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:

• L.D. Landau and E.M. Lifshitz, *The Classical Theory of Fields*

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

1 Electrodynamics and Special Relativity 4  
1.1 Introduction ............................................. 4  
1.2 The Lorentz Transformation ................................. 9  
1.3 An interlude on 3-vectors and suffix notation .......... 12  
1.4 4-vectors and 4-tensors ................................ 17  
1.5 Lorentz tensors ........................................... 25  
1.6 Proper time and 4-velocity ................................ 27  

2 Electrodynamics and Maxwell’s Equations 29  
2.1 Natural units ............................................. 29  
2.2 Gauge potentials and gauge invariance .................... 30  
2.3 Maxwell’s equations in 4-tensor notation ................. 32  
2.4 Lorentz transformation of \( \vec{E} \) and \( \vec{B} \) .................. 40  
2.5 The Lorentz force ......................................... 41  
2.6 Action principle for charged particles .................... 43  
2.7 Gauge invariance of the action ............................ 46  
2.8 Canonical momentum, and Hamiltonian ................... 47  

3 Particle Motion in Static Electromagnetic Fields 48  
3.1 Description in terms of potentials ....................... 49  
3.2 Particle motion in static uniform \( \vec{E} \) and \( \vec{B} \) fields .............. 50  

4 Action Principle for Electrodynamics 55  
4.1 Invariants of the electromagnetic field .................... 55  
4.2 Action for electrodynamics ................................ 58  
4.3 Inclusion of sources ....................................... 61  
4.4 Energy density and energy flux ........................... 64  
4.5 Energy-momentum tensor .................................. 65  
4.6 Energy-momentum tensor for the electromagnetic field .... 71  
4.7 Inclusion of massive charged particles ................... 75  

5 Coulomb’s Law 77  
5.1 Potential of a point charge ............................... 77  
5.2 Electrostatic energy ...................................... 79
5.3 Field of a uniformly moving charge ........................................ 80
5.4 Motion of a charge in a Coulomb potential .......................... 83

6 Electromagnetic Waves .................................................. 88
6.1 Wave equation .............................................................. 88
6.2 Monochromatic plane waves ............................................ 92
6.3 Motion of a point charge in a linearly-polarised E.M. wave .......... 95
6.4 Circular and elliptical polarisation ...................................... 96
6.5 General superposition of plane waves .................................. 98
6.6 Gauge invariance and electromagnetic fields ......................... 105
6.7 Fourier decomposition of electrostatic fields ......................... 109
6.8 Waveguides ................................................................. 110
6.9 Resonant cavities .......................................................... 117

7 Fields Due to Accelerating Charges ................................... 119
7.1 Retarded potentials ....................................................... 119
7.2 Lienard-Wiechert potentials ............................................ 122
7.3 Electric and magnetic fields of a moving charge ..................... 128
7.4 Radiation by accelerated charges ....................................... 131
7.5 Applications of Larmor formula ......................................... 133
7.6 Angular distribution of the radiated power ............................ 136
7.7 Frequency distribution of radiated energy .............................. 141
7.8 Frequency spectrum for relativistic circular motion .................. 145
7.9 Frequency spectrum for periodic motion ............................. 148
7.10 Cerenkov radiation ......................................................... 150
7.11 Thompson scattering ..................................................... 154

8 Radiating Systems ......................................................... 156
8.1 Fields due to localised oscillating sources ............................ 156
8.2 Electric dipole radiation ................................................ 160
8.3 Higher multipoles ......................................................... 164
8.4 Linear antenna ............................................................ 170

9 Electromagnetism and Quantum Mechanics ......................... 173
9.1 The Schrödinger equation and gauge transformations ............. 173
9.2 Magnetic monopoles ...................................................... 176
9.3 Dirac quantisation condition ........................................... 180

10 Local Gauge Invariance and Yang-Mills Theory .................. 181
  10.1 Relativistic quantum mechanics ...................................... 181
  10.2 Yang-Mills theory .................................................... 185
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 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

$$\vec{r}' = \vec{r} - \vec{v}t.$$  

(1.1)

James Clerk Maxwell
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:

$$t' = t.$$  \hfill (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

$$\vec{r}' = M \cdot \vec{r} - \vec{v}t, \quad t' = t,$$  \hfill (1.3)

where $M$ is an orthogonal $3 \times 3$ constant matrix acting by matrix multiplication on the components of the position vector:

$$\vec{r} \leftrightarrow \begin{pmatrix} x \\ y \\ z \end{pmatrix}, \quad M \cdot \vec{r} \leftarrow \begin{pmatrix} x \\ y \\ z \end{pmatrix},$$  \hfill (1.4)

where $M^T M = 1$, and $M^T$ denotes the transpose of the matrix $M$.

Returning to our simplifying assumption that the two frames are parallel, i.e. that $M = I$, 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

$$\vec{u}' = \vec{u} - \vec{v}.$$  \hfill (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

$$c' = c - v.$$  \hfill (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)

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

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

(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.

\(^1\)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.
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

$$\vec{B} = \frac{1}{\omega} \vec{E} \times \vec{E}. \quad (1.13)$$

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 under 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.\[^3\] 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

\[^2\]Strictly, as will be explained later, we should say covariant rather than invariant.

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

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

(1.14)

Thus the Maxwell equations that we now write rather compactly as

\[ \nabla \cdot \vec{D} = 4\pi \rho, \quad \nabla \cdot \vec{B} = 0, \]
\[ \nabla \times \vec{H} - \frac{\partial \vec{D}}{\partial t} = 4\pi \vec{J}, \quad \nabla \times \vec{E} + \frac{\partial \vec{B}}{\partial t} = 0, \]

(1.15)  (1.16)

took, in 1865, the highly inelegant forms

\[ \frac{\partial f}{\partial x} + \frac{\partial g}{\partial y} + \frac{\partial h}{\partial z} = 4\pi e, \]
\[ \frac{\partial F}{\partial x} + \frac{\partial G}{\partial y} + \frac{\partial H}{\partial z} = 0, \]

(1.17)

for the two equations in (1.15), and

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

(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.

\(^4\)Here, we are writing the equations in the so-called “Natural Units,” which we shall be using throughout this course.
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 marketplace, 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

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

in the frame \( S \), and by

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

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

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

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

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

$$x' = A(x - vt) .$$

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

$$x = A(x' + vt') .$$

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 in vacuo to traverse the distance. In fact, the metre is nowadays defined to be the distance travelled by light 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

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

(1.31)

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

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

(1.32)

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

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

(1.33)
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

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

(1.34)

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

\[
\vec{A} \cdot \vec{B} = A_1B_1 + A_2B_2 + A_3B_3 = \sum_{i=1}^{3} A_iB_i.
\]

(1.35)

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 occuring 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

\[
\vec{A} \cdot \vec{B} = A_iB_i.
\]

(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
\begin{equation}
\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). \tag{1.40}
\end{equation}
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
\begin{equation}
\epsilon_{ijk} = -\epsilon_{ikj} = -\epsilon_{jik} = -\epsilon_{kji}. \tag{1.41}
\end{equation}
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
\begin{equation}
\epsilon_{123} = \epsilon_{231} = \epsilon_{312} = +1, \quad \epsilon_{132} = \epsilon_{321} = \epsilon_{213} = -1, \tag{1.42}
\end{equation}
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
\begin{equation}
(\vec{A} \times \vec{B})_i = \epsilon_{ijk} A_j B_k. \tag{1.43}
\end{equation}
For example, the $i = 1$ component (the $x$ component) is given by
\begin{equation}
(\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, \tag{1.44}
\end{equation}
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
\begin{equation}
[\vec{A} \times (\vec{B} \times \vec{C})]_i = \epsilon_{ijk} A_j (\vec{B} \times \vec{C})_k = \epsilon_{ijk} \epsilon_{k\ell m} A_j B_\ell C_m. \tag{1.45}
\end{equation}
For convenience, we may cycle the indices on the second $\epsilon$ tensor around and write this as
\begin{equation}
[\vec{A} \times (\vec{B} \times \vec{C})]_i = \epsilon_{ijk} \epsilon_{\ell m k} A_j B_\ell C_m. \tag{1.46}
\end{equation}
There is an extremely useful identity, which can be proved simply by considering all possible values of the free indices $i, j, \ell, m$:
\begin{equation}
\epsilon_{ijk} \epsilon_{\ell m k} = \delta_{\ell \ell} \delta_{jm} - \delta_{im} \delta_{j\ell}. \tag{1.47}
\end{equation}
Using this in (1.46), we have
\[
[\vec{A} \times (\vec{B} \times \vec{C})]_i = (\delta_{ij}\delta_{km} - \delta_{im}\delta_{jk})A_jB_kC_m,
\]
\[
= \delta_{ij}\delta_{km}A_jB_kC_m - \delta_{im}\delta_{jk}A_jB_kC_m,
\]
\[
= B_iA_jC_j - C_iA_jB_j,
\]
\[
= (\vec{A} \cdot \vec{C})B_i - (\vec{A} \cdot \vec{B})C_i.
\] (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
\[
\text{div}\vec{A} = \vec{\nabla} \cdot \vec{A} = \partial_i A_i, \quad (\text{curl}\vec{A})_i = (\vec{\nabla} \times \vec{A})_i = \epsilon_{ijk}\partial_j A_k.
\] (1.52)

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
\[
\nabla^2 = \partial_i \partial_i.
\] (1.53)

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
\[
r^2 = x_i x_j.
\] (1.54)

If we now act with \(\partial_i\) on both sides, we get
\[
2r \partial_i r = 2x_j \partial_i x_j = 2x_j \delta_{ij} = 2x_i.
\] (1.55)

Thus we have the very useful result that
\[
\partial_i r = \frac{x_i}{r}.
\] (1.56)
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 $M$ acting on the
column vector whose components are $x$, $y$ and $z$:

$$
\begin{pmatrix}
  x' \\
  y' \\
  z'
\end{pmatrix}
= M
\begin{pmatrix}
  x \\
  y \\
  z
\end{pmatrix},
$$

(1.57)

where

$$
M^T M = 1.
$$

(1.58)

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

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

(1.59)

Matrices satisfying the equation (1.58) are called orthogonal matrices. If they are of
dimension $n \times n$, they are called $O(n)$ matrices. Thus the 3-dimensional rotation matrices
are called $O(3)$ matrices.\(^5\)

In index notation, we can write $M$ as $M_{ij}$, where $i$ labels the rows and $j$ labels the
columns:

$$
M = \begin{pmatrix}
  M_{11} & M_{12} & M_{13} \\
  M_{21} & M_{22} & M_{23} \\
  M_{31} & M_{32} & M_{33}
\end{pmatrix}.
$$

(1.60)

The rotation (1.57) can then be expressed as

$$
x'_i = M_{ij} x_j,
$$

(1.61)

\(^5\)There is a little subtlety that we have glossed over, here. If we take the determinant of (1.58), and
use the facts that $\det(AB) = (\det A)(\det B)$ and $\det(A^T) = \det A$, we see that $(\det M)^2 = 1$ and hence
$\det M = \pm 1$. The matrices with $\det M = +1$ are called $SO(n)$ matrices in $n$ dimensions, where the “S”
stands for “special,” meaning unit determinant. It is actually $SO(n)$ matrices that are pure rotations. The
transformations with $\det M = -1$ are actually rotations combined with a reflection of the coordinates (such
as $x \to -x$). Thus, the pure rotation group in 3 dimensions is $SO(3)$.
and the orthogonality condition (1.58) is

\[ M_{ki} M_{kj} = \delta_{ij}. \]  

(1.62)

(Note that if \( M \) has components \( M_{ij} \) then its transpose \( M^T \) has components \( M_{ji} \).) One can directly see in the index notation that the orthogonality condition (1.62) implies, together with (1.61), that the quadratic form \( x_i x_i = x^2 + y^2 + z^2 \) is invariant under these rotations and reflections:

\[ x_i' x_i' = M_{ij} M_{ik} x_j x_k = \delta_{jk} x_j x_k = x_j x_j = x_i x_i. \]  

(1.63)

As stated above, the components of any 3-vector transform the same way under rotations as do the components of the position vector \( \vec{r} \). Thus, if \( \vec{A} \) and \( \vec{B} \) are 3-vectors, then after a rotation by the matrix \( M \) we shall have

\[ A_i' = M_{ij} A_j, \quad B_i' = M_{ij} B_j. \]  

(1.64)

If we calculate the scalar product of \( \vec{A} \) and \( \vec{B} \) after the rotation, we shall therefore have

\[ A_i' B_i' = M_{ij} A_j M_{ik} B_k. \]  

(1.65)

(Note the choice of a different dummy suffix in the expression for \( B_i' \)!) Using the orthogonality condition (1.62), we therefore have that

\[ A_i' B_i' = A_j B_k \delta_{jk} = A_j B_j. \]  

(1.66)

Thus the scalar product of any two 3-vectors is invariant under rotations of the coordinate axes. That is to say, \( A_i B_i \) is a scalar quantity, and by definition a scalar is invariant under rotations.

It is useful to count up how many independent parameters are needed to specify the most general possible rotation matrix \( M \). Looking at (1.60), we can see that a general \( 3 \times 3 \) matrix has 9 components. But our matrix \( M \) is required to be orthogonal, i.e. it must satisfy \( M^T M - I = 0 \). How many equations does this amount to? Naively, it is a \( 3 \times 3 \) matrix equation, and so implies 9 conditions. But this is not correct, since the left-hand side of \( M^T M - I = 0 \) is in fact a symmetric matrix. (Take the transpose, and verify this.) A \( 3 \times 3 \) symmetric matrix has \( (3 \times 4)/2 = 6 \) independent components, and so setting a symmetric \( 3 \times 3 \) matrix to zero implies only 6 independent equations rather than 9. Thus the orthogonality condition imposes 6 constraints on the 9 components of a general \( 3 \times 3 \) matrix, and so that leaves

\[ 9 - 6 = 3 \]  

(1.67)
as the number of independent components of a $3 \times 3$ orthogonal matrix. It is easy to see that this is the correct counting; to specify a general rotation in 3-dimensional space, we need two angles to specify an axis (for example, the latitude and longitude), and a third angle to specify the rotation around that axis.

The above are just a few simple examples of the use of index notation in order to write 3-vector and 3-tensor expressions in Cartesian 3-tensor analysis. It is a very useful notation when one needs to deal with complicated expressions. As we shall now see, there is a very natural generalisation to the case of vector and tensor analysis in 4-dimensional Minkowski spacetime.

### 1.4 4-vectors and 4-tensors

The Lorentz transformations given in (1.32) are linear in the space and time coordinates. They can be written more succinctly if we first define the set of four spacetime coordinates denoted by $x^\mu$, where $\mu$ is an index, or label, that ranges over the values 0, 1, 2 and 3. The case $\mu = 0$ corresponds to the time coordinate $t$, while $\mu = 1, 2$ and 3 corresponds to the space coordinates $x, y$ and $z$ respectively. Thus we have

$$\begin{align*}
(x^0, x^1, x^2, x^3) &= (t, x, y, z).
\end{align*}$$

Of course, once the abstract index label $\mu$ is replaced, as here, by the specific index values 0, 1, 2 and 3, one has to be very careful when reading a formula to distinguish between, for example, $x^2$ meaning the symbol $x$ carrying the spacetime index $\mu = 2$, and $x^2$ meaning the square of $x$. It should generally be obvious from the context which is meant.

