarkably, fully consistent with the special relativity. The Maxwell theory of electromagnetism is the first, and in many ways the most important, example of what is known as a classical relativistic field theory.

Our emphasis in this course will be on establishing the formalism within which the relativistic invariance of electrodynamics is made manifest, and thereafter exploring the relativistic features of the theory.

In Newtonian mechanics, the fundamental laws of physics, such as the dynamics of moving objects, are valid in all inertial frames (i.e. all non-accelerating frames). If \( S \) is an inertial frame, then the set of all inertial frames comprises all frames that are in uniform motion relative to \( S \). Suppose that two inertial frames \( S \) and \( S' \), are parallel, and that their origins coincide at \( t = 0 \). If \( S' \) is moving with uniform velocity \( \vec{v} \) relative to \( S \), then a point \( P \) with position vector \( \vec{r} \) with respect to \( S \) will have position vector \( \vec{r}' \) with respect to \( S' \), where

\[
\vec{r}' = \vec{r} - \vec{v} t. \tag{1.1}
\]

Of course, it is always understood in Newtonian mechanics that time is absolute, and so the times \( t \) and \( t' \) measured by observers in the frames \( S \) and \( S' \) are the same:

\[
t' = t. \tag{1.2}
\]

The transformations (1.1) and (1.2) form part of what is called the **Galilean Group**. The full Galilean group includes also rotations of the spatial Cartesian coordinate system, so that we can define

\[
\vec{r}' = \mathbf{M} \cdot \vec{r} - \vec{v} t, \quad t' = t, \tag{1.3}
\]

where \( \mathbf{M} \) is an orthogonal \( 3 \times 3 \) constant matrix acting by matrix multiplication on the components of the position vector:

\[
\vec{r} \leftrightarrow \begin{pmatrix} x \\ y \\ z \end{pmatrix}, \quad \mathbf{M} \cdot \vec{r} \leftrightarrow \mathbf{M} \begin{pmatrix} x \\ y \\ z \end{pmatrix}, \tag{1.4}
\]

where \( \mathbf{M}^T \mathbf{M} = 1 \), and \( \mathbf{M}^T \) denotes the transpose of the matrix \( \mathbf{M} \).
Returning to our simplifying assumption that the two frames are parallel, i.e. that \( M = 1 \), it follows that if a particle having position vector \( \vec{r} \) in \( S \) moves with velocity \( \vec{u} = d\vec{r}/dt \), then its velocity \( \vec{u}' = d\vec{r}'/dt \) as measured with respect to the frame \( S' \) is given by

\[
\vec{u}' = \vec{u} - \vec{v}.
\]  

(1.5)

Suppose, for example, that \( \vec{v} \) lies along the \( x \) axis of \( S \); i.e. that \( S' \) is moving along the \( x \) axis of \( S \) with speed \( v = |\vec{v}| \). If a beam of light were moving along the \( x \) axis of \( S \) with speed \( c \), then the prediction of Newtonian mechanics and the Galilean transformation would therefore be that in the frame \( S' \), the speed \( c' \) of the light beam would be

\[
c' = c - v.
\]  

(1.6)

Of course, as is well known, this contradicts experiment. As far as we can tell, with experiments of ever-increasing accuracy, the true state of affairs is that the speed of the light beam is the same in all inertial frames. Thus the predictions of Newtonian mechanics and the Galilean transformation are falsified by experiment.

Of course, it should be emphasised that the discrepancies between experiment and the Galilean transformations are rather negligible if the relative speed \( v \) between the two inertial frames is of a typical “everyday” magnitude, such as the speed of a car or a plane. But if \( v \) begins to become appreciable in comparison to the speed of light, then the discrepancy becomes appreciable too.