The invariant quadratic form appearing on the left-hand side of (1.33) can now be written in a nice way, if we first introduce the 2-index quantity $\eta_{\mu\nu}$, defined by

$$\begin{align*}
\eta_{00} &= -1, & \eta_{11} = \eta_{22} = \eta_{33} &= 1,
\end{align*}$$

with $\eta_{\mu\nu} = 0$ if $\mu \neq \nu$. Note that $\eta_{\mu\nu}$ is symmetric:

$$\eta_{\mu\nu} = \eta_{\nu\mu}.$$  

Using $\eta_{\mu\nu}$, the quadratic form on the left-hand side of (1.33) can be rewritten as

$$x^2 + y^2 + z^2 - t^2 = \sum_{\mu=0}^{3} \sum_{\nu=0}^{3} \eta_{\mu\nu} x^{\mu} x^{\nu}.$$  

---

6The choice to put the index label $\mu$ as a superscript, rather than a subscript, is purely conventional. But, unlike the situation with many arbitrary conventions, in this case the coordinate index is placed upstairs in all modern literature.
In the same way as we previously associated 2-index objects in 3-dimensional Euclidean space with $3 \times 3$ matrices, so here too we can associate $\eta_{\mu\nu}$ with a $4 \times 4$ matrix $\eta$:

$$\eta = \begin{pmatrix}
\eta_{00} & \eta_{01} & \eta_{02} & \eta_{03} \\
\eta_{10} & \eta_{11} & \eta_{12} & \eta_{13} \\
\eta_{20} & \eta_{21} & \eta_{22} & \eta_{23} \\
\eta_{30} & \eta_{31} & \eta_{32} & \eta_{33}
\end{pmatrix} = \begin{pmatrix}
-1 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 \\
0 & 0 & 1 & 0 \\
0 & 0 & 0 & 1
\end{pmatrix}.$$

(1.72)

Thus one can think of the rows of the matrix on the right as being labelled by the index $\mu$ and the columns being labelled by the index $\nu$.

At this point, it is convenient again to introduce the Einstein Summation Convention, now for four-dimensional spacetime indices. This makes the writing of expressions such as (1.71) much less cumbersome. The summation convention works as follows:

In an expression such as (1.71), if an index appears exactly twice in a term, then it will be understood that the index is summed over the natural index range (0, 1, 2, 3 in our present case), and the explicit summation symbol will be omitted. An index that occurs twice in a term, thus is understood to be summed over, is called a Dummy Index.

Since in (1.71) both $\mu$ and $\nu$ occur exactly twice, we can rewrite the expression, using the Einstein summation convention, as simply

$$x^2 + y^2 + z^2 - t^2 = \eta_{\mu\nu} x^\mu x^\nu.$$

(1.73)

On might at first think there would be a great potential for ambiguity, but this is not the case. The point is that in any valid vectorial (or, more generally, tensorial) expression, the only time that a particular index can ever occur exactly twice in a term is when it is summed over. Thus, there is no ambiguity resulting from agreeing to omit the explicit summation symbol, since it is logically inevitable that a summation is intended.\textsuperscript{7} Note that the pair of dummy indices will always occur with one index upstairs and the other downstairs, in any valid expression.

Now let us return to the Lorentz transformations. The pure boosts written in (1.32), being linear in the space and time coordinates, can be written in the form

$$x'^\mu = \Lambda^\mu_{\ \nu} x^\nu,$$

(1.74)

\textsuperscript{7}As a side remark, it should be noted that in a valid vectorial or tensorial expression, a specific index can \textbf{NEVER} appear more than twice in a given term. If you have written down a term where a given index occurs 3, 4 or more times then there is no need to look further at it; it is \textbf{WRONG}. Thus, for example, it is totally meaningless to write $\eta_{\mu\nu} x^\mu x^\mu$. If you ever find such an expression in a calculation then you must stop, and go back to find the place where an error was made.
where $\Lambda^\mu_\nu$ are constants, and the Einstein summation convention is operative for the dummy index $\nu$. By comparing (1.74) carefully with (1.32), we can see that the components $\Lambda^\mu_\nu$ are given by

\begin{align*}
\Lambda^0_0 &= \gamma, & \Lambda^0_i &= -\gamma v_i, \\
\Lambda^i_0 &= -\gamma v_i, & \Lambda^i_j &= \delta_{ij} + \frac{\gamma - 1}{v^2} v_i v_j, \quad (1.75)
\end{align*}

where $\delta_{ij}$ is the Kronecker delta symbol,

\begin{equation}
\delta_{ij} = 1 \text{ if } i = j, \quad \delta_{ij} = 0 \text{ if } i \neq j. \quad (1.76)
\end{equation}

A couple of points need to be explained here. Firstly, we are introducing Latin indices here, namely the $i$ and $j$ indices, which range only over the three spatial index values, $i = 1, 2$ and 3. Thus the 4-index $\mu$ can be viewed as $\mu = (0, i)$, where $i = 1, 2$ and 3. This piece of notation is useful because the three spatial index values always occur on a completely symmetric footing, whereas the time index value $\mu = 0$ is a bit different. This can be seen, for example, in the definition of $\eta_{\mu\nu}$ in (1.72) or (1.69).

The second point is that when we consider spatial indices (for example when $\mu$ takes the values $i = 1, 2$ or 3), it actually makes no difference whether we write the index $i$ upstairs or downstairs. Sometimes, as in (1.75), it will be convenient to be rather relaxed about whether we put spatial indices upstairs or downstairs. By contrast, when the index takes the value 0, it is very important to be careful about whether it is upstairs or downstairs. The reason why we can be cavalier about the Latin indices, but not the Greek, will become clearer as we proceed.

We already saw that the Lorentz boost transformations (1.32), re-expressed in terms of $\Lambda^\mu_\nu$ in (1.75), have the property that $\eta_{\mu\nu} x^\mu x^\nu = \eta_{\mu\nu} x'^\mu x'^\nu$. Thus from (1.74) we have

\begin{equation}
\eta_{\mu\nu} x^\mu x^\nu = \eta_{\mu\nu} \Lambda^\mu_\rho \Lambda^\nu_\sigma x^\rho x^\sigma. \quad (1.77)
\end{equation}

(Note that we have been careful to choose two different dummy indices for the two implicit summations over $\rho$ and $\sigma$!) On the left-hand side, we can replace the dummy indices $\mu$ and $\nu$ by $\rho$ and $\sigma$, and thus write

\begin{equation}
\eta_{\rho\sigma} x^\rho x^\sigma = \eta_{\mu\nu} \Lambda^\mu_\rho \Lambda^\nu_\sigma x^\rho x^\sigma. \quad (1.78)
\end{equation}

This can be grouped together as

\begin{equation}
(\eta_{\rho\sigma} - \eta_{\mu\nu} \Lambda^\mu_\rho \Lambda^\nu_\sigma)x^\rho x^\sigma = 0, \quad (1.79)
\end{equation}
and, since it is true for any $x^\mu$, we must have that

$$
\eta_{\mu\nu} \Lambda^{\mu}_{\;\rho} \Lambda^{\nu}_{\;\sigma} = \eta_{\rho\sigma} \tag{1.80}
$$

(This can also be verified directly from (1.75).) The full set of $\Lambda$‘s that satisfy (1.80) are the Lorentz Transformations. The Lorentz Boosts, given by (1.75), are examples, but they are just a subset of the full set of Lorentz transformations that satisfy (1.80). Essentially, the additional Lorentz transformations consist of rotations of the three-dimensional spatial coordinates. Thus, one can really say that the Lorentz boosts (1.75) are the “interesting” Lorentz transformations, i.e. the ones that rotate space and time into one another. The remainder are just rotations of our familiar old 3-dimensional Euclidean space.

We can count the number of independent parameters in a general Lorentz transformation in the same way we did for the 3-dimensional rotations in the previous section. We start with $\Lambda^{\mu}_{\;\nu}$, which can be thought of as a $4 \times 4$ matrix with rows labelled by $\mu$ and columns labelled by $\nu$. Thus

$$
\Lambda^{\mu}_{\;\nu} \longrightarrow \Lambda = \begin{pmatrix}
\Lambda^0_0 & \Lambda^0_1 & \Lambda^0_2 & \Lambda^0_3 \\
\Lambda^1_0 & \Lambda^1_1 & \Lambda^1_2 & \Lambda^1_3 \\
\Lambda^2_0 & \Lambda^2_1 & \Lambda^2_2 & \Lambda^2_3 \\
\Lambda^3_0 & \Lambda^3_1 & \Lambda^3_2 & \Lambda^3_3
\end{pmatrix} \tag{1.81}
$$

These $4 \times 4 = 16$ components are subject to the conditions (1.80). In matrix notation, (1.80) clearly translates into

$$
\Lambda^T \eta \Lambda - \eta = 0 \tag{1.82}
$$

This is itself a $4 \times 4$ matrix equation, but not all its components are independent since the left-hand side is a symmetric matrix. (Verify this by taking its transpose.) Thus (1.82) contains $(4 \times 5)/2 = 10$ independent conditions, implying that the most general Lorentz transformation has

$$
16 - 10 = 6 \tag{1.83}
$$

independent parameters.

Notice that if $\eta$ had been simply the $4 \times 4$ unit matrix, then (1.82) would have been a direct 4-dimensional analogue of the 3-dimensional orthogonality condition (1.58). In other words, were it not for the minus sign in the $00$ component of $\eta$, the Lorentz transformations would just be spatial rotations in 4 dimensions, and they would be elements of the group $O(4)$. The counting of the number of independent such transformations would be identical to the one given above, and so the group $O(4)$ of orthogonal $4 \times 4$ matrices is characterised by 6 independent parameters.
Because of the minus sign in $\eta$, the group of $4 \times 4$ matrices satisfying (1.82) is called $O(1,3)$, with the numbers 1 and 3 indicating the number of time and space dimensions respectively. Thus the four-dimensional Lorentz Group is $O(1,3)$.

Obviously, the subset of $\Lambda$ matrices of the form

$$\Lambda = \begin{pmatrix} 1 & 0 \\ 0 & \mathbf{M} \end{pmatrix},$$

which is shorthand for $\Lambda = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & M_{11} & M_{12} & M_{13} \\ 0 & M_{21} & M_{22} & M_{23} \\ 0 & M_{31} & M_{32} & M_{33} \end{pmatrix}$, (1.84)

where $\mathbf{M}$ is any $3 \times 3$ orthogonal matrix, satisfies (1.82). This $O(3)$ subgroup of the $O(1,3)$ Lorentz group describes the pure rotations (and reflections) in the 3-dimensional spatial directions. The 3 parameters characterising these transformations, together with the 3 parameters of the velocity vector characterising the pure boost Lorentz transformations (1.75), comprise the total set of $3+3 = 6$ parameters of the general Lorentz transformations.

It is useful to note that any Lorentz transformation $\Lambda_{\mu\nu}$ can be decomposed into a product of a pure Lorentz boost and a pure spatial rotation. Thus we can write a general Lorentz transformation $\Lambda_{\mu\nu}$ in the form

$$\Lambda_{\mu\nu} = \Lambda_{\mu\rho}(R) \Lambda_{\rho\nu}(B),$$

where $\Lambda_{\rho\nu}(B)$ denotes a pure boost, of the form (1.75), and $\Lambda_{\mu\rho}(R)$ denotes a pure spatial rotation, of the form (1.84).

The decomposition given in (1.85) has been organised in the form of a pure Lorentz boost, followed by a pure spatial rotation. One could instead make a decomposition of $\Lambda_{\mu\nu}$ in the opposite order, as a pure spatial rotation followed by a pure Lorentz boost:

$$\Lambda_{\mu\nu} = \tilde{\Lambda}_{\mu\rho}(B) \tilde{\Lambda}_{\rho\nu}(R),$$

where the pure boost and pure rotation transformations will, in general, differ from those in the previous decomposition (1.85), which is why they are written with tildes in (1.86). In other words, the pure boost and the pure spatial rotation matrices do not commute in general.

An example of a decomposition into boost times rotation appears in homework 1, where you are asked to re-express the composition of a pure boost along $x$ followed by a pure boost along $y$ in the form (1.85).

The coordinates $x^\mu = (x^0, x^i)$ live in a four-dimensional spacetime, known as Minkowski Spacetime. This is the four-dimensional analogue of the three-dimensional Euclidean Space.
described by the Cartesian coordinates \( x^i = (x, y, z) \). The quantity \( \eta_{\mu\nu} \) is called the \textit{Minkowski Metric}, and for reasons that we shall see presently, it is called a \textit{tensor}. It is called a metric because it provides the rule for measuring distances in the four-dimensional Minkowski spacetime. The distance, or to be more precise, the \textit{interval}, between two infinitesimally-separated points \((x^0, x^1, x^2, x^3)\) and \((x^0 + dx^0, x^1 + dx^1, x^2 + dx^2, x^3 + dx^3)\) in spacetime is written as \( ds \), and is given by

\[
    ds^2 = \eta_{\mu\nu} dx^\mu dx^\nu. \tag{1.87}
\]

Clearly, this is the Minkowskian generalisation of the three-dimensional distance \( ds_E \) between neighbouring points \((x, y, z)\) and \((x + dx, y + dy, z + dz)\) in Euclidean space, which, by Pythagoras’ theorem, is given by

\[
    ds_E^2 = dx^2 + dy^2 + dz^2 = \delta_{ij} dx^i dx^j. \tag{1.88}
\]

The Euclidean metric (1.88) is invariant under arbitrary constant rotations of the \((x, y, z)\) coordinate system. (This is clearly true because the distance between the neighbouring points must obviously be independent of how the axes of the Cartesian coordinate system are oriented.) By the same token, the Minkowski metric (1.87) is invariant under arbitrary Lorentz transformations. In other words, as can be seen to follow immediately from (1.80), the spacetime interval \( ds'^2 = \eta_{\mu\nu} dx'^\mu dx'^\nu \) calculated in the primed frame is identical to the interval \( ds^2 \) calculated in the unprimed frame

\[
    ds'^2 = \eta_{\mu\nu} dx'^\mu dx'^\nu = \eta_{\mu\nu} \Lambda^\mu_{\rho} \Lambda^\nu_{\sigma} dx^\rho dx^\sigma, \\
    = \eta_{\rho\sigma} dx^\rho dx^\sigma = ds^2. \tag{1.89}
\]

For this reason, we do not need to distinguish between \( ds^2 \) and \( ds'^2 \), since it is the same in all inertial frames. It is what is called a \textit{Lorentz Scalar}.

The Lorentz transformation rule of the coordinate differential \( dx^\mu \), i.e.

\[
    dx'^\mu = \Lambda^\mu_{\nu} dx^\nu, \tag{1.90}
\]

can be taken as the prototype for more general 4-vectors. Thus, we may define any set of four quantities \( U^\mu \), for \( \mu = 0, 1, 2 \) and \( 3 \), to be the components of a Lorentz 4-vector (often, we shall just abbreviate this to simply a 4-vector) if they transform, under Lorentz transformations, according to the rule

\[
    U'^\mu = \Lambda^\mu_{\nu} U^\nu. \tag{1.91}
\]
The Minkowski metric $\eta_{\mu\nu}$ may be thought of as a $4 \times 4$ matrix, whose rows are labelled by $\mu$ and columns labelled by $\nu$, as in (1.72). Clearly, the inverse of this matrix takes the same form as the matrix itself. We denote the components of the inverse matrix by $\eta^{\mu\nu}$. This is called, not surprisingly, the inverse Minkowski metric. Clearly it satisfies the relation

$$\eta_{\mu\nu} \eta^{\nu\rho} = \delta^\rho_\mu,$$

where the 4-dimensional Kronecker delta is defined to equal 1 if $\mu = \rho$, and to equal 0 if $\mu \neq \rho$. Note that like $\eta_{\mu\nu}$, the inverse $\eta^{\mu\nu}$ is symmetric also: $\eta^{\mu\nu} = \eta^{\nu\mu}$.

The Minkowski metric and its inverse may be used to lower or raise the indices on other quantities. Thus, for example, if $U^\mu$ are the components of a 4-vector, then we may define

$$U_\mu = \eta_{\mu\nu} U^\nu.$$

(1.93)

This is another type of 4-vector. Two distinguish the two, we call a 4-vector with an upstairs index a contravariant 4-vector, while one with a downstairs index is called a covariant 4-vector. Note that if we raise the lowered index in (1.93) again using $\eta^{\mu\nu}$, then we get back to the starting point:

$$\eta^{\mu\nu} U_\nu = \eta^{\mu\nu} \eta_{\rho\nu} U^\rho = \delta^\rho_\mu U^\rho = U^\mu.$$

(1.94)

It is for this reason that we can use the same symbol $U$ for the covariant 4-vector $U_\mu = \eta_{\mu\nu} U^\nu$ as we used for the contravariant 4-vector $U^\mu$.