By contrast, it turns out that Maxwell’s equations of electromagnetism do predict a constant speed of light, independent of the choice of inertial frame. To be precise, let us begin with the free-space Maxwell’s equations,

\[
\begin{align*}
\nabla \cdot \vec{E} &= \frac{1}{\epsilon_0} \rho, & \nabla \times \vec{B} - \mu_0 \epsilon_0 \frac{\partial \vec{E}}{\partial t} &= \mu_0 \vec{J}, \\
\nabla \cdot \vec{B} &= 0, & \nabla \times \vec{E} + \frac{\partial \vec{B}}{\partial t} &= 0,
\end{align*}
\]

(1.7)

where \( \vec{E} \) and \( \vec{B} \) are the electric and magnetic fields, \( \rho \) and \( \vec{J} \) are the charge density and current density, and \( \epsilon_0 \) and \( \mu_0 \) are the permittivity and permeability of free space.\(^1\)

To see the electromagnetic wave solutions, we can consider a region of space where there are no sources, i.e. where \( \rho = 0 \) and \( \vec{J} = 0 \). Then we shall have

\[
\nabla \times (\nabla \times \vec{E}) = -\frac{\partial}{\partial t} \nabla \times \vec{B} = -\mu_0 \epsilon_0 \frac{\partial^2 \vec{E}}{\partial t^2}.
\]

(1.8)

\(^1\)The equations here are written using the system of units known as SI, which could be said to stand for “Super Inconvenient.” In these units, the number of unnecessary dimensionful “fundamental constants” is maximised. We shall pass speedily to more convenient units a little bit later.
But using the vector identity $\nabla \times (\nabla \times \vec{E}) = \nabla (\nabla \cdot \vec{E}) - \nabla^2 \vec{E}$, it follows from $\nabla \cdot \vec{E} = 0$ that the electric field satisfies the wave equation

$$\nabla^2 \vec{E} - \mu_0 \epsilon_0 \frac{\partial^2 \vec{E}}{\partial t^2} = 0.$$  

(1.9)

This admits plane-wave solutions of the form

$$\vec{E} = \vec{E}_0 e^{i(\vec{k} \cdot \vec{r} - \omega t)},$$

(1.10)

where $\vec{E}_0$ and $\vec{k}$ are constant vectors, and $\omega$ is also a constant, where

$$k^2 = \mu_0 \epsilon_0 \omega^2.$$  

(1.11)

Here $k$ means $|\vec{k}|$, the magnitude of the wave-vector $\vec{k}$. Thus we see that the waves travel at speed $c$ given by

$$c = \frac{\omega}{k} = \frac{1}{\sqrt{\mu_0 \epsilon_0}}.$$  

(1.12)

Putting in the numbers, this gives $c \approx 3 \times 10^8$ metres per second, i.e. the familiar speed of light.

A similar calculation shows that the magnetic field $\vec{B}$ also satisfies an identical wave equation, and in fact $\vec{B}$ and $\vec{E}$ are related by

$$\vec{B} = \frac{1}{\omega} \vec{k} \times \vec{E}.$$  

(1.13)

The situation, then, is that if the Maxwell equations (1.7) hold in a given frame of reference, then they predict that the speed of light will be $c \approx 3 \times 10^8$ metres per second in that frame. Therefore, if we assume that the Maxwell equations hold in all inertial frames, then they predict that the speed of light will have that same value in all inertial frames. Since this prediction is in agreement with experiment, we can reasonably expect that the Maxwell equations will indeed hold in all inertial frames. Since the prediction contradicts the implications of the Galilean transformations, it follows that the Maxwell equations are not invariant under Galilean transformations. This is just as well, since the Galilean transformations are wrong!

In fact, as we shall see, the transformations that correctly describe the relation between observations in different inertial frames in uniform motion are the Lorentz Transformations of Special Relativity. Furthermore, even though the Maxwell equations were written down in the pre-relativity days of the nineteenth century, they are in fact perfectly invariant\(^2\) under

\(^2\)Strictly, as will be explained later, we should say covariant rather than invariant.
the Lorentz transformations. No further modification is required in order to incorporate Maxwell’s theory of electromagnetism into special relativity.

However, the Maxwell equations as they stand, written in the form given in equation (1.7), do not look manifestly covariant with respect to Lorentz transformations. This is because they are written in the language of 3-vectors. To make the Lorentz transformations look nice and simple, we should instead express them in terms of 4-vectors, where the extra component is associated with the time direction.

Actually, before proceeding it is instructive to take a step back and look at what the Maxwell equations actually looked like in Maxwell’s 1865 paper *A Dynamical Theory of the Electromagnetic Field*, published in the Philosophical Transactions of the Royal Society of London. It must be recalled that in 1865 three-dimensional vectors had not yet been invented, and so everything was written out explicitly in terms of the $x$, $y$ and $z$ components. To make matters worse, Maxwell used a different letter of the alphabet for each component of each field. In terms of the now-familiar electric vector fields $\vec{E}$, $\vec{D}$, the magnetic fields $\vec{B}$, $\vec{H}$, the current density $\vec{J}$ and the charge density $\rho$, Maxwell’s chosen names for the components were

$$\vec{E} = (P, Q, R), \quad \vec{D} = (f, g, h),$$
$$\vec{B} = (F, G, H), \quad \vec{H} = (\alpha, \beta, \gamma),$$
$$\vec{J} = (p, q, r), \quad \rho = e.$$  \hspace{1cm} (1.14)

Thus the Maxwell equations that we now write rather compactly as

$$\nabla \cdot \vec{D} = 4\pi \rho, \quad \nabla \cdot \vec{B} = 0,$$
$$\nabla \times \vec{H} - \frac{\partial \vec{D}}{\partial t} = 4\pi \vec{J}, \quad \nabla \times \vec{E} + \frac{\partial \vec{B}}{\partial t} = 0,$$  \hspace{1cm} (1.15)

took, in 1865, the highly inelegant forms

$$\frac{\partial f}{\partial x} + \frac{\partial g}{\partial y} + \frac{\partial h}{\partial z} = 4\pi e,$$
$$\frac{\partial F}{\partial x} + \frac{\partial G}{\partial y} + \frac{\partial H}{\partial z} = 0.$$  \hspace{1cm} (1.17)

---

3Vectors were invented independently by Josiah Willard Gibbs, and Oliver Heaviside, around the end of the 19th century.

4Here, we are writing the equations in the so-called “Natural Units,” which we shall be using throughout this course.
for the two equations in (1.15), and

\[
\begin{align*}
\frac{\partial \gamma}{\partial y} - \frac{\partial \beta}{\partial z} - \frac{\partial f}{\partial t} &= 4\pi p, \\
\frac{\partial \alpha}{\partial z} - \frac{\partial \gamma}{\partial x} - \frac{\partial g}{\partial t} &= 4\pi q, \\
\frac{\partial \beta}{\partial x} - \frac{\partial \alpha}{\partial y} - \frac{\partial h}{\partial t} &= 4\pi r, \\
\frac{\partial R}{\partial y} - \frac{\partial Q}{\partial z} + \frac{\partial F}{\partial t} &= 0, \\
\frac{\partial P}{\partial z} - \frac{\partial R}{\partial x} + \frac{\partial G}{\partial t} &= 0, \\
\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} + \frac{\partial H}{\partial t} &= 0, \\
\end{align*}
\]

(1.18)

for the two vector-valued equations in (1.16). Not only does Maxwell’s way of writing his
equations, in (1.17) and (1.18), look like a complete mess, but it also completely fails to
make manifest the familiar fact that the equations are symmetric under arbitrary rotations
of the three-dimensional \((x, y, z)\) coordinate system. Of course in the 3-vector notation of
(1.15) and (1.16) this rotational symmetry is completely manifest; that is precisely what
the vectors notation was invented for, to make manifest the rotational symmetry of three-
dimensional equations like the Maxwell equations. The symmetry is, of course, actually
there in Maxwell’s equations (1.17) and (1.18), but it is completely obscure and non-obvious.