In a similar fashion, we may define the quantities $\Lambda^\mu_{\nu}$ by

$$\Lambda^\mu_{\nu} = \eta_{\mu\rho} \eta^{\nu\sigma} \Lambda^\rho_{\sigma}.$$

(1.95)

It is then clear that (1.80) can be restated as

$$\Lambda^\mu_{\nu} \Lambda^\rho_{\nu} = \delta^\rho_\mu.$$

(1.96)

Notice two points concerning raising and lowering indices with $\eta$. The first is that if we have a vector-valued or tensor-valued equation, such as $A^\mu = B^\mu$, or $S_{\mu\nu} = T_{\mu\nu}$ or whatever, we can raise or lower these free indices at will, as long as we raise them on both sides of the equation at the same time. Thus, for example,

$$S_{\mu\nu} = T_{\mu\nu} \iff S^\mu_{\nu} = T^\mu_{\nu} \iff S^{\mu\nu} = T^{\mu\nu}.$$

(1.97)

The second point is that if there is an index contraction in a term, we can freely “see-saw” the pair of dummy indices, moving the upper index down and simultaneously the lower index up. Thus, for example,

$$A^\mu B_\mu = A_\mu B^\mu,$$

(1.98)
and so on. The reason for emphasising these two points is just to make clear that one does not need to make a song and dance about raising or lowering free indices, or see-sawing dummy index positions. After any such operations, a valid covariant equation will always have the properties that every matching free index will be in the same location (upstairs or downstairs) in every term in the equation. Furthermore, every dummy index pair will always have one occurrence of that index upstairs, and the other downstairs.\(^8\)

We can invert the Lorentz transformation \(x'_{\mu} = \Lambda_{\mu\nu} x^\nu\), by multiplying both sides by \(\Lambda^{\mu\rho}\) and using (1.96) to give \(x'^{\mu} \Lambda_{\mu\rho} = \delta^\rho_\nu x^\nu = x^\rho\), and hence, after relabelling,

\[
x^{\mu} = \Lambda^{\mu}_{\nu} x'^{\nu}.
\]  

(1.99)

It now follows from (1.91) that the components of the covariant 4-vector \(U_{\mu}\) defined by (1.93) transform under Lorentz transformations according to the rule

\[
U'_{\mu} = \Lambda_{\mu\nu} U_{\nu}.
\]  

(1.100)

Any set of 4 quantities \(U_{\mu}\) which transform in this way under Lorentz transformations will be called a covariant 4-vector.

Using (1.99), we can see that the gradient operator \(\partial/\partial x^{\mu}\) transforms as a covariant 4-vector. Using the chain rule for partial differentiation we have

\[
\frac{\partial}{\partial x'^{\mu}} = \frac{\partial x'^{\nu}}{\partial x^{\mu}} \frac{\partial}{\partial x^{\nu}}.
\]  

(1.101)

But from (1.99) we have (after a relabelling of indices) that

\[
\frac{\partial x'^{\nu}}{\partial x^{\mu}} = \Lambda^{\nu}_{\mu},
\]  

(1.102)

and hence (1.101) gives

\[
\frac{\partial}{\partial x'^{\mu}} = \Lambda^{\nu}_{\mu} \frac{\partial}{\partial x^{\nu}}.
\]  

(1.103)

As can be seen from (1.100), this is precisely the transformation rule for a a covariant 4-vector. The gradient operator arises sufficiently often that it is useful to use a special

\(\ast\)These statements apply to equations written in the four-dimensionally covariant language, with Greek indices \(\mu, \nu, \ldots\) ranging over 0, 1, 2 and 3. As has already been emphasised, if one decomposes a four-dimensionally covariant expression into the 1+3 language of time plus three spatial directions (denoted by the Latin spatial indices \(i, j, \ldots\)), then one is completely free to write the Latin indices upstairs or downstairs, unmatched between different terms. In the context of our Minkowski spacetime discussions in this course, the only reason for caring about the distinction between upstairs and downstairs indices is because of the time direction. Thus we must respect the upstairs/downstairs rules for four-dimensional covariance, but it is unimportant in the 1+3 language we use for purely three-dimensional rotational covariance.
symbol to denote it. We therefore define

\[ \partial_\mu \equiv \frac{\partial}{\partial x^\mu}. \]  

(1.104)

Thus the Lorentz transformation rule (1.103) is now written as

\[ \partial'_\mu = \Lambda_\mu^\nu \partial_\nu. \]  

(1.105)

1.5 Lorentz tensors

Having seen how contravariant and covariant 4-vectors transform under Lorentz transformations (as given in (1.91) and (1.100) respectively), we can now define the transformation rules for more general objects called tensors. These objects carry multiple indices, and each one transforms with a \( \Lambda \) factor, of either the (1.91) type if the index is upstairs, or of the (1.100) type if the index is downstairs. Thus, for example, a tensor \( T_{\mu\nu} \) transforms under Lorentz transformations according to the rule

\[ T'_{\mu\nu} = \Lambda_\mu^\rho \Lambda_\nu^\sigma T_{\rho\sigma}. \]  

(1.106)

More generally, a tensor \( T^{\mu_1 \cdots \mu_m}{}_{\nu_1 \cdots \nu_n} \) will transform according to the rule

\[ T'^{\mu_1 \cdots \mu_m}{}_{\nu_1 \cdots \nu_n} = \Lambda^{\mu_1 \rho_1} \cdots \Lambda^{\mu_m \rho_m} \Lambda_{\nu_1 \sigma_1} \cdots \Lambda_{\nu_n \sigma_n} T^{\mu_1 \cdots \rho_m}{}_{\sigma_1 \cdots \sigma_n}. \]  

(1.107)

Note that scalars are just special cases of tensors with no indices, while vectors are special cases with just one index.

It is easy to see that products of tensors give rise again to tensors. For example, if \( U^\mu \) and \( V^\nu \) are two contravariant vectors then \( T^{\mu\nu} \equiv U^\mu V^\nu \) is a tensor, since, using the known transformation rules for \( U^\mu \) and \( V^\nu \) we have

\[ T'^{\mu\nu} = U'^\mu V'^\nu = \Lambda^\mu_\rho U^\rho \Lambda_\nu^\sigma V^\sigma, \]

\[ = \Lambda^{\mu_\rho} \Lambda_\nu^\sigma T^{\rho\sigma}. \]  

(1.108)

Note that the gradient operator \( \partial_\mu \) can also be used to map a tensor into another tensor. For example, if \( U_\mu \) is a vector field (i.e. a vector that changes from place to place in space-time) then \( S_{\mu\nu} \equiv \partial_\mu U_\nu \) is a tensor field. As always, the way to check that that something is a tensor is to check that it transforms in the proper way under Lorentz transformations. So in this case, one needs to check that it transforms in the way an \( (m, n) = (0, 2) \) tensor in eqn (1.107) does.

We make also define the operation of Contraction, which reduces a tensor to one with a smaller number of indices. A contraction is performed by setting an upstairs index on a
tensor equal to a downstairs index. The Einstein summation convention then automatically comes into play, and the result is that one has an object with one fewer upstairs indices and one fewer downstairs indices. Furthermore, a simple calculation shows that the new object is itself a tensor. Consider, for example, a tensor $T^\mu_\nu$. This, of course, transforms as
\[
T'^\mu_\mu = \Lambda^\mu_\rho \Lambda^\nu_\sigma T^\rho_\sigma
\] (1.109)
under Lorentz transformations. If we form the contraction and define $\phi \equiv T^\mu_\mu$, then we see that under Lorentz transformations we shall have
\[
\phi' \equiv T'^\mu_\mu = \Lambda^\mu_\rho \Lambda^\nu_\sigma T^\rho_\sigma = \delta^\sigma_\rho T^\rho_\sigma = \phi.
\] (1.110)
Since $\phi' = \phi$, it follows, by definition, that $\phi$ is a scalar.

An essentially identical calculation shows that for a tensor with arbitrary numbers of upstairs and downstairs indices, if one makes an index contraction of one upstairs with one downstairs index, the result is a tensor with the corresponding reduced numbers of indices. Of course multiple contractions work in the same way.

The Minkowski metric $\eta_{\mu\nu}$ is itself a tensor, but of a rather special type, known as an invariant tensor. This is because, unlike a generic 2-index tensor, the Minkowski metric is identical in all Lorentz frames. To see this, let us first write out how it would transform under Lorentz transformations, using the usual transformation rules in (1.107):
\[
\eta'^\mu_\nu = \Lambda^\mu_\rho \Lambda^\nu_\sigma \eta^{\rho\sigma}.
\] (1.111)
Our goal is to show that in fact $\eta'^\mu_\nu = \eta_{\mu\nu}$, i.e. that it is actually invariant under Lorentz transformations. Now although the right-hand side of (1.111) looks reminiscent of what one has in (1.80), which is the defining property of the $\Lambda^\mu_\nu$ Lorentz transformations, it is not the same. Specifically, in (1.111) the indices of the two $\Lambda$ transformations are contracted with $\eta$ on their second indices, rather than on the first indices as in (1.80). We can easily work out what the right-hand side of (1.111) is by going through the following steps. First, we rewrite (1.80) in matrix language as $\Lambda^T \eta \Lambda = \eta$. Then right-multiply by $\Lambda^{-1}$ and left-multiply by $\eta^{-1}$; this gives $\eta^{-1} \Lambda^T \eta = \Lambda^{-1}$. Next left-multiply by $\Lambda$ and right-multiply by $\eta^{-1}$, which gives $\Lambda \eta^{-1} \Lambda^T = \eta^{-1}$. (This is the analogue for the Lorentz transformations of the proof, for ordinary orthogonal matrices, that $M^T M = 1$ implies $M M^T = 1$.) Converting back to index notation gives $\Lambda^\mu_\rho \Lambda^\nu_\sigma \eta^{\rho\sigma} = \eta^{\mu\nu}$. After some index raising and lowering, this gives
\[
\Lambda^\mu_\rho \Lambda^\nu_\sigma \eta_{\rho\sigma} = \eta_{\mu\nu}.
\] (1.112)
Applying this result to the right-hand side of (1.111) therefore gives the desired result,

$$\eta'_{\mu\nu} = \eta_{\mu\nu}. \quad (1.113)$$

Thus we have shown that the tensor \(\eta_{\mu\nu}\) is actually \textit{invariant} under Lorentz transformations. The same is also true for the inverse metric \(\eta^{\mu\nu}\).

We already saw that the gradient operator \(\partial_{\mu} \equiv \partial/\partial x^{\mu}\) transforms as a covariant vector. If we define, in the standard way, \(\partial^{\mu} \equiv \eta^{\mu\nu} \partial_{\nu}\), then it is evident from what we have seen above that the operator

$$\Box \equiv \partial^{\mu} \partial_{\mu} = \eta^{\mu\nu} \partial_{\mu} \partial_{\nu} \quad (1.114)$$

transforms as a scalar under Lorentz transformations. This is a very important operator, which is otherwise known as the wave operator, or d’Alembertian:

$$\Box = -\partial_{0} \partial_{0} + \partial_{i} \partial_{i} = -\frac{\partial^2}{\partial t^2} + \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}. \quad (1.115)$$

It is worth commenting further at this stage about a remark that was made earlier. Notice that in (1.115) we have been cavalier about the location of the Latin indices, which of course range only over the three spatial directions \(i = 1, 2, 3\). We can get away with this because the metric that is used to raise or lower the Latin indices is just the Minkowski metric restricted to the index values 1, 2 and 3. But since we have

$$\eta_{00} = -1, \quad \eta_{ij} = \delta_{ij}, \quad \eta_{0i} = \eta_{i0} = 0, \quad (1.116)$$

this means that Latin indices are lowered and raised using the Kronecker delta \(\delta_{ij}\) and its inverse \(\delta^{ij}\). But these are just the components of the unit matrix, and so raising or lowering Latin indices has no effect. It is because of the minus sign associated with the \(\eta_{00}\) component of the Minkowski metric that we have to pay careful attention to the process of raising and lowering Greek indices. Thus, we can get away with writing \(\partial_{i} \partial_{i}\), but we cannot write \(\partial_{\mu} \partial_{\mu}\).

### 1.6 Proper time and 4-velocity

We defined the Lorentz-invariant interval \(ds\) between infinitesimally-separated spacetime events by

$$ds^2 = \eta_{\mu\nu} dx^{\mu} dx^{\nu} = -dt^2 + dx^2 + dy^2 + dz^2. \quad (1.117)$$

This is the Minkowskian generalisation of the spatial interval in Euclidean space. Note that \(ds^2\) can be positive, negative or zero. These cases correspond to what are called spacelike, timelike or null separations, respectively.
On occasion, it is useful to define the negative of $ds^2$, and write

$$d\tau^2 = -ds^2 = -\eta_{\mu\nu} dx^\mu dx^\nu = dt^2 - dx^2 - dy^2 - dz^2.$$  \hfill (1.118)

This is called the *Proper Time* interval, and $\tau$ is the proper time. Since $ds$ is a Lorentz scalar, it is obvious that $d\tau$ is a scalar too.

We know that $dx^\mu$ transforms as a contravariant 4-vector. Since $d\tau$ is a scalar, it follows that

$$U^\mu \equiv \frac{dx^\mu}{d\tau}$$  \hfill (1.119)

is a contravariant 4-vector also. If we think of a particle following a path, or *worldline* in spacetime parameterised by the proper time $\tau$, i.e. it follows the path $x^\mu = x^\mu(\tau)$, then $U^\mu$ defined in (1.119) is called the *4-velocity* of the particle.

It is useful to see how the 4-velocity is related to the usual notion of 3-velocity of a particle. By definition, the 3-velocity $\vec{u}$ is a 3-vector with components $u^i$ given by

$$u^i = \frac{dx^i}{dt}.$$  \hfill (1.120)

From (1.118), it follows that

$$d\tau^2 = dt^2[1 - (\frac{dx}{dt})^2 - (\frac{dy}{dt})^2 - (\frac{dz}{dt})^2] = dt^2(1 - u^2),$$  \hfill (1.121)

where $u = |\vec{u}|$, or in other words, $u = \sqrt{\vec{u} \cdot \vec{u}}$. In view of the definition of the $\gamma$ factor in (1.31), it is natural to define

$$\gamma \equiv \frac{1}{\sqrt{1 - u^2}}.$$  \hfill (1.122)

Thus we have $d\tau = dt/\gamma$, and so from (1.119) the 4-velocity can be written as

$$U^\mu = \frac{dt}{d\tau} \frac{dx^\mu}{dt} = \gamma \frac{dx^\mu}{dt}.$$  \hfill (1.123)

Since $dx^0/dt = 1$ and $dx^i/dt = u^i$, we therefore have that

$$U^0 = \gamma, \quad U^i = \gamma u^i.$$  \hfill (1.124)

Note that $U^\mu U_\mu = -1$, since, from (1.118), we have

$$U^\mu U_\mu = \eta_{\mu\nu} U^\mu U^\nu = \frac{\eta_{\mu\nu} dx^\mu dx^\nu}{(d\tau)^2} = \frac{-(d\tau)^2}{(d\tau)^2} = -1.$$  \hfill (1.125)

We shall sometimes find it convenient to rewrite (1.124) as

$$U^\mu = (\gamma, \gamma u^i) \quad \text{or} \quad U^\mu = (\gamma, \gamma \vec{u}).$$  \hfill (1.126)
Having set up the 4-vector formalism, it is now completely straightforward write down how velocities transform under Lorentz transformations. We know that the 4-velocity $U^\mu$ will transform according to (1.91), and this is identical to the way that the coordinates $x^\mu$ transform:

$$U'^\mu = \Lambda^{\mu}_{\nu} U^\nu, \quad x'^\mu = \Lambda^{\mu}_{\nu} x^\nu.$$  \hspace{1cm} (1.127)

Therefore, if we want to know how the 3-velocity transforms, we need only write down the Lorentz transformations for $(t, x, y, z)$, and then replace $(t, x, y, z)$ by $(U^0, U^1, U^2, U^3)$. Finally, using (1.126) to express $(U^0, U^1, U^2, U^3)$ in terms of $\vec{u}$ will give the result.