So the moral of the story is one not only wants equations that have the nice symmetries,
but one wants to write them in a notation that makes these symmetries manifest. It is
worth bearing this in mind when we pursue our goal of re-writing the Maxwell equations in
a notation that does even more, and makes their symmetry under Lorentz transformations
manifest. In order to give a nice elegant treatment of the Lorentz transformation properties
of the Maxwell equations, we should first therefore reformulate special relativity in terms
of 4-vectors and 4-tensors. Since there are many different conventions on offer in the mar-
ketplace, we shall begin with a review of special relativity in the notation that we shall be
using in this course.

### 1.2 The Lorentz Transformation

The derivation of the Lorentz transformation follows from Einstein’s two postulates:

- The laws of physics are the same for all inertial observers.
- The speed of light is the same for all inertial observers.
To derive the Lorentz transformation, let us suppose that we have two inertial frames \(S\) and \(S'\), whose origins coincide at time zero, that is to say, at \(t = 0\) in the frame \(S\), and at \(t' = 0\) in the frame \(S'\). If a flash of light is emitted at the origin at time zero, then it will spread out over a spherical wavefront given by

\[
x^2 + y^2 + z^2 - c^2t^2 = 0
\]  
(1.19)

in the frame \(S\), and by

\[
x'^2 + y'^2 + z'^2 - c^2t'^2 = 0
\]  
(1.20)

in the frame \(S'\). Note that, following the second of Einstein’s postulates, we have used the same speed of light \(c\) for both inertial frames. Our goal is to derive the relation between the coordinates \((x, y, z, t)\) and \((x', y', z', t')\) in the two inertial frames.

Consider for simplicity the case where \(S'\) is parallel to \(S\), and moves along the \(x\) axis with velocity \(v\). Clearly we must have

\[
y' = y, \quad z' = z.
\]  
(1.21)

Furthermore, the transformation between \((x, t)\) and \((x', t')\) must be a linear one, since otherwise it would not be translation-invariant or time-translation invariant. Thus we may say that

\[
x' = Ax + Bt, \quad t' = Cx + Dt,
\]  
(1.22)

for constants \(A\), \(B\), \(C\) and \(D\) to be determined.

Now, if \(x' = 0\), this must, by definition, correspond to the equation \(x = vt\) in the frame \(S\), and so from the first equation in (1.22) we have \(B = -Av\), and so we have

\[
x' = A(x - vt).
\]  
(1.23)

By the same token, if we exchange the roles of the primed and the unprimed frames, and consider the origin \(x = 0\) for the frame \(S\), then this will correspond to \(x' = -vt'\) in the frame \(S'\). (If the origin of the frame \(S'\) moves along \(x\) in the frame \(S\) with velocity \(v\), then the origin of the frame \(S\) must be moving along \(x'\) in the frame \(S'\) with velocity \(-v\).) It follows that we must have

\[
x = A(x' + vt').
\]  
(1.24)

Note that it must be the same constant \(A\) in both these equations, since the two really just correspond to reversing the direction of the \(x\) axis, and the physics must be the same for the two cases.
Now we bring in the postulate that the speed of light is the same in the two frames, so if we have \(x = ct\) then this must imply \(x' = ct'\). Solving the resulting two equations

\[
ct' = A(c - v)t, \quad ct = A(c + v)t'
\]

for \(A\), we obtain

\[
A = \frac{1}{\sqrt{1 - v^2/c^2}}.
\]

Solving \(x^2 - c^2t^2 = x'^2 - c^2t'^2\) for \(t'\), after using (1.23), we find \(t'^2 = A^2(t - vx/c^2)^2\) and hence

\[
t' = A(t - \frac{v}{c^2}x).
\]

(We must choose the positive square root since it must reduce to \(t' = +t\) if the velocity \(v\) goes to zero.) At this point we shall change the name of the constant \(A\) to the conventional one \(\gamma\), and thus we arrive at the Lorentz transformation

\[
x' = \gamma(x - vt), \quad y' = y, \quad z' = z, \quad t' = \gamma(t - \frac{v}{c^2}x),
\]

where

\[
\gamma = \frac{1}{\sqrt{1 - v^2/c^2}},
\]

in the special case where \(S'\) is moving along the \(x\) direction with velocity \(v\).

At this point, for notational convenience, we shall introduce the simplification of working in a system of units in which the speed of light is set equal to 1. We can do this because the speed of light is the same for all inertial observers, and so we may as well choose to measure length in terms of the time it takes for light \textit{in vacuo} to traverse the distance. In fact, the metre is nowadays \textit{defined} to be the distance travelled by light \textit{in vacuo} in \(1/299,792,458\) of a second. By making the small change of taking the light-second as the basic unit of length, rather than the \(1/299,792,458'\)th of a light-second, we end up with a system of units in which \(c = 1\). Alternatively, we could measure time in “light metres,” where the unit is the time taken for light to travel 1 metre. In these units, the Lorentz transformation (1.28) becomes

\[
x' = \gamma(x - vt), \quad y' = y, \quad z' = z, \quad t' = \gamma(t - vx),
\]

where

\[
\gamma = \frac{1}{\sqrt{1 - v^2}}.
\]

It will be convenient to generalise the Lorentz transformation (1.30) to the case where the frame \(S'\) is moving with (constant) velocity \(\vec{v}\) in an arbitrary direction, rather than
specifically along the $x$ axis. It is rather straightforward to do this. We know that there is a complete rotational symmetry in the three-dimensional space parameterised by the $(x, y, z)$ coordinate system. Therefore, if we can first rewrite the special case described by (1.30) in terms of 3-vectors, where the 3-vector velocity $\vec{v}$ happens to be simply $\vec{v} = (v, 0, 0)$, then generalisation will be immediate. It is easy to check that with $\vec{v}$ taken to be $(v, 0, 0)$, the Lorentz transformation (1.30) can be written as

$$
\vec{r}' = \vec{r} + \frac{\gamma - 1}{v^2} (\vec{v} \cdot \vec{r}) \vec{v} - \gamma \vec{v} t, \quad t' = \gamma (t - \vec{v} \cdot \vec{r}),
$$

with $\gamma = (1 - v^2)^{-1/2}$ and $v \equiv |\vec{v}|$, and with $\vec{r} = (x, y, z)$. Since these equations are manifestly covariant under 3-dimensional spatial rotations (i.e. they are written entirely in a 3-vector notation), it must be that they are the correct form of the Lorentz transformations for an arbitrary direction for the velocity 3-vector $\vec{v}$.

The Lorentz transformations (1.32) are what are called the pure boosts. It is easy to check that they have the property of preserving the spherical light-front condition, in the sense that points on the expanding spherical shell given by $r^2 = t^2$ of a light-pulse emitted at the origin at $t = 0$ in the frame $S$ will also satisfy the equivalent condition $r'^2 = t'^2$ in the primed reference frame $S'$. (Note that $r^2 = x^2 + y^2 + z^2$.) In fact, a stronger statement is true: The Lorentz transformation (1.32) satisfies the equation

$$
x^2 + y^2 + z^2 - t^2 = x'^2 + y'^2 + z'^2 - t'^2.
$$

1.3 An interlude on 3-vectors and suffix notation

Before describing the 4-dimensional spacetime approach to special relativity, it may be helpful to give a brief review of some analogous properties of 3-dimensional Euclidean space, and Cartesian vector analysis.

Consider a 3-vector $\vec{A}$, with $x$, $y$ and $z$ components denoted by $A_1$, $A_2$ and $A_3$ respectively. Thus we may write

$$
\vec{A} = (A_1, A_2, A_3).
$$

It is convenient then to denote the set of components by $A_i$, for $i = 1, 2, 3$.