Consider, for simplicity, the case where $S'$ is moving along the $x$ axis with velocity $v$. The Lorentz transformation for $U^\mu$ can therefore be read off from (1.30) and (1.31):

\begin{align*}
U'^0 & = \gamma_v (U^0 - vU^1), \\
U'^1 & = \gamma_v (U^1 - vU^0), \\
U'^2 & = U^2, \\
U'^3 & = U^3,
\end{align*}

where we are now using $\gamma_v \equiv (1 - v^2)^{-1/2}$ to denote the gamma factor of the Lorentz transformation, to distinguish it from the $\gamma$ constructed from the 3-velocity $\vec{u}$ of the particle in the frame $S$, which is defined in (1.122). Thus from (1.126) we have

\begin{align*}
\gamma' & = \gamma \gamma_v (1 - v u_x), \\
\gamma' u'_x & = \gamma \gamma_v (u_x - v), \\
\gamma' u'_y & = \gamma u_y, \\
\gamma' u'_z & = \gamma u_z,
\end{align*}

(1.129)

where, of course, $\gamma' = (1 - u'^2)^{-1/2}$ is the analogue of $\gamma$ in the frame $S'$. Thus we find

\begin{align*}
u'_x & = \frac{u_x - v}{1 - vu_x}, \quad u'_y = \frac{u_y}{\gamma_v (1 - vu_x)}, \quad u'_z = \frac{u_z}{\gamma_v (1 - vu_x)}. \hspace{1cm} (1.130)
\end{align*}

2 Electrodynamics and Maxwell’s Equations

2.1 Natural units

We saw earlier that the supposition of the universal validity of Maxwell’s equations in all inertial frames, which in particular would imply that the speed of light should be the same in all frames, is consistent with experiment. It is therefore reasonable to expect that Maxwell’s
equations should be compatible with special relativity. However, written in their standard form (1.7), this compatibility is by no means apparent. Our next task will be to re-express the Maxwell equations, in terms of 4-tensors, in a way that makes their Lorentz covariance manifest.

We shall begin by changing units from the S.I. system in which the Maxwell equations are given in (1.7). The first step is to change to Gaussian units, by performing the rescalings

\[ \vec{E} \rightarrow \frac{1}{\sqrt{4\pi \varepsilon_0}} \vec{E}, \quad \vec{B} \rightarrow \sqrt{\frac{\mu_0}{4\pi}} \vec{B}, \]

\[ \rho \rightarrow \sqrt{4\pi \varepsilon_0} \rho, \quad \vec{J} \rightarrow \sqrt{4\pi \varepsilon_0} \vec{J}. \]  \hspace{1cm} (2.1)

Bearing in mind that the speed of light is given by

\[ c = \frac{1}{\sqrt{\mu_0 \varepsilon_0}}, \]

we see that the Maxwell equations (1.7) become

\[ \vec{\nabla} \cdot \vec{E} = 4\pi \rho, \quad \vec{\nabla} \times \vec{B} - \frac{1}{c} \partial \vec{E} \partial t = \frac{4\pi}{c} \vec{J}, \]

\[ \vec{\nabla} \cdot \vec{B} = 0, \quad \vec{\nabla} \times \vec{E} + \frac{1}{c} \partial \vec{B} \partial t = 0, \]  \hspace{1cm} (2.2)

Finally, we pass from Gaussian units to Natural units, by choosing our units of length and time so that \( c = 1, \) as we did in our discussion of special relativity. Thus, in natural units, the Maxwell equations become

\[ \vec{\nabla} \cdot \vec{E} = 4\pi \rho, \quad \vec{\nabla} \times \vec{B} - \partial \vec{E} \partial t = 4\pi \vec{J}, \]  \hspace{1cm} (2.3)

\[ \vec{\nabla} \cdot \vec{B} = 0, \quad \vec{\nabla} \times \vec{E} + \frac{\partial \vec{B}}{\partial t} = 0, \]  \hspace{1cm} (2.4)

The equations (2.3), which have sources on the right-hand side, are called the Field Equations. The equations (2.4) are called Bianchi Identities. We shall elaborate on this a little later.

### 2.2 Gauge potentials and gauge invariance

We already remarked that the two Maxwell equations (2.4) are known as Bianchi identities. They are not field equations, since there are no sources; rather, they impose constraints on the electric and magnetic fields. The first equation in (2.4), i.e. \( \vec{\nabla} \cdot \vec{B} = 0, \) can be solved by writing

\[ \vec{B} = \vec{\nabla} \times \vec{A}, \]  \hspace{1cm} (2.5)

where \( \vec{A} \) is the magnetic 3-vector potential. Note that (2.5) \textit{identically} solves \( \vec{\nabla} \cdot \vec{B} = 0, \) because of the vector identity that \textit{div curl} \( \equiv 0. \) Substituting (2.5) into the second equation
in (2.4), we obtain
\[ \nabla \times (\vec{E} + \frac{\partial \vec{A}}{\partial t}) = 0. \] (2.6)

This can be solved, again identically, by writing
\[ \vec{E} + \frac{\partial \vec{A}}{\partial t} = -\nabla \phi, \] (2.7)

where \( \phi \) is the electric scalar potential. Thus we can solve the Bianchi identities (2.4) by writing \( \vec{E} \) and \( \vec{B} \) in terms of scalar and 3-vector potentials \( \phi \) and \( \vec{A} \):
\[ \vec{E} = -\nabla \phi - \frac{\partial \vec{A}}{\partial t}, \quad \vec{B} = \nabla \times \vec{A}. \] (2.8)

Although we have now “disposed of” the two Maxwell equations in (2.4), it has been achieved at a price, in that there is a redundancy in the choice of gauge potentials \( \phi \) and \( \vec{A} \). First, we may note that that \( \vec{B} \) in (2.8) is unchanged if we make the replacement
\[ \vec{A} \rightarrow \vec{A} + \nabla \lambda, \] (2.9)

where \( \lambda \) is an arbitrary function of position and time. The expression for \( \vec{E} \) will also be invariant, if we simultaneously make the replacement
\[ \phi \rightarrow \phi - \frac{\partial \lambda}{\partial t}. \] (2.10)

To summarise, if a given set of electric and magnetic fields \( \vec{E} \) and \( \vec{B} \) are described by a scalar potential \( \phi \) and 3-vector potential \( \vec{A} \) according to (2.8), then the identical physical situation (i.e. identical electric and magnetic fields) is equally well described by a new pair of scalar and 3-vector potentials, related to the original pair by the Gauge Transformations given in (2.9) and (2.10), where \( \lambda \) is an arbitrary function of position and time.

We can in fact use the gauge invariance to our advantage, by making a convenient and simplifying gauge choice for the scalar and 3-vector potentials. We have one arbitrary function (i.e. \( \lambda(t, \vec{r}) \)) at our disposal, and so this allows us to impose one functional relation on the potentials \( \phi \) and \( \vec{A} \). For our present purposes, the most useful gauge choice is to use this freedom to impose the Lorenz gauge condition,
\[ \nabla \cdot \vec{A} + \frac{\partial \phi}{\partial t} = 0. \] (2.11)

Note that, contrary to the belief of many physicists, this gauge choice was introduced by the Danish physicist Ludvig Valentin Lorenz, and not the Dutch physicist Hendrik Antoon Lorentz who is responsible for the Lorentz transformation. Adding to the confusion is that
Unlike many other gauge choices that one encounters, the Lorenz gauge condition is, as we shall see later, Lorentz invariant.

Substituting (2.8) into the remaining Maxwell equations (i.e. (2.3), and using the Lorenz gauge condition (2.11), we therefore find

\[
\nabla^2 \phi - \frac{\partial^2 \phi}{\partial t^2} = -4\pi \rho, \\
\nabla^2 \vec{A} - \frac{\partial^2 \vec{A}}{\partial t^2} = -4\pi \vec{J}.
\]

(2.12)

The important thing, which we shall make use of shortly, is that in each case we have on the left-hand side the d’Alembertian operator \( \Box = \partial^\mu \partial_\mu \), which we discussed earlier.

2.3 Maxwell’s equations in 4-tensor notation

The next step is to write the Maxwell equations in terms of four-dimensional quantities. Since the 3-vectors \( \vec{E} \) and \( \vec{B} \) describing the electric and magnetic fields have three components each, there is clearly no way in which they can be “assembled” into 4-vectors. However, we may note that in four dimensions a two-index antisymmetric tensor has \((4 \times 3)/2 = 6\) independent components. Since this is equal to \(3 + 3\), it suggests that perhaps we should be grouping the electric and magnetic fields together into a single 2-index antisymmetric tensor.
tensor. This is in fact exactly what is needed. Thus we introduce a tensor $F_{\mu\nu}$, satisfying

$$F_{\mu\nu} = -F_{\nu\mu}.$$  \hfill (2.13)

It turns out that we should define its components in terms of $\vec{E}$ and $\vec{B}$ as follows:

$$F_{0i} = -E_i, \quad \text{(which implies} \quad F_{i0} = E_i), \quad F_{ij} = \epsilon_{ijk} B_k.$$  \hfill (2.14)

Here $\epsilon_{ijk}$ is the usual totally-antisymmetric tensor of 3-dimensional vector calculus. It is equal to +1 if $(ijk)$ is an even permutation of (123), to = −1 if it is an odd permutation, and to zero if it is no permutation (i.e. if two or more of the indices $(ijk)$ are equal). In other words, we have

$$F_{23} = B_1, \quad F_{31} = B_2, \quad F_{12} = B_3,$$

$$F_{32} = -B_1, \quad F_{13} = -B_2, \quad F_{21} = -B_3.$$  \hfill (2.15)

Viewing $F_{\mu\nu}$ as a matrix $F$ with rows labelled by $\mu$ and columns labelled by $\nu$, we shall have

$$F = \begin{pmatrix} F_{00} & F_{01} & F_{02} & F_{03} \\ F_{10} & F_{11} & F_{12} & F_{13} \\ F_{20} & F_{21} & F_{22} & F_{23} \\ F_{30} & F_{31} & F_{32} & F_{33} \end{pmatrix} = \begin{pmatrix} 0 & -E_1 & -E_2 & -E_3 \\ E_1 & 0 & B_3 & -B_2 \\ E_2 & -B_3 & 0 & B_1 \\ E_3 & B_2 & -B_1 & 0 \end{pmatrix}.$$  \hfill (2.16)

We also need to combine the charge density $\rho$ and the 3-vector current density $\vec{J}$ into a four-dimensional quantity. This is easy; we just define a 4-vector $J^\mu$, whose spatial components $J^i$ are just the usual 3-vector current components, and whose time component $J^0$ is equal to the charge density $\rho$:

$$J^0 = \rho, \quad J^i = J^i.$$  \hfill (2.17)

A word of caution is in order here. Although we have defined objects $F_{\mu\nu}$ and $J^\mu$ that have the appearance of a 4-tensor and a 4-vector, we are only entitled to call them such if we have verified that they transform in the proper way under Lorentz transformations. In fact they do, and we shall justify this a little later.

For now, we shall proceed to see how the Maxwell equations look when expressed in terms of $F_{\mu\nu}$ and $J^\mu$. The answer is that they become

$$\partial_\mu F^{\mu\nu} = -4\pi J^\nu,$$  \hfill (2.18)

$$\partial_\mu F_{\nu\rho} + \partial_\nu F_{\rho\mu} + \partial_\rho F_{\mu\nu} = 0.$$  \hfill (2.19)
Two very nice things have happened. First of all, the original four Maxwell equations 
(2.3) and (2.4) have become just two four-dimensional equations; (2.18) is the field equation, 
and (2.19) is the Bianchi identity. Secondly, the equations are manifestly Lorentz 
covariant; i.e. they transform tensorially under Lorentz transformations. This means that 
they keep exactly the same form in all Lorentz frames. If we start with (2.18) and (2.19) 
in the unprimed frame $S$, then we know that in the frame $S'$, related to $S$ by the Lorentz 
transformation (1.74), the equations will look identical, except that they will now have primes on all the quantities. Furthermore, we know precisely how the primed quantities are 
related to the unprimed:

$$F'_{\mu\nu} = \Lambda_{\mu}^\rho \Lambda_{\nu}^\sigma F_{\rho\sigma}, \quad J'^\mu = \Lambda^\mu_\nu J^\nu,$$

(2.20)

eqst., where $\Lambda_{\mu}^\nu$ describes the Lorentz transformation from the frame $S$ to the frame $S'$.

We should first verify that indeed (2.18) and (2.19) are equivalent to the Maxwell equations 
(2.3) and (2.4). Consider first (2.18). This equation is vector-valued, since it has the 
free index $\nu$. Therefore, to reduce it down to three-dimensional equations, we have two 
cases to consider, namely $\nu = 0$ or $\nu = j$. For $\nu = 0$ we have

$$\partial_\mu F^\mu_0 = \partial_0 F^{00} + \partial_i F^{i0} = \partial_0 F^{00} = -4\pi J^0,$$

(2.21)

which therefore corresponds (see (2.14) and (2.17)) to

$$-\partial_i E_i = -4\pi \rho, \quad \text{i.e.} \quad \vec{\nabla} \cdot \vec{E} = 4\pi \rho.$$

(2.22)

For $\nu = j$, we shall have

$$\partial_\mu F^{\mu j} = \partial_0 F^{0j} + \partial_i F^{ij} = -4\pi J^j,$$

(2.23)

which gives

$$\partial_0 E_j + \epsilon_{ijk} \partial_i B_k = -4\pi J^j.$$

(2.24)

This is just

$$-\frac{\partial \vec{E}}{\partial t} + \vec{\nabla} \times \vec{B} = 4\pi \vec{J}.$$

(2.25)

Thus (2.18) is equivalent to the two Maxwell field equations in (2.3).

Turning now to (2.19), it follows from the antisymmetry (2.13) of $F_{\mu\nu}$ that the left-hand 
side is totally antisymmetric in $(\mu\nu\rho)$ (i.e. it changes sign under any exchange of a pair of 
indices). Therefore there are two distinct inequivalent assignments of indices, after we make 

\footnote{Recall that the $i$’th component of $\vec{\nabla} \times \vec{V}$ is given by $(\vec{\nabla} \times \vec{V})_i = \epsilon_{ijk} \partial_j V_k$ for any 3-vector $\vec{V}$.}
the 1 + 3 decomposition $\mu = (0, i)$ etc.: Either one of the indices is a 0 with the other two Latin, or else all three are Latin. Consider first $(\mu, \nu, \rho) = (0, i, j)$:

$$\partial_0 F_{ij} + \partial_i F_{j0} + \partial_j F_{0i} = 0 ,$$

(2.26)

which, from (2.14), means

$$\epsilon_{ijk} \frac{\partial B_k}{\partial t} + \partial_i E_j - \partial_j E_i = 0 .$$

(2.27)

Since this is antisymmetric in $ij$ there is no loss of generality involved in contracting with $\epsilon_{ij\ell}$, which gives

$$2 \frac{\partial B_\ell}{\partial t} + 2 \epsilon_{ij\ell} \partial_i E_j = 0 .$$

(2.28)

This is just the statement that

$$\vec{\nabla} \times \vec{E} + \frac{\partial \vec{B}}{\partial t} = 0 ,$$

(2.29)

which is the second of the Maxwell equations in (2.4).

The other distinct possibility for assigning decomposed indices in (2.19) is to take $(\mu, \nu, \rho) = (i, j, k)$, giving

$$\partial_i F_{jk} + \partial_j F_{ki} + \partial_k F_{ij} = 0 .$$

(2.30)

Since this is totally antisymmetric in $(i, j, k)$, no generality is lost by contracting it with $\epsilon_{ijk}$, giving

$$3 \epsilon_{ijk} \partial_i F_{jk} = 0 .$$

(2.31)

From (2.14), this implies

$$3 \epsilon_{ijk} \epsilon_{j\ell k} \partial_\ell B_\ell = 0 , \quad \text{and hence} \quad 6 \partial_i B_i = 0 .$$

(2.32)

This has just reproduced the first Maxwell equation in (2.4), i.e. $\vec{\nabla} \cdot \vec{B} = 0$.

We have now demonstrated that the equations (2.18) and (2.19) are equivalent to the four Maxwell equations (2.3) and (2.4). Since (2.18) and (2.19) are written in a four-dimensional notation, it is highly suggestive that they are indeed Lorentz covariant. However, we should be a little more careful, in order to be sure about this point. Not every set of objects $V^\mu$ can be viewed as a Lorentz 4-vector, after all. The test is whether they transform properly, as in (1.91), under Lorentz transformations.

We may begin by considering the quantities $J^\mu = (\rho, J^i)$. Note first that by applying $\partial_\nu$ to the Maxwell field equation (2.18), we get identically zero on the left-hand side, since

$\epsilon_{ijm} \epsilon_{k\ell m} = \delta_{ik} \delta_j \ell - \delta_{j\ell} \delta_{ik}$, and hence $\epsilon_{ijm} \epsilon_{kjm} = 2 \delta_{ik}$.\footnote{Recall that $\epsilon_{ijm} \epsilon_{k\ell m} = \delta_{ik} \delta_j \ell - \delta_{j\ell} \delta_{ik}$, and hence $\epsilon_{ijm} \epsilon_{kjm} = 2 \delta_{ik}$.}
partial derivatives commute and $F^{\mu\nu}$ is antisymmetric. Thus from the right-hand side we get

$$\partial_\mu J^\mu = 0.$$  \hspace{1cm} (2.33)

This is the equation of charge conservation. Decomposed into the 3 + 1 language, it takes the familiar form

$$\frac{\partial \rho}{\partial t} + \vec{\nabla} \cdot \vec{J} = 0.$$  \hspace{1cm} (2.34)

By integrating over a closed 3-volume $V$ and using the divergence theorem on the second term, we learn that the rate of change of charge inside $V$ is balanced by the flow of charge through its boundary $S$:

$$\frac{\partial}{\partial t} \int_V \rho \, d^3x = - \int_S \vec{J} \cdot d\vec{S},$$  \hspace{1cm} (2.35)

where $d^3x = dx dy dz$ is the spatial 3-volume element, and $d\vec{S} = (dy dz, dz dx, dx dy)$ is the 2-area element.

Now we are in a position to show that $J^\mu = (\rho, \vec{J})$ is indeed a 4-vector. Considering $J^0 = \rho$ first, we may note that

$$dQ \equiv \rho \, dx dy dz$$  \hspace{1cm} (2.36)

is clearly Lorentz invariant, since it is an electric charge. Clearly, all Lorentz observers will agree on the number of electrons in a specified closed spatial region, and so they will agree on the amount of charge. Another quantity that is Lorentz invariant is

$$dv = dt dx dy dz,$$  \hspace{1cm} (2.37)

the 4-volume element of an infinitesimal volume in spacetime. This can be seen from the fact that the Jacobian $J$ of the transformation from $dv$ to $dv' = dt' dx' dy' dz'$ is given by

$$J = \det \left( \frac{\partial x'^\mu}{\partial x^\nu} \right) = \det(\Lambda^\mu_\nu).$$  \hspace{1cm} (2.38)

Now the defining property (1.80) of the Lorentz transformation can be written in a matrix notation as

$$\Lambda^T \eta \Lambda = \eta,$$  \hspace{1cm} (2.39)

and hence taking the determinant, we get $(\det \Lambda)^2 = 1$ and hence

$$\det \Lambda = \pm 1.$$  \hspace{1cm} (2.40)

Assuming that we restrict attention to Lorentz transformations without reflections, then they will be connected to the identity (we can take the boost velocity $\vec{v}$ to zero and/or
the rotation angle to zero and continuously approach the identity transformation), and so det Λ = 1. Thus it follows from (2.38) that for Lorentz transformations without reflections, the 4-volume element \( dv = dt dx dy dz \) is invariant.

Comparing \( dQ = \rho dx dy dz \) and \( dv = dt dx dy dz \), both of which we have argued are Lorentz invariant, we can conclude that, just as \( dt \) transforms as the 0 component of a 4-vector, so the charge density \( \rho \) must transform as the 0 component of a 4-vector under Lorentz transformations. Thus writing, as we did, that \( J^0 = \rho \), is justified.

In the same way, we may consider the spatial components \( J^i \) of the putative 4-vector \( J^\mu \). Considering \( J^1 \), for example, we know that \( J^1 dx dy \) is the current flowing through the area element \( dx dy \). Therefore in time \( dt \), there will have been a flow of charge \( J^1 dt dx dy \). Being a charge, this must be Lorentz invariant, and so it follows from the known Lorentz invariance of \( dv = dt dx dy dz \) that \( J^1 \) must transform the same way as \( dx \) under Lorentz transformations. That is, \( J^1 \) must transform as the 1 component of a 4-vector. Similar arguments apply to \( J^2 \) and \( J^3 \). (It is important in this argument that, because of the charge-conservation equation (2.33) or (2.35), the flow of charges we are discussing when considering the \( J^i \) components are the same charges we discussed when considering the \( J^0 \) component.)

We have now argued that \( J^\mu = (\rho, J^i) \) is indeed a Lorentz 4-vector, where \( \rho \) is the charge density and \( J^i \) the 3-vector current density.

Actually, the argument we have presented for showing that \( J^\mu \) is a 4-vector is a little sketchy. One should really be rather more careful about getting the orientations of the 3-area elements orthogonal to the 0, 1, 2 and 3 directions right. This involves defining the infinitesimal 3-area 4-vector

\[
d\Sigma_\mu = (dx dy dz, -dt dy dz, -dt dz dx, -dt dx dy).
\]  

Charge conservation can then be described in 4-dimensional terms, by considering an arbitrary 4-volume \( V_4 \) in spacetime, which is bounded by a 3-dimensional surface \( S_3 \). Integrating \( \partial_\mu J^\mu = 0 \) over \( V_4 \) and then using the 4-dimensional analogue of the divergence theorem gives

\[
0 = \int_{V_4} \partial_\mu J^\mu d^4 x = \int_{S_3} J^\mu d\Sigma_\mu , \tag{2.42}
\]

\(^{11}\)One can verify, by carefully generalising the usual proof of the 3-dimensional divergence theorem to four dimensions, that the signs given in the definition of \( d\Sigma_\mu \) in (2.41) are correct, where the convention is that a 3-area element such as \( dt dx dy \) is positively oriented, in the sense that \( dt dx dy \) would give a positive number when integrated over a 3-area in the \((t, x, y)\) hyperplane.
where \( d^4x = dt dx dy dz \) is the infinitesimal 4-volume element.

Now consider a 4-volume \( V_4 \) comprising the entire infinite spatial 3-volume sandwiched between an initial timelike surface at \( t = t_1 \) and a final timelike surface at \( t = t_2 \). Thus the bounding 3-surface \( S_3 \) is like an infinite-radius hyper-cylinder, comprising the two end-caps given by the infinite spatial hyperplane at \( t = t_1 \) with \(-\infty \leq x \leq \infty, -\infty \leq y \leq \infty \) and \(-\infty \leq z \leq \infty \), and the infinite spatial hyperplane at \( t = t_2 \), again with \(-\infty \leq x \leq \infty, -\infty \leq y \leq \infty \) and \(-\infty \leq z \leq \infty \); and then finally the the side of the cylinder, with \( t_1 \leq t \leq t_2 \) and \( x, y \) and \( z \) all being at infinity (the “sphere at infinity”). We shall assume that the charge and current densities are localised, and that they fall off at spatial infinity. Thus from (2.42) we then have

\[
0 = -\int_{\text{Cap 1}} J^0 dx dy dz + \int_{\text{Cap 2}} J^0 dx dy dz + \int_{\text{Sphere at infinity}} J^i d\Sigma_i. \tag{2.43}
\]

(The minus sign on the first term is because in the divergence theorem the 3-area element \( d\Sigma_i \) points outwards from the 4-volume \( V_4 \) on all of the boundary surface.) The fall-off assumptions imply the final integral vanishes, and so we have

\[
\int_{\text{Cap 1}} J^0 dx dy dz = \int_{\text{Cap 2}} J^0 dx dy dz, \quad \text{i.e.} \quad \int_{t=t_1} \rho dx dy dz = \int_{t=t_2} \rho dx dy dz, \tag{2.44}
\]

and hence we conclude that the total charge at time \( t_1 \) is the same as the total charge at time \( t_2 \). That is to say, charge is conserved.

The conclusion about charge conservation would be the same for any choice of inertial observer. That is to say, the conclusion must be independent of the choice of Lorentz frame. This implies that \( J^\mu \ d\Sigma_\mu \) must be Lorentz invariant. Since one could consider all possible timelike slicings for the “sandwich” in (2.42), one is essentially saying that \( J^\mu \ d\Sigma_\mu \) must be a Lorentz scalar where the 4-vector \( d\Sigma_\mu \) can be oriented arbitrarily. By the quotient theorem (see homework 2), it therefore follows that \( J^\mu = (\rho, J^i) \) must be a 4-vector.

At this point, we recall that by choosing the Lorenz gauge (2.11), we were able to reduce the Maxwell field equations (2.3) to (2.12). Furthermore, we can write these equations together as

\[
\Box A^\mu = -4\pi J^\mu, \tag{2.45}
\]

where

\[
A^\mu = (\phi, \vec{A}), \tag{2.46}
\]

where the d’Alembertian, or wave operator, \( \Box = \partial^\mu \partial_\mu = \partial_t \partial_t - \vec{\partial}^2 \) was introduced in (1.115). We saw that it is manifestly a Lorentz scalar operator, since it is built from the contraction
of indices on the two Lorentz-vector gradient operators. Since we have already established
that $J^\mu$ is a 4-vector, it therefore follows that $A^\mu$ is a 4-vector. Note, en passant, that the
Lorenz gauge condition (2.11) that we imposed earlier translates, in the four-dimensional
language, into
\[ \partial_\mu A^\mu = 0, \]  
(2.47)
which is nicely Lorentz invariant.

The final step is to note that our definition (2.14) is precisely consistent with (2.46) and
(2.8), if we write
\[ F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu. \]  
(2.48)
First, we note from (2.46) that because of the $\eta_{00} = -1$ needed when lowering the 0 index,
we shall have
\[ A_\mu = (\phi, \vec{A}). \]  
(2.49)
Therefore we find
\[ F_{0i} = \partial_0 A_i - \partial_i A_0 = \frac{\partial A_i}{\partial t} + \partial_i \phi = -E_i, \]
\[ F_{ij} = \partial_i A_j - \partial_j A_i = \epsilon_{ijk}(\vec{\nabla} \times \vec{A})_k = \epsilon_{ijk} B_k. \]  
(2.50)

In summary, we have shown that $J^\mu$ is a 4-vector, and hence, using (2.45), that $A^\mu$ is a
4-vector. Then, it is manifest from (2.48) that $F_{\mu\nu}$ is a 4-tensor. Hence, we have established
that the Maxwell equations, written in the form (2.18) and (2.19), are indeed expressed in
terms of 4-tensors and 4-vectors, and so the manifest Lorentz covariance of the Maxwell
equations is established.

Finally, it is worth remarking that in the 4-tensor description, the way in which the gauge
invariance arises is very straightforward. First, it is manifest that the Bianchi identity (2.19)
is solved identically by writing
\[ F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu, \]  
(2.51)
for some 4-vector $A_\mu$. This is because (2.19) is totally antisymmetric in $\mu\nu\rho$, and so, when
(2.51) is substituted into it, one gets identically zero since partial derivatives commute.
(Try making the substitution and verify this explicitly. The vanishing because of the com-
mutativity of partial derivatives is essentially the same as the reason why $\text{curl grad} \equiv 0$
and $\text{div curl} \equiv 0$.) It is also clear from (2.51) that $F_{\mu\nu}$ will be unchanged if we make the
replacement
\[ A_\mu \rightarrow A_\mu + \partial_\mu \lambda, \]  
(2.52)
where $\lambda$ is an arbitrary function of position and time. Again, the reason is that partial
derivatives commute. Comparing (2.52) with (2.49), we see that (2.52) implies
\[
\phi \longrightarrow \phi - \frac{\partial \lambda}{\partial t}, \quad A_i \longrightarrow A_i + \partial_i \lambda, \quad (2.53)
\]
and so we have reproduced the gauge transformations (2.9) and (2.10).

It should have become clear by now that all the familiar features of the Maxwell equations are equivalently described in the spacetime formulation in terms of 4-vectors and 4-tensors. The only difference is that everything is described much more simply and elegantly in the four-dimensional language.

2.4 Lorentz transformation of $\vec{E}$ and $\vec{B}$

Although for many purposes the four-dimensional description of the Maxwell equations is the most convenient, it is sometimes useful to revert to the original description in terms of $\vec{E}$ and $\vec{B}$. For example, we may easily derive the Lorentz transformation properties of $\vec{E}$ and $\vec{B}$, making use of the four-dimensional formulation. In terms of $F_{\mu\nu}$, there is no work needed to write down its behaviour under Lorentz transformations. Raising the indices for convenience, we shall have
\[
F'_{\mu\nu} = \Lambda^\mu_\rho \Lambda^\nu_\sigma F^{\rho\sigma}. \quad (2.54)
\]
From this, and the fact (see (2.14) that $F^{0i} = E_i$, $F^{ij} = \epsilon_{ijk}B_k$, we can then immediately read of the Lorentz transformations for $\vec{E}$ and $\vec{B}$.

From the expressions (1.75) for the most general Lorentz boost transformation, we may first calculate $\vec{E}'$, calculated from
\[
E'_i &= F'^{0i} = \Lambda^0_\rho \Lambda_i^i F^{\rho\sigma}, \\
&= \Lambda^0_0 \Lambda^i_k F^{0k} + \Lambda^0_k \Lambda_i^0 F^{k0} + \Lambda^0_0 \Lambda^i_k F^{k\ell}, \\
&= \gamma \left( \delta_{ik} + \frac{\gamma - 1}{v^2} v_i v_k \right) E_k - \gamma^2 v_i v_k E_k - \gamma v_k \left( \delta_{i\ell} + \frac{\gamma - 1}{v^2} v_i v_\ell \right) \epsilon_{k\ell m} B_m, \\
&= \gamma E_i + \gamma \epsilon_{ijk} v_j B_k - \frac{\gamma - 1}{v^2} v_i v_k E_k. \quad (2.55)
\]
(Note that because $F^{\mu\nu}$ is antisymmetric, there is no $F^{00}$ term on the right-hand side on
the second line.) Thus, in terms of 3-vector notation, the Lorentz boost transformation of the electric field is given by
\[
\vec{E}' = \gamma (\vec{E} + \vec{v} \times \vec{B}) - \frac{\gamma - 1}{v^2} (\vec{v} \cdot \vec{E}) \vec{v}. \quad (2.56)
\]
An analogous calculation shows that the Lorentz boost transformation of the magnetic field is given by

\[ \vec{B}' = \gamma (\vec{B} - \vec{v} \times \vec{E}) - \frac{\gamma - 1}{v^2} (\vec{v} \cdot \vec{B}) \vec{v}. \]  

(2.57)

Suppose, for example, that in the frame \( S \) there is just a magnetic field \( \vec{B} \), while \( \vec{E} = 0 \). An observer in a frame \( S' \) moving with uniform velocity \( \vec{v} \) relative to \( S \) will therefore observe not only a magnetic field, given by

\[ \vec{B}' = \gamma \vec{B} - \frac{\gamma - 1}{v^2} (\vec{v} \cdot \vec{B}) \vec{v}, \]  

(2.58)

but also an electric field, given by

\[ \vec{E}' = \gamma \vec{v} \times \vec{B}. \]  

(2.59)

This, of course, is the principle of the dynamo.\(^{12}\)

It is instructive to write out the Lorentz transformations explicitly in the case when the boost is along the \( x \) direction, \( \vec{v} = (v, 0, 0) \). Equations (2.56) and (2.57) become

\[ \begin{align*}
E'_x &= E_x, & E'_y &= \gamma (E_y - vB_z), & E'_z &= \gamma (E_z + vB_y), \\
B'_x &= B_x, & B'_y &= \gamma (B_y + vE_z), & B'_z &= \gamma (B_z - vE_y).
\end{align*} \]  

(2.60)

2.5 The Lorentz force

Consider a point particle following the path, or worldline, \( x^\mu = x^\mu(\tau) \) in Minkowski spacetime. As we saw earlier, its 4-velocity is given by

\[ U^\mu = \frac{dx^\mu(\tau)}{d\tau} = (\gamma, \gamma \vec{u}), \quad \text{where} \quad \gamma = \frac{1}{\sqrt{1 - u^2}}, \]  

(2.61)

where \( u^i = dx^i/dt \) is its 3-velocity. Multiplying (2.61) by the rest mass \( m \) of the particle gives another 4-vector, namely the 4-momentum

\[ p^\mu = mU^\mu = (m\gamma, m\gamma \vec{u}). \]  

(2.62)

The 0 component \( p^0 = m\gamma \) is called the relativistic energy \( E \), and the spatial components \( p^i = m\gamma u^i \) are called the relativistic 3-momentum. Note that since \( U^\mu U_\mu = -1 \), we shall have

\[ p^\mu p_\mu = -m^2. \]  

(2.63)

\(^{12}\)In a practical dynamo the rotor is moving with a velocity \( \vec{v} \) which is much less than the speed of light, i.e. \( |\vec{v}| \ll 1 \) in natural units. This means that the gamma factor \( \gamma = (1 - v^2)^{-1/2} \) is approximately equal to unity in such cases.
We now define the relativistic 4-force \( f^\mu \) acting on the particle to be
\[
f^\mu = \frac{dp^\mu}{d\tau},
\]  
where \( \tau \) is the proper time. Clearly \( f^\mu \) is indeed a 4-vector, since it is the 4-vector \( dp^\mu \) divided by the scalar \( d\tau \).