The scalar product between two vectors $\vec{A}$ and $\vec{B}$ is given by

$$
\vec{A} \cdot \vec{B} = A_1 B_1 + A_2 B_2 + A_3 B_3 = \sum_{i=1}^{3} A_i B_i.
$$

This expression can be written more succinctly using the Einstein Summation Convention. The idea is that when writing valid expressions using vectors, or more generally tensors,
on every occasion that a summation of the form \( \sum_{i=1}^{3} \) is performed, the summand is an expression in which the summation index \( i \) occurs exactly twice. Furthermore, there will be no occasion when an index occurs exactly twice in a given term and a sum over \( i \) is not performed. Therefore, we can abbreviate the writing by simply omitting the explicit summation symbol, since we know as soon as we see an index occurring exactly twice in a term of an equation that it must be accompanied by a summation symbol. Thus we can abbreviate (1.35) and just write the scalar product as

\[ \vec{A} \cdot \vec{B} = A_i B_i. \]  

(1.36)

The index \( i \) here is called a “dummy suffix.” It is just like a local summation variable in a computer program; it doesn’t matter if it is called \( i \), or \( j \) or anything else, as long as it doesn’t clash with any other index that is already in use.

The next concept to introduce is the Kronecker delta tensor \( \delta_{ij} \). This is defined by

\[ \delta_{ij} = 1 \quad \text{if} \quad i = j, \quad \delta_{ij} = 0 \quad \text{if} \quad i \neq j, \]  

(1.37)

Thus

\[ \delta_{11} = \delta_{22} = \delta_{33} = 1, \quad \delta_{12} = \delta_{13} = \cdots = 0. \]  

(1.38)

Note that \( \delta_{ij} \) is a symmetric tensor: \( \delta_{ij} = \delta_{ji} \). The Kronecker delta clearly has the replacement property

\[ A_i = \delta_{ij} A_j, \]  

(1.39)

since by (1.37) the only non-zero term in the summation over \( j \) is the term when \( j = i \).

Now consider the vector product \( \vec{A} \times \vec{B} \). We have

\[ \vec{A} \times \vec{B} = (A_2 B_3 - A_3 B_2, A_3 B_1 - A_1 B_3, A_1 B_2 - A_2 B_1). \]  

(1.40)

To write this using index notation, we first define the 3-index totally-antisymmetric tensor \( \epsilon_{ijk} \). Total antisymmetry means that the tensor changes sign if any pair of indices is swapped. For example

\[ \epsilon_{ijk} = -\epsilon_{ikj} = -\epsilon_{jik} = -\epsilon_{kji}. \]  

(1.41)

Given this total antisymmetry, we actually only need to specify the value of one non-zero component in order to pin down the definition completely. We shall define \( \epsilon_{123} = +1 \). From the total antisymmetry, it then follows that

\[ \epsilon_{123} = \epsilon_{231} = \epsilon_{312} = +1, \quad \epsilon_{132} = \epsilon_{321} = \epsilon_{213} = -1, \]  

(1.42)
with all other components vanishing.

It is now evident that in index notation, the \( i \)'th component of the vector product \( \vec{A} \times \vec{B} \) can be written as

\[
(\vec{A} \times \vec{B})_i = \epsilon_{ijk} A_j B_k .
\]  

(1.43)

For example, the \( i = 1 \) component (the \( x \) component) is given by

\[
(\vec{A} \times \vec{B})_1 = \epsilon_{1jk} A_j B_k = \epsilon_{123} A_2 B_3 + \epsilon_{132} A_3 B_2 = A_2 B_3 - A_3 B_2 ,
\]  

(1.44)

in agreement with the \( x \)-component given in (1.40).

Now, let us consider the vector triple product \( \vec{A} \times (\vec{B} \times \vec{C}) \). The \( i \) component is therefore given by

\[
[\vec{A} \times (\vec{B} \times \vec{C})]_i = \epsilon_{ijk} \epsilon_{k\ell m} A_\ell B_\ell C_m .
\]  

(1.45)

For convenience, we may cycle the indices on the second \( \epsilon \) tensor around and write this as

\[
[\vec{A} \times (\vec{B} \times \vec{C})]_i = \epsilon_{ijk} \epsilon_{\ell mk} A_\ell B_\ell C_m .
\]  

(1.46)

There is an extremely useful identity, which can be proved simply by considering all possible values of the free indices \( i, j, \ell, m \):

\[
\epsilon_{ijk} \epsilon_{\ell mk} = \delta_{i\ell} \delta_{jm} - \delta_{im} \delta_{j\ell} .
\]  