Using (2.62), we can write the 4-force as
\[
f^\mu = \left( m\gamma^3 \vec{u} \cdot \frac{d\vec{u}}{d\tau}, m\gamma^3 \vec{u} \cdot \frac{d\vec{u}}{d\tau} \vec{u} + m\gamma \frac{d\vec{u}}{d\tau} \right).
\]  
It follows that if we move to the instantaneous rest frame of the particle, i.e. the frame in which \( \vec{u} = 0 \) at the particular moment we are considering, then \( f^\mu \) reduces to
\[
f^\mu \bigg|_{\text{rest frame}} = (0, \vec{F}),
\]
where
\[
\vec{F} = m\frac{d\vec{u}}{dt}
\]
is the Newtonian force measured in the rest frame of the particle. Thus, we should interpret the 4-force physically as describing the Newtonian 3-force when measured in the instantaneous rest frame of the accelerating particle.

If we now suppose that the particle has electric charge \( e \), and that it is moving under the influence of an electromagnetic field \( F_{\mu\nu} \), then its motion is given by the Lorentz force equation
\[
f^\mu = eF^{\mu\nu} U_\nu.
\]  
One can more or less justify this equation on the grounds of “what else could it be?”, since we know that there must exist a relativistic equation (i.e. a Lorentz covariant equation) that describes the motion. In fact it is easy to see that (2.68) is correct. We calculate the spatial components:
\[
f^i = eF^{i\nu} U_\nu = eF^{i0} U_0 + eF^{ij} U_j,
\]  
\[
= e(-E_i)(-\gamma) + e\epsilon_{ijk} B_k \gamma u_j,
\]
and thus
\[
\vec{f} = e\gamma (\vec{E} + \vec{u} \times \vec{B}).
\]

\(^{13}\)Note that we can replace the proper time \( \tau \) by the coordinate time \( t \) in the instantaneous rest frame, since \( d\tau = dt/\gamma \), and \( \gamma = 1 \) when \( \vec{u} = 0 \).
But $f^\mu = dp^\mu /d\tau$, and so $\vec{f} = d\vec{p}/d\tau = \gamma d\vec{p}/dt$ (recall from section 1.6 that $d\tau = dt/\gamma$) and so we have
\[
\frac{d\vec{p}}{dt} = e (\vec{E} + \vec{u} \times \vec{B}),
\]
where $d\vec{p}/dt$ is the rate of change of the relativistic 3-momentum $\vec{p} = m\gamma u\vec{u}$. This is indeed the standard expression for the motion of a charged particle under the Lorentz force.

2.6 Action principle for charged particles

In this section, we shall show how the equations of motion for a charged particle moving in an electromagnetic field can be derived from an action principle. To begin, we shall consider an uncharged particle of mass $m$, with no forces acting on it. It will, of course, move in a straight line. It turns out that its equation of motion can be derived from the Lorentz-invariant action
\[
S = -m \int_{\tau_1}^{\tau_2} d\tau,
\]
where $\tau$ is the proper time along the trajectory $x^\mu(\tau)$ of the particle, starting at proper time $\tau = \tau_1$ and ending at $\tau = \tau_2$. The action principle then states that if we consider all possible paths between the initial and final spacetime points on the path, then the actual path followed by the particle will be such that the action $S$ is stationary. In other words, if we consider small variations of the path around the actual path, then to first order in the variations we shall have $\delta S = 0$.

To see how this works, we note that $d\tau^2 = dt^2 - dx^i dx^i = dt^2(1 - v_i v_i) = dt^2(1 - v^2)$, where $v_i = dx^i/dt$ is the 3-velocity of the particle. Thus $d\tau = dt/\gamma$ where $\gamma = (1 - v^2)^{-1/2}$ and so
\[
S = -m \int_{t_1}^{t_2} \frac{dt}{\gamma} = -m \int_{t_1}^{t_2} (1 - v^2)^{1/2} dt = -m \int_{t_1}^{t_2} (1 - \dot{x}^i \dot{x}^i)^{1/2} dt.
\]
In other words, the Lagrangian $L$, for which $S = \int_{t_1}^{t_2} L dt$, is given by
\[
L = -\frac{m}{\gamma} = -m(1 - \dot{x}^i \dot{x}^i)^{1/2}.
\]

As a check, if we expand (2.74) for small velocities (i.e. small compared with the speed of light, so $|\dot{x}^i| << 1$), we shall have
\[
L = -m + \frac{1}{2} mv^2 + \cdots.
\]
Since the Lagrangian is given by $L = T - V$ we see that $T$ is just the usual kinetic energy $\frac{1}{2} mv^2$ for a non-relativistic particle of mass $m$, while the potential energy is just $m$. Of course if we were not using units where the speed of light were unity, this energy would be
Since it is just a constant, this rest-mass energy of the particle does not affect the equations of motion that will follow from the action principle.

Now let us consider small variations $\delta x^i(t)$ around the path $x^i(t)$ followed by the particle. The action will vary according to

$$\delta S = m \int_{t_1}^{t_2} (1 - \dot{x}^j \dot{x}_j)^{-1/2} \dot{x}^i \delta \dot{x}^i \, dt .$$

Integrating by parts then gives

$$\delta S = -m \int_{t_1}^{t_2} \frac{d}{dt} \left[(1 - \dot{x}^j \dot{x}_j)^{-1/2} \dot{x}^i\right] \delta x^i \, dt + m \left[(1 - \dot{x}^j \dot{x}_j)^{-1/2} \dot{x}^i \delta x^i\right]_{t_1}^{t_2} .$$

As usual in an action principle, we restrict to variations of the path that vanish at the endpoints, so $\delta x^i(t_1) = \delta x^i(t_2) = 0$ and the boundary term can be dropped. The variation $\delta x^i$ is allowed to be otherwise arbitrary in the time interval $t_1 < t < t_2$, and so we conclude from the requirement of stationary action $\delta S = 0$ that

$$\frac{d}{dt} \left(m(1 - \dot{x}^j \dot{x}_j)^{-1/2} \dot{x}^i\right) = 0 .$$

Now, recalling that we define $\gamma = (1 - v^2)^{-1/2}$, we see that

$$\frac{d}{dt}(m \gamma \vec{v}) = 0 ,$$

or, in other words,

$$\frac{d}{dt} \vec{p} = 0 ,$$

where $\vec{p} = m \gamma \vec{v}$ is the relativistic 3-momentum. We have, of course, derived the equation for straight-line motion in the absence of any forces acting.

Now we extend the discussion to the case of a particle of mass $m$ and charge $e$, moving under the influence of an electromagnetic field $F_{\mu\nu}$. This field will be written in terms of a 4-vector potential:

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu .$$

The action will now be the sum of the free-particle action (2.73) above plus a term describing the interaction of the particle with the electromagnetic field. The total action turns out to be

$$S = \int_{\tau_1}^{\tau_2} \left(-md\tau + eA_\mu dx^\mu\right) .$$

Note that it is again Lorentz invariant.

From (2.49) we have $A^\mu = (\phi, \vec{A})$ and hence $A_\mu = (-\phi, \vec{A})$, and so

$$A_\mu dx^\mu = A_\mu \frac{dx^\mu}{dt} \, dt = (A_0 + A_i \dot{x}^i) \, dt = (-\phi + A_i \dot{x}^i) \, dt .$$
Thus we have

\[ S = \int_{t_1}^{t_2} L dt \]

with the Lagrangian \( L \) given by

\[ L = -m(1 - \dot{x}^j \dot{x}^j)^{1/2} - e\phi + eA_i \dot{x}^i, \quad (2.84) \]

where potentials \( \phi \) and \( A_i \) depend on \( t \) and \( x \). The first-order variation of the action under a variation \( \delta x^i \) in the path gives

\[
\delta S = \int_{t_1}^{t_2} \left[ m(1 - \dot{x}^j \dot{x}^j)^{-1/2} \dot{x}^i \delta x^i - e\partial_i \phi \delta x^i + eA_i \delta \dot{x}^i + e\partial_j A_i \dot{x}^j \delta x^i \right] dt,
\]

\( = \int_{t_1}^{t_2} \left[ -\frac{d}{dt}(m\gamma \dot{x}^i) - e\partial_i \phi - e\frac{dA_i}{dt} + e\partial_i A_j \dot{x}^j \right] \delta x^i dt. \quad (2.85) \)

(We have dropped the boundary terms immediately, since \( \delta x^i \) is again assumed to vanish at the endpoints.) Thus the principle of stationary action \( \delta S = 0 \) implies

\[
\frac{d(m\gamma \dot{x}^i)}{dt} = -e\partial_i \phi - \frac{dA_i}{dt} + e\partial_i A_j \dot{x}^j. \quad (2.86)
\]

Now, the total time derivative \( \frac{dA_i}{dt} \) has two contributions, and we may write it as

\[
\frac{dA_i}{dt} = \partial A_i \frac{dt}{dt} + \partial_j A_i \frac{dx^j}{dt} = \partial A_i \frac{dt}{dt} + \partial_j A_i \dot{x}^j. \quad (2.87)
\]

This arises because first of all, \( A_i \) can depend explicitly on the time coordinate; this contribution is \( \partial A_i / \partial t \). Additionally, \( A_i \) depends on the spatial coordinates \( x^i \), and along the path followed by the particle, \( x^i \) depends on \( t \) because the path is \( x^i = x^i(t) \). This accounts for the second term.

Putting all this together, we have

\[
\frac{d(m\gamma \dot{x}^i)}{dt} = e \left( -\partial_i \phi - \partial A_i \right) + e(\partial_i A_j - \partial_j A_i) \dot{x}^j, \]

\[ = e(E_i + \epsilon_{ijk} \dot{x}^j B_k). \quad (2.88) \]

In other words, we have

\[
\frac{d\vec{p}}{dt} = e(\vec{E} + \vec{v} \times \vec{B}), \quad (2.89)
\]

which is the Lorentz force equation (2.71).

It is worth noting that although we gave a “three-dimensional” derivation of the equations of motion following from the action (2.82), we can also instead directly derive the four-dimensional equation \( dp^\mu / d\tau = eF^{\mu\nu}U_\nu \). To begin, we note that since \( d\tau^2 = -\eta_{\mu\nu} dx^\mu dx^\nu \), its variation variation under a variation of the path \( x^\mu(\tau) \) in spacetime gives \( 2d\tau \delta(d\tau) = -2\eta_{\mu\nu} d(\delta x^\mu) dx^\nu \), and so dividing by \( 2d\tau \) gives

\[
\delta(d\tau) = -\eta_{\mu\nu} \frac{dx^\nu}{d\tau} d(\delta x^\mu), \]

\[ = -U_\mu d(\delta x^\mu), \quad (2.90) \]
where \( U_\mu \) is the 4-velocity. Thus the variation of the action (2.82) gives

\[
\delta S = \int_{\tau_1}^{\tau_2} \left( mU_\mu d\delta x^\mu + eA_\mu d\delta x^\mu + e\partial_\nu A_\mu \delta x^\nu dx^\mu \right),
\]

\[
= \int_{\tau_1}^{\tau_2} \left( d(mU_\mu \delta x^\mu) - mdU_\mu \delta x^\mu + d(eA_\mu \delta x^\mu) - edA_\mu \delta x^\mu + e\partial_\mu A_\nu \delta x^\mu dx^\nu \right),
\]

\[
= \int_{\tau_1}^{\tau_2} \left( -mdU_\mu \delta x^\mu - edA_\mu \delta x^\mu + e\partial_\mu A_\nu \delta x^\mu dx^\nu \right),
\]

\[
= \int_{\tau_1}^{\tau_2} \left( -m \frac{dU_\mu}{d\tau} - e \frac{dA_\mu}{d\tau} + e\partial_\mu A_\nu \frac{dx^\nu}{d\tau} \right) dx^\mu d\tau,
\]

(2.91)

where we have dropped the boundary terms \( \int_{\tau_1}^{\tau_2} d(mU_\mu \delta x^\mu + eA_\mu \delta x^\mu) \) in getting to the third line, since they integrate to give \( [mU_\mu \delta x^\mu + eA_\mu \delta x^\mu]_{\tau_1}^{\tau_2} \) and therefore vanish, as we assume \( \delta x^\mu = 0 \) at the initial and final proper times \( \tau_1 \) and \( \tau_2 \). Now by the chain rule

\[
\frac{dA_\mu}{d\tau} = \partial_\nu A_\mu \frac{dx^\nu}{d\tau} = \partial_\nu A_\mu U^\nu,
\]

(2.92)

and so

\[
\delta S = \int_{\tau_1}^{\tau_2} \left( -m \frac{dU_\mu}{d\tau} - e\partial_\nu A_\mu U^\nu + e\partial_\mu A_\nu U^\nu \right) \delta x^\mu d\tau,
\]

\[
= \int_{\tau_1}^{\tau_2} \left( -m \frac{dU_\mu}{d\tau} + eF_{\mu\nu} U^\nu \right) d\tau.
\]

(2.93)

Requiring \( \delta S = 0 \) for all variations (that vanish at the endpoints) we therefore obtain the equation of motion

\[
m \frac{dU_\mu}{d\tau} = eF_{\mu\nu} U^\nu.
\]

(2.94)

Thus we have reproduced the Lorentz force equation in its four-dimensionally covariant form

\[
\frac{dp^\mu}{d\tau} = eF^{\mu\nu} U_\nu,
\]

(2.95)

where \( p^\mu = mU^\mu \) is the 4-momentum.

### 2.7 Gauge invariance of the action

In writing down the relativistic action (2.82) for a charged particle we had to make use of the 4-vector potential \( A_\mu \). This is itself not physically observable, since, as we noted earlier, \( A_\mu \) and \( A'_\mu = A_\mu + \partial \lambda \) describe the same physics, where \( \lambda \) is any arbitrary function in spacetime, since \( A_\mu \) and \( A'_\mu \) give rise to the same electromagnetic field \( F_{\mu\nu} \). One might worry, therefore, that the action itself would be gauge dependent, and therefore might not properly describe the required physical situation. However, all is in fact well. This already
can be seen from the fact that, as we demonstrated, the variational principle for the action (2.82) does in fact produce the correct gauge-invariant Lorentz force equation (2.71).