(1.47)

Using this in (1.46), we have

\[
[\vec{A} \times (\vec{B} \times \vec{C})]_i = (\delta_{i\ell} \delta_{jm} - \delta_{im} \delta_{j\ell}) A_\ell B_\ell C_m ,
\]

\[
= \delta_{i\ell} \delta_{jm} A_\ell B_\ell C_m - \delta_{im} \delta_{j\ell} A_\ell B_\ell C_m ,
\]

\[
= B_i A_j C_j - C_i A_j B_j ,
\]

\[
= (\vec{A} \cdot \vec{C}) B_i - (\vec{A} \cdot \vec{B}) C_i .
\]  

(1.48)

In other words, we have proven that

\[
\vec{A} \times (\vec{B} \times \vec{C}) = (\vec{A} \cdot \vec{C}) \vec{B} - (\vec{A} \cdot \vec{B}) \vec{C} .
\]  

(1.49)

It is useful also to apply the index notation to the gradient operator \( \vec{\nabla} \). This is a vector-valued differential operator, whose components are given by

\[
\vec{\nabla} = \left( \frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z} \right) .
\]  

(1.50)

In terms of the index notation, we may therefore say that the \( i \)'th component \( (\vec{\nabla})_i \) of the vector \( \vec{\nabla} \) is given by \( \partial / \partial x_i \). In order to make the writing a little less clumsy, it is useful to rewrite this as

\[
\partial_i = \frac{\partial}{\partial x_i} .
\]  

(1.51)
Thus, the \( i \)'th component of \( \vec{\nabla} \) is \( \partial_i \).

It is now evident that the divergence and the curl of a vector \( \vec{A} \) can be written in index notation as

\[
\text{div}\, \vec{A} = \vec{\nabla} \cdot \vec{A} = \partial_i A_i ,
\]

\( \text{curl}\, \vec{A} = (\vec{\nabla} \times \vec{A})_i = \epsilon_{ijk} \partial_j A_k . \)

The Laplacian, \( \nabla^2 = \vec{\nabla} \cdot \vec{\nabla} = \partial^2 / \partial x^2 + \partial^2 / \partial y^2 + \partial^2 / \partial z^2 \), is given by

\[
\nabla^2 = \partial_i \partial_i .
\]

By the rules of partial differentiation, we have \( \partial_i x_j = \delta_{ij} \). If we consider the position vector \( \vec{r} = (x, y, z) \), then we have \( r^2 = x^2 + y^2 + z^2 \), which can be written as

\[
r^2 = x_j x_j .
\]

If we now act with \( \partial_i \) on both sides, we get

\[
2r \partial_i r = 2x_j \partial_i x_j = 2x_j \delta_{ij} = 2x_i .
\]

Thus we have the very useful result that

\[
\partial_i r = x_i .
\]

So far, we have not given any definition of what a 3-vector actually is, and now is the time to remedy this. We may define a 3-vector \( \vec{A} \) as an ordered triplet of real quantities, \( \vec{A} = (A_1, A_2, A_3) \), which transforms under rigid rotations of the Cartesian axes in the same way as does the position vector \( \vec{r} = (x, y, z) \). Now, any rigid rotation of the Cartesian coordinate axes can be expressed as a constant \( 3 \times 3 \) orthogonal matrix \( \mathbf{M} \) acting on the column vector whose components are \( x, y \) and \( z \):

\[
\begin{pmatrix}
x' \\
y' \\
z'
\end{pmatrix} = \mathbf{M} \begin{pmatrix} x \\ y \\ z \end{pmatrix},
\]

where

\[
\mathbf{M}^T \mathbf{M} = 1 .
\]

An example would be the matrix describing a rotation by a (constant) angle \( \theta \) around the \( z \) axis, for which we would have

\[
\mathbf{M} = \begin{pmatrix} \cos \theta & \sin \theta & 0 \\ -\sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{pmatrix} .
\]