It is instructive also to examine the effects of a gauge transformation directly at the level of the action. If we make the gauge transformation $A_\mu \rightarrow A'_\mu = A_\mu + \partial_\mu \lambda$, we see from (2.82) that the action $S$ transforms to $S'$ given by

$$S' = \int_{\tau_1}^{\tau_2} (-md\tau + eA_\mu dx^\mu + e\partial_\mu \lambda dx^\mu),$$

$$= S + e \int_{\tau_1}^{\tau_2} \partial_\mu \lambda dx^\mu = e \int_{\tau_1}^{\tau_2} d\lambda,$$

and so

$$S' = S + e[\lambda(\tau_2) - \lambda(\tau_1)].$$

The simplest situation to consider is where we restrict ourselves to gauge transformations that vanish at the endpoints, in which case the action will be gauge invariant, $S' = S$. Even if $\lambda$ is non-vanishing at the endpoints, we see from (2.97) that $S$ and $S'$ merely differ by a constant that depends solely on the values of $\lambda$ at $\tau_1$ and $\tau_2$. Clearly, the addition of this constant has no effect on the equations of motion that one derives from $S'$.

### 2.8 Canonical momentum, and Hamiltonian

Given any Lagrangian $L(x^i, \dot{x}^i, t)$ one defines the canonical momentum $\pi_i$ as

$$\pi_i = \frac{\partial L}{\partial \dot{x}^i}. \tag{2.98}$$

The relativistic Lagrangian for the charged particle is given by (2.84), and so we have

$$\pi_i = m(1 - \dot{x}^j \dot{x}^j)^{-1/2} \dot{x}^i + eA_i, \tag{2.99}$$

or, in other words,

$$\pi_i = m\gamma \dot{x}^i + eA_i, \tag{2.100}$$

$$= p_i + eA_i, \tag{2.101}$$

where $p_i$ as usual is the standard mechanical relativistic 3-momentum of the particle.

As usual, the Hamiltonian for the system is given by

$$H = \pi_i \dot{x}^i - L, \tag{2.102}$$

and so we find

$$H = m\gamma \dot{x}^i \dot{x}^i + \frac{m}{\gamma} + e\phi, \tag{2.103}$$

$$= m\gamma v^2 + \frac{m}{\gamma} + e\phi.$$
Now, \( m\gamma v^2 + m/\gamma = m\gamma(v^2 + (1 - v^2)) = m\gamma \), so we have

\[
H = m\gamma + e\phi. \tag{2.104}
\]

As always, the Hamiltonian is to be viewed as a function of the coordinates \( x^i \) and the canonical momenta \( \pi_i \). To express \( \gamma \) in terms of \( \pi_i \), we note from (2.100) that \( m\gamma \dot{x}^i = \pi_i - eA_i \), and so squaring, we get \( m^2\gamma^2 v^2 = m^2v^2/(1 - v^2) = (\pi_i - eA_i)^2 \). Solving for \( v^2 \), and hence for \( \gamma \), we find that \( m^2\gamma^2 = (\pi_i - eA_i)^2 + m^2 \), and so finally, from (2.104), we arrive at the Hamiltonian

\[
H = \sqrt{(\pi_i - eA_i)^2 + m^2} + e\phi, \tag{2.105}
\]

with \( H \) expressed as a function of the coordinates \( x^i \) and the canonical momenta \( \pi_i \).

Note that Hamilton’s equations, which will necessarily give rise to the same Lorentz force equations of motion we encountered previously, are given by

\[
\frac{\partial H}{\partial \pi_i} = \dot{x}^i, \quad \frac{\partial H}{\partial x^i} = -\dot{\pi}_i. \tag{2.106}
\]

As a check of the correctness of the Hamiltonian (2.105) we may examine it in the non-relativistic limit when \((\pi_i - eA_i)^2\) is much less than \(m^2\). We then extract an \(m^2\) factor from inside the square root in \(\sqrt{(\pi_i - eA_i)^2 + m^2}\) and expand to get

\[
H = m\sqrt{1 + (\pi_i - eA_i)^2/m^2} + e\phi, \\
= m + \frac{1}{2m} (\pi_i - eA_i)^2 + e\phi + \cdots. \tag{2.107}
\]

The first term is the rest-mass energy, which is just a constant, and the remaining terms presented explicitly in (2.107) give the standard non-relativistic Hamiltonian for a charged particle

\[
H_{\text{non-rel.}} = \frac{1}{2m} (\pi_i - eA_i)^2 + e\phi. \tag{2.108}
\]

This should be familiar from quantum mechanics, when one writes down the Schrödinger equation for the wave function for a charged particle in an electromagnetic field.

### 3 Particle Motion in Static Electromagnetic Fields

In this chapter, we discuss the motion of a charged particle in static (i.e. time-independent) electromagnetic fields.
3.1 Description in terms of potentials

If we are describing static electric and magnetic fields, \( \vec{E} = \vec{E}(\vec{r}) \) and \( \vec{B} = \vec{B}(\vec{r}) \), it is natural (and always possible) to describe them in terms of scalar and 3-vector potentials that are also static, \( \phi = \phi(\vec{r}) \), \( \vec{A} = \vec{A}(\vec{r}) \). Thus we write

\[
\begin{align*}
\vec{E} &= -\vec{\nabla}\phi - \frac{\partial \vec{A}}{\partial t} = -\vec{\nabla}\phi(\vec{r}), \\
\vec{B} &= \vec{\nabla} \times \vec{A}(\vec{r}).
\end{align*}
\] (3.1)

We can still perform gauge transformations, as given in (2.9) and (2.10). The most general gauge transformation that preserves the time-independence of the potentials is therefore given by taking the parameter \( \lambda \) to be of the form

\[
\lambda(\vec{r}, t) = \lambda(\vec{r}) + k t,
\] (3.2)

where \( k \) is an arbitrary constant. This implies that \( \phi \) and \( \vec{A} \) will transform according to

\[
\phi \rightarrow \phi - k, \quad \vec{A} \rightarrow \vec{A} + \vec{\nabla}\lambda(\vec{r}).
\] (3.3)

Note, in particular, that the electrostatic potential \( \phi \) can just be shifted by an arbitrary constant. This is the familiar freedom that one typically uses to set \( \phi = 0 \) at infinity.

Recall that the Hamiltonian for a particle of mass \( m \) and charge \( e \) in an electromagnetic field is given by (2.104)

\[
H = m \gamma + e\phi,
\] (3.4)

where \( \gamma = (1 - v^2)^{-1/2} \). In the present situation with static fields, the Hamiltonian does not depend explicitly on time, i.e. \( \partial H/\partial t = 0 \). It then follows that the Hamiltonian is conserved (i.e. it is the same at all times) since we have (by the chain rule)

\[
\frac{dH}{dt} = \frac{\partial H}{\partial t} + \frac{\partial H}{\partial x^i} \dot{x}^i + \frac{\partial H}{\partial \pi_i} \dot{\pi}_i,
\]

\[
= 0 - \dot{\pi}_i \dot{x}^i + \dot{x}^i \dot{\pi}_i = 0.
\] (3.5)

(We used the Hamilton equations (2.106) in getting to the second line.) This time-independent quantity \( H \) is then just the energy \( \mathcal{E} \) of the system:

\[
\mathcal{E} \equiv H = m \gamma + e\phi.
\] (3.6)

We may think of the first term in \( \mathcal{E} \) as being the mechanical term,

\[
\mathcal{E}_{\text{mech}} = m \gamma,
\] (3.7)
since this is just the total energy of a particle of rest mass $m$ moving with velocity $\vec{v}$. The second term, $e\phi$, is the contribution to the total energy from the electric field. Note that the magnetic field, described by the 3-vector potential $\vec{A}$, does not contribute to the conserved energy. This is because the magnetic field $\vec{B}$ does no work on the charge:

Recall that the Lorentz force equation can be written as

$$\frac{d(m\gamma v^i)}{dt} = e(E_i + \epsilon_{ijk} v^j B_k).$$  \hspace{1cm} (3.8)

Multiplying by $v^i$ we therefore have

$$m\gamma v^i \frac{dv^i}{dt} + m v^i v^i \frac{d\gamma}{dt} = e v^i E_i.$$  \hspace{1cm} (3.9)

Now $\gamma = (1 - v^2)^{-1/2}$, so

$$\frac{d\gamma}{dt} = (1 - v^2)^{-3/2} v^i \frac{dv^i}{dt} = \gamma^3 \gamma v^i \frac{dv^i}{dt},$$  \hspace{1cm} (3.10)

and so (3.9) gives

$$m \frac{d\gamma}{dt} = e v^i E_i.$$  \hspace{1cm} (3.11)

Since $E_{\text{mech}} = m\gamma$, and $m$ is a constant, we therefore have

$$\frac{dE_{\text{mech}}}{dt} = e \vec{v} \cdot \vec{E}.$$  \hspace{1cm} (3.12)

Thus, the mechanical energy of the particle is changed only by the electric field, and not by the magnetic field.

Note that another (and equivalent) derivation of the constancy of $\mathcal{E} = m\gamma + e\phi$ is as follows:

$$\frac{d\mathcal{E}}{dt} = \frac{d(m\gamma)}{dt} + e \frac{d\phi}{dt}$$

$$= \frac{dE_{\text{mech}}}{dt} + e \partial_i \phi \frac{dx^i}{dt},$$

$$= e \vec{v} \cdot \vec{E} - e \vec{v} \cdot \vec{E} = 0.$$  \hspace{1cm} (3.13)

### 3.2 Particle motion in static uniform $\vec{E}$ and $\vec{B}$ fields

Let us consider the case where a charged particle is moving in static (i.e. time-independent) uniform $\vec{E}$ and $\vec{B}$ fields. In other words, $\vec{E}$ and $\vec{B}$ are constant vectors, independent of time and of position. In this situation, it is easy to write down explicit expressions for the corresponding scalar and 3-vector potentials. For the scalar potential, we can take

$$\phi = -\vec{E} \cdot \vec{r} = -E_i x^i.$$  \hspace{1cm} (3.14)
Clearly this gives the correct electric field, since

$$-\partial_i \phi = \partial_i (E_j x^j) = E_j \partial_i x^j = E_j \delta_{ij} = E_i .$$  \hspace{1cm} (3.15)

(It is, of course, essential that $E_j$ is constant for this calculation to be valid.)

Turning now to the uniform $\vec{B}$ field, it is easily seen that this can be written as $\vec{B} = \vec{\nabla} \times \vec{A}$, with the 3-vector potential given by

$$\vec{A} = \frac{1}{2} \vec{B} \times \vec{r} .$$  \hspace{1cm} (3.16)

It is easiest to check this using index notation. We have

$$(\vec{\nabla} \times \vec{A})_i = \epsilon_{ijk} \partial_j A_k = \epsilon_{ijk} \partial_j \left( \frac{1}{2} \epsilon_{k\ell m} B_\ell x^m \right) ,$$

$$= \frac{1}{2} \epsilon_{ijk} \epsilon_{\ell m k} B_\ell \partial_j x^m = \frac{1}{2} \epsilon_{ijk} \epsilon_{\ell jk} B_\ell ,$$

$$= \delta_{i\ell} B_\ell = B_i .$$  \hspace{1cm} (3.17)

Of course the potentials we have written above are not unique, since we can still perform gauge transformations. If we restrict attention to transformations that maintain the time-independence of $\phi$ and $\vec{A}$, then for $\phi$ the only remaining freedom is to add an arbitrary constant to $\phi$. For the 3-vector potential, we can still add $\vec{\nabla} \lambda(\vec{r})$ to $\vec{A}$, where $\lambda(\vec{r})$ is an arbitrary function of position. It is sometimes helpful, for calculational reasons, to do this. Suppose, for example, that the uniform $\vec{B}$ field lies along the $z$ axis: $\vec{B} = (0,0,B)$. From (3.16), we may therefore write the 3-vector potential

$$\vec{A} = (-\frac{1}{2} B y, \frac{1}{2} B x, 0) .$$  \hspace{1cm} (3.18)

Another choice is to take $\vec{A}' = \vec{A} + \vec{\nabla} \lambda(\vec{r})$, with $\lambda = -\frac{1}{2} B x y$. This gives

$$\vec{A}' = (-B y, 0, 0) .$$  \hspace{1cm} (3.19)

One easily verifies that indeed $\vec{\nabla} \times \vec{A}' = (0,0,B)$.

### 3.2.1 Motion in a static uniform electric field

From the Lorentz force equation, we shall have

$$\frac{d\vec{p}}{dt} = e \vec{E} ,$$  \hspace{1cm} (3.20)

where $\vec{p} = m \gamma \vec{v}$ is the relativistic 3-momentum. Without loss of generality, we may take the electric field to lie along the $x$ axis, and so we will have

$$\frac{dp_x}{dt} = e E , \quad \frac{dp_y}{dt} = 0 , \quad \frac{dp_z}{dt} = 0 .$$  \hspace{1cm} (3.21)
Since $p_y$ and $p_z$ are therefore constants, we can without loss of generality rotate the coordinate system around the $x$ axis so that $p_z = 0$. Thus we may integrate (3.21 to give

$$p_x = eEt, \quad p_y = \bar{p}, \quad p_z = 0,$$

(3.22)

where $\bar{p}$ is a constant. Note that when integrating $dp_x/dt$, we have fixed the unimportant constant of integration by choosing the origin for the time coordinate $t$ such that $p_x = 0$ at $t = 0$.

Recalling that the 4-momentum is given by $p^\mu = (m\gamma, \vec{p}) = (\mathcal{E}_{\text{mech}}, \bar{p})$, and that $p^\mu p_\mu = m^2 U^\mu U_\mu = -m^2$, we see that $-\mathcal{E}_{\text{mech}}^2 + \vec{p} \cdot \vec{p} = -m^2$, and so

$$\mathcal{E}_{\text{mech}} = \sqrt{m^2 + p_x^2 + p_y^2} = \sqrt{m^2 + \bar{p}^2 + (eEt)^2}.$$  (3.23)

Hence we may write

$$\mathcal{E}_{\text{mech}} = \sqrt{\mathcal{E}_0^2 + (eEt)^2},$$  (3.24)

where $\mathcal{E}_0^2 = m^2 + \bar{p}^2$ is the square of the mechanical energy at time $t = 0$.

We have $\vec{p} = m\gamma \vec{v} = \mathcal{E}_{\text{mech}} \vec{v}$, and so $p_x = \mathcal{E}_{\text{mech}} dx/dt$ and therefore

$$\frac{dx}{dt} = \frac{p_x}{\mathcal{E}_{\text{mech}}} = \frac{eEt}{\sqrt{\mathcal{E}_0^2 + (eEt)^2}},$$  (3.25)

which can be integrated to give

$$x = \frac{1}{eE} \sqrt{\mathcal{E}_0^2 + (eEt)^2}.$$  (3.26)

(The constant of integration has been absorbed into a choice of origin for the $x$ coordinate.)

Note from (3.25) that the $x$-component of the 3-velocity asymptotically approaches 1 as $t$ goes to infinity. Thus the particle is accelerated closer and closer to the speed of light, but never reaches it.

We also have

$$\frac{dy}{dt} = \frac{p_y}{\mathcal{E}_{\text{mech}}} = \frac{\bar{p}}{\sqrt{\mathcal{E}_0^2 + (eEt)^2}}.$$  (3.27)

This can be integrated by changing variable from $t$ to $u$, defined by

$$eEt = \mathcal{E}_0 \sinh u.$$  (3.28)

This gives $y = \bar{p} u/(eE)$, and hence

$$y = \frac{\bar{p}}{eE} \arcsinh \left( \frac{eEt}{\mathcal{E}_0} \right).$$  (3.29)

(Again, the constant of integration has been absorbed into the choice of origin for $y$.)
The solutions (3.26) and (3.29) for $x$ and $y$ as functions of $t$ can be combined to give $x$ as a function of $y$, leading to

$$x = \frac{\mathcal{E}_0}{eE} \cosh \left( \frac{eEy}{\bar{p}} \right).$$  \hspace{1cm} (3.30)

This is a catenary.

In the non-relativistic limit when $|v| << 1$, we have $\bar{p} \approx m\bar{v}$ and then, expanding (3.30) we find the standard “Newtonian” parabolic motion

$$x \approx \text{constant} + \frac{eE}{2m\bar{v}^2} y^2. \hspace{1cm} (3.31)$$

### 3.2.2 Motion in a static uniform magnetic field

From the Lorentz force equation we shall have

$$\frac{d\bar{p}}{dt} = e\bar{v} \times \bar{B}. \hspace{1cm} (3.32)$$

Recalling (3.11), we see that in the absence of an electric field we shall have $\gamma = \text{constant}$, and hence $d\bar{p}/dt = d(m\gamma\bar{v})/dt = m\gamma d\bar{v}/dt$, leading to

$$\frac{d\bar{v}}{dt} = \frac{e}{m\gamma} \bar{v} \times \bar{B} = \frac{e}{\mathcal{E}} \bar{v} \times \bar{B}, \hspace{1cm} (3.33)$$

since $\mathcal{E} = m\gamma + e\phi = m\gamma$ (a constant) here.

Without loss of generality we may choose the uniform $\bar{B}$ field to lie along the $z$ axis: $\bar{B} = (0, 0, B)$. Defining

$$\omega = \frac{eB}{\mathcal{E}} = \frac{eB}{m\gamma}, \hspace{1cm} (3.34)$$

we then find

$$\frac{dv_x}{dt} = \omega v_y, \hspace{0.5cm} \frac{dv_y}{dt} = -\omega v_x, \hspace{0.5cm} \frac{dv_z}{dt} = 0. \hspace{1cm} (3.35)$$

From this, it follows that

$$\frac{d(v_x+i v_y)}{dt} = -i\omega (v_x+i v_y), \hspace{1cm} (3.36)$$

and so the first two equations in (3.35) can be integrated to give

$$v_x+i v_y = v_0 e^{-i(\omega t+\alpha)}, \hspace{1cm} (3.37)$$

where $v_0$ is a real constant, and $\alpha$ is a constant (real) phase. Thus after further integrations we obtain

$$x = x_0 + r_0 \sin(\omega t + \alpha), \hspace{0.5cm} y = y_0 + r_0 \cos(\omega t + \alpha), \hspace{0.5cm} z = z_0 + \bar{v}_z t, \hspace{1cm} (3.38)$$
for constants \( r_0, x_0, y_0, z_0 \) and \( \tilde{v}_z \), with

\[
    r_0 = \frac{\frac{v_0}{\omega}} = \frac{m\gamma v_0}{eB} = \frac{\tilde{p}}{eB},
\]

(3.39)

where \( \tilde{p} \) is the magnitude of the relativistic 3-momentum in the \((x, y)\) plane. The particle therefore follows a helical path, of radius \( r_0 \), twisting along the \( z \) axis.

### 3.2.3 Motion in uniform \( \vec{E} \) and \( \vec{B} \) fields

Having considered the case of particle motion in a uniform \( \vec{E} \) field, and in a uniform \( \vec{B} \) field, we may also consider the situation of motion in uniform \( \vec{E} \) and \( \vec{B} \) fields together. To discuss this in detail is quite involved, and we shall not pursue it extensively here. In fact a relatively simple way to study this general case is to work directly in the 4-dimensional language, solving

\[
    \frac{dU^\mu}{d\tau} = \frac{e}{m} F^\mu_\nu U^\nu,
\]

(3.40)

where \( U = \frac{d\xi^\mu}{d\tau} \) is the 4-velocity of the particle. Since we are assuming \( \vec{E} \) and \( \vec{B} \) are uniform, constant, fields it follows that \( F^\mu_\nu \) is a constant tensor. See Homework 4, where this approach is explored further.

One can, of course, still approach the problem from a 3-dimensional standpoint. The equations can become quite complicated in general. Here, we consider the situation where we take

\[
    \vec{B} = (0, 0, B), \quad \vec{E} = (0, E_y, E_z),
\]

(3.41)

(there is no loss of generality in choosing axes so that this is the case), and we make the simplifying assumption that the motion is non-relativistic, i.e. \(|\vec{v}| \ll 1\). The equations of motion will therefore be

\[
    m \frac{d\vec{v}}{dt} = e(\vec{E} + \vec{v} \times \vec{B}),
\]

(3.42)

and so

\[
    m\ddot{x} = eB\dot{y}, \quad m\ddot{y} = eE_y - eB\dot{x}, \quad m\ddot{z} = eE_z.
\]

(3.43)

We can immediately solve for \( z \), finding

\[
    z = \frac{e}{2m} E_z t^2 + \tilde{v}_t t, \quad (3.44)
\]

where we have chosen the \( z \) origin so that \( z = 0 \) at \( t = 0 \). The \( x \) and \( y \) equations can be combined into

\[
    \frac{d}{dt}(\dot{x} + i\dot{y}) + i\omega(\dot{x} + i\dot{y}) = \frac{ie}{m} E_y,
\]

(3.45)
where $\omega = eB/m$. Thus we find
\begin{equation}
\dot{x} + i \dot{y} = ae^{-i\omega t} + \frac{e}{m\omega} E_y = ae^{-i\omega t} + \frac{E_y}{B}.
\end{equation}

Choosing the origin of time so that $a$ is real, we have
\begin{equation}
\dot{x} = a \cos \omega t + \frac{E_y}{B}, \quad \dot{y} = -a \sin \omega t.
\end{equation}

Taking the time averages, we see that
\begin{equation}
\langle \dot{x} \rangle = \frac{E_y}{B}, \quad \langle \dot{y} \rangle = 0.
\end{equation}

The averaged velocity along the $x$ direction is called the drift velocity. Notice that it is perpendicular to $\vec{E}$ and $\vec{B}$. It can be written in general as
\begin{equation}
\vec{v}_{\text{drift}} = \frac{\vec{E} \times \vec{B}}{B^2}.
\end{equation}

For our assumption that $|\vec{v}| << 1$ to be valid, we must have $|\vec{E} \times \vec{B}| << B^2$, i.e. $|E_y| << |B|$.

Integrating (3.47) once more, we find
\begin{equation}
x = \frac{a}{\omega} \sin \omega t + \frac{E_y}{B} t, \quad y = \frac{a}{\omega} (\cos \omega t - 1),
\end{equation}
where the origins of $x$ and $y$ have been chosen so that $x = y = 0$ at $t = 0$. These equations describe the projection of the particle’s motion onto the $(x, y)$ plane. The curve is called a trochoid. If $|a| > E_y/B$ there will be loops in the motion, and in the special case $a = -E_y/B$ the curve becomes a cycloid, with cusps:
\begin{equation}
x = \frac{E_y}{\omega B} (\omega t - \sin \omega t), \quad y = \frac{E_y}{\omega B} (1 - \cos \omega t).
\end{equation}

4 Action Principle for Electrodynamics

In this section, we shall show how the Maxwell equations themselves can be derived from an action principle. We shall also introduce the notion of the energy-momentum tensor for the electromagnetic field. We begin with a discussion of Lorentz invariant quantities that can be built from the Maxwell field strength tensor $F_{\mu\nu}$.

4.1 Invariants of the electromagnetic field

As we shall now show, it is possible to build two independent Lorentz invariants that are quadratic in the electromagnetic field. One of these will turn out to be just what is needed in order to construct an action for electrodynamics.
4.1.1 The first invariant

The first quadratic invariant is very simple; we may write

\[ I_1 \equiv F_{\mu \nu} F^{\mu \nu}. \]  

(4.1)

Obviously this is Lorentz invariant, since it is built from the product of two Lorentz tensors, with all indices contracted. It is instructive to see what this looks like in terms of the electric and magnetic fields. From the expressions given in (2.14), we see that

\[ I_1 = F_{0i} F^{0i} + F_{ij} F^{ij}, \]

\[ = 2F_{0i} F^{0i} + F_{ij} F^{ij} = -2E_i E_i + \epsilon_{ijk} B_k \epsilon_{ij\ell} B_\ell, \]

\[ = -2E_i E_i + 2B_i B_i, \]

(4.2)

and so

\[ I_1 = F_{\mu \nu} F^{\mu \nu} = 2(\vec{B}^2 - \vec{E}^2). \]

(4.3)

One could, of course, verify from the Lorentz transformations (2.56) and (2.57) for \( \vec{E} \) and \( \vec{B} \) that indeed \((\vec{B}^2 - \vec{E}^2)\) was invariant, i.e. \( I_1' = I_1 \) under Lorentz transformations. This would be quite an involved computation. However, the great beauty of the 4-dimensional language is that there is absolutely no work needed at all; one can see by inspection that \( F_{\mu \nu} F^{\mu \nu} \) is Lorentz invariant.

4.1.2 The second invariant

The second quadratic invariant that we can write down is given by

\[ I_2 \equiv \frac{1}{2} \epsilon^{\mu \nu \rho \sigma} F_{\mu \nu} F_{\rho \sigma}. \]

(4.4)

First, we need to explain the tensor \( \epsilon^{\mu \nu \rho \sigma} \). This is the four-dimensional Minkowski spacetime generalisation of the totally-antisymmetric tensor \( \epsilon^{ijk} \) of three-dimensional Cartesian tensor analysis. The tensor \( \epsilon^{\mu \nu \rho \sigma} \) is also totally antisymmetric in all its indices. That means that it changes sign if any two indices are exchanged. For example,\(^{14}\)

\[ \epsilon^{\mu \nu \rho \sigma} = -\epsilon^{\nu \mu \rho \sigma} = -\epsilon^{\mu \nu \sigma \rho} = -\epsilon^{\rho \mu \nu \sigma}. \]

(4.5)

\(^{14}\)Beware that in an odd dimension, such as 3, the process of “cycling” the indices on \( \epsilon_{ijk} \) (for example, pushing one off the right-hand end and bringing it to the front) is an even permutation; \( \epsilon_{kij} = \epsilon_{ijk} \). By contrast, in an even dimension, such as 4, the process of cycling is an odd permutation; \( \epsilon^{\sigma \mu \nu \rho} = -\epsilon^{\mu \rho \sigma \nu}. \) This is an elementary point, but easily overlooked if one is familiar only with three dimensions!
Since all the non-vanishing components of $\epsilon^{\mu\nu\rho\sigma}$ are related by the antisymmetry, we need only specify one non-vanishing component in order to define the tensor completely. We shall define

$$\epsilon^{0123} = -1,$$

or, equivalently

$$\epsilon^{0123} = +1.$$  \hspace{1cm} (4.6)

Thus $\epsilon^{\mu\nu\rho\sigma}$ is $-1$, $+1$ or $0$ according to whether $(\mu\nu\rho\sigma)$ is an even permutation of $(0123)$, and odd permutation, or no permutation at all. We use this definition of $\epsilon^{\mu\nu\rho\sigma}$ in all frames. This can be done because, like the Minkowski metric $\eta_{\mu\nu}$, the tensor $\epsilon^{\mu\nu\rho\sigma}$ is an invariant tensor, as we shall now discuss.

Actually, to be more precise, $\epsilon^{\mu\nu\rho\sigma}$ is an invariant pseudo-tensor. This means that under Lorentz transformations that are connected to the identity (pure boosts and/or pure rotations), it is truly an invariant tensor. However, it reverses its sign under Lorentz transformations that involve a reflection. To see this, let us calculate what the transformation of $\epsilon^{\mu\nu\rho\sigma}$ would be if we assume it behaves as an ordinary Lorentz tensor:

$$\epsilon'^{\mu\nu\rho\sigma} \equiv \Lambda^\mu_\alpha \Lambda^\nu_\beta \Lambda^\rho_\gamma \Lambda^\sigma_\delta \epsilon^\alpha\beta\gamma\delta,$$

$$= (\det \Lambda) \epsilon^{\mu\nu\rho\sigma}.$$  \hspace{1cm} (4.7)

The last equality can easily be seen by writing out all the terms. (It is easier to play around with the analogous identity in 2 or 3 dimensions, to convince oneself of it in an example with fewer terms to write down.) Now, we already saw in section 2.3 that $\det \Lambda = \pm 1$, with $\det \Lambda = +1$ for pure boosts and/or rotations, and $\det \Lambda = -1$ if there is a reflection as well. (See the discussion leading up to equation (2.40).) Thus we see from (4.7) that $\epsilon^{\mu\nu\rho\sigma}$ behaves like an invariant tensor, taking the same values in all Lorentz frames, provided there is no reflection. (Lorentz transformations connected to the identity, i.e. where there is no reflection, are sometimes called proper Lorentz transformations.) In practice, we shall almost always be considering only proper Lorentz transformations, and so the distinction between a tensor and a pseudo-tensor will not concern us.

Returning now to the second quadratic invariant, (4.4), we shall have

$$I_2 = \frac{1}{2} \epsilon^{\mu\nu\rho\sigma} F_{\mu\nu} F_{\rho\sigma} = \frac{1}{2} \times 4 \times \epsilon^{0ijk} F_{0i} F_{jk},$$

$$= 2(-\epsilon_{ijk})(-E_i)\epsilon_{jkl} B_\ell,$$

$$= 4 E_i B_i = 4 \vec{E} \cdot \vec{B}.$$  \hspace{1cm} (4.8)

Thus, to summarise, we have the two quadratic invariants

$$I_1 = F_{\mu\nu} F^{\mu\nu} = 2(\vec{B}^2 - \vec{E}^2),$$

$$I_2 = \frac{1}{2} \epsilon^{\mu\nu\rho\sigma} F_{\mu\nu} F_{\rho\sigma} = 4 \vec{E} \cdot \vec{B}.$$  \hspace{1cm} (4.9)
Since the two quantities $I_1$ and $I_2$ are (manifestly) Lorentz invariant, this means that, even though it is not directly evident in the three-dimensional language without quite a lot of work, the two quantities

$$\vec{B}^2 - \vec{E}^2, \quad \text{and} \quad \vec{E} \cdot \vec{B}$$

(4.10)

are Lorentz invariant; i.e. they take the same values in all Lorentz frames. This has a number of consequences. For example

1. If $\vec{E}$ and $\vec{B}$ are perpendicular in one Lorentz frame, then they are perpendicular in all Lorentz frames.

2. In particular, if there exists a Lorentz frame where the electromagnetic field is purely electric ($\vec{B} = 0$), or purely magnetic ($\vec{E} = 0$), then $\vec{E}$ and $\vec{B}$ are perpendicular in any other frame.

3. If $|\vec{E}| > |\vec{B}|$ in one frame, then it is true in all frames. Conversely, if $|\vec{E}| < |\vec{B}|$ in one frame, then it is true in all frames.

4. By making an appropriate Lorentz transformation, we can, at a given point, make $\vec{E}$ and $\vec{B}$ equal to any values we like, subject only to the conditions that we cannot alter the values of $(\vec{B}^2 - \vec{E}^2)$ and $\vec{E} \cdot \vec{B}$ at that point.

### 4.2 Action for electrodynamics

We have already discussed the action principle for a charged particle moving in an electromagnetic field. In that discussion, the electromagnetic field was just a specified background, which, of course, would be a solution of the Maxwell equations. We can also derive the Maxwell equations themselves from an action principle, as we shall now show.

We begin by introducing the notion of Lagrangian density. This is a quantity that is integrated over a three-dimensional spatial volume (typically, all of 3-space) to give the Lagrangian:

$$L = \int \mathcal{L} d^3x$$

(4.11)

Then, the Lagrangian is integrated over a time interval $t_1 \leq t \leq t_2$ to give the action,

$$S = \int_{t_1}^{t_2} L dt = \int \mathcal{L} dt$$

(4.12)

Consider first the vacuum Maxwell equations without sources,

$$\partial_\mu F^{\mu\nu} = 0, \quad \partial_\mu F_{\nu\rho} + \partial_\nu F_{\rho\mu} + \partial_\rho F_{\mu\nu} = 0.$$  

(4.13)