Willoughby Seago

Theoretical Physics

Classical Electrodynamics

May 18, 2022

COURSE NOTES

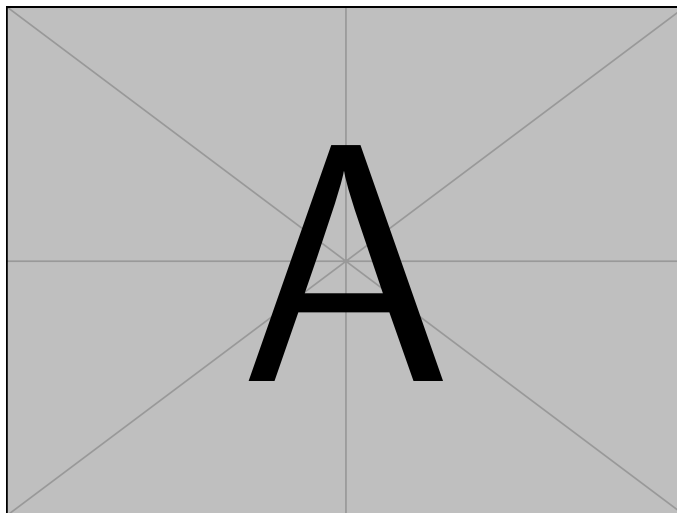# Classical Electrodynamics

Willoughby Seago

May 18, 2022

These are my notes from the course classical electrodynamics. I took this course as a part of the theoretical physics degree at the University of Edinburgh.

These notes were last updated at 17:19 on August 24, 2022. For notes on other topics see .

Chapters

Page

Contents

Page

List of Figures

Page

One

Introduction

Content

Classical electrodynamics is the study of relativistic electromagnetism. The classical refers to the fact that this course does *not* cover quantum electrodynamics, QED. However, the content covered in this course should leave a student well prepared to start studying QED and other quantum field theories.

Broadly there are four sections to this course, we start with the fundamentals, reviewing non-relativistic electrodynamics, and special relativity, before combining them in the next section, relativistic electrodynamics. We then study action principles, an incredibly important area of study if going on to do any field theory. We then have a short section on Green's functions, before turning to our final topic, radiation.

This course assumes prior knowledge of the following subjects:

- Maxwell's equations. ($*$)

- The scalar and vector potentials φ and $\boldsymbol{A}$. ($* \, \S$)

- Conservation laws, such as conservation of charge and energy. ($* \, \S$)

- The Poynting vector, which is proportional to $\boldsymbol{E} \times \boldsymbol{B}$. ($*$)

- Gauge transformations, such as $\boldsymbol{A} \to \boldsymbol{A} + \nabla \chi$ and $\varphi \to \varphi - \partial_t \chi$. ($* \, \P \, \|$)

- Relativistic notation, such as $x^\mu = (ct, \boldsymbol{x})$, $x_\mu = (ct, -\boldsymbol{x})$, $\partial_\mu = \left(\frac{1}{c} \partial_t, \nabla \right)$, and $\partial^\mu = \left(\frac{1}{c} \partial_t, -\nabla \right)$. ($\dagger \, \P \, \|$)

- Proper time, τ. ($\dagger \, \P$)

- Lagrangian dynamics. ($\S$)

- Green's functions. ($\# \, ⁇$)

- Vector calculus. ($\neg$)

- Tensors and index notation. ($\neg \, \text{¡} \, \P$)

These topics are covered in the following courses, all of which have notes which can be found at https://github.com/WilloughbySeago/Uni-Notes:

- $*$ Electromagnetism.

† Relativity, Nuclear, and Particle Physics (in particular the relativity bit).

§ Lagrangian Dynamics.

¶ General Relativity.

‖ Quantum Theory.

\# Fourier Analysis and Statistics (in particular the Fourier analysis bit).

⁇ Methods of Mathematical Physics.

¬ Dynamics and Vector Calculus (in particular the vector calculus part).

¡ Methods of Theoretical Physics.

1.2 Conventions and Units

For most of the course we work in Heaviside–Lorentz units, which we will introduce in the first section. These are the same units as used in the Quantum Theory course. Equations *not* in Heaviside–Lorentz units will be flagged with a label giving the appropriate units, so for example an equation in SI units will have an accompanying $\boxed{\text{SI}}$, or an equation in Gaussian units will have $\boxed{\text{Gaussian}}$.

We follow the particle physics convention of using the metric signature $(+---)$, meaning the Minkowski metric is given by $\eta_{\mu\nu} = \eta^{\mu\nu} = \mathrm{diag}(+1, -1, -1, -1)$. This matches all other courses on relativity, including Relativity Nuclear and Particle Physics, Quantum Theory, and General Relativity.

Part I

Fundamentals and Review

Two

Electrodynamics

2.1 Maxwell's Equations

Recall that in SI units **Maxwell's equations** are

$$\nabla \cdot \boldsymbol{E} = \frac{\rho}{\varepsilon_0}, \qquad \boxed{\text{SI}} \qquad (2.1.1)$$

$$\nabla \cdot \boldsymbol{B} = 0, \qquad \boxed{\text{SI}} \qquad (2.1.2)$$

$$\nabla \times \boldsymbol{E} = -\frac{\partial \boldsymbol{B}}{\partial t}, \qquad \boxed{\text{SI}} \qquad (2.1.3)$$

$$\nabla \times \boldsymbol{B} = \mu_0 \boldsymbol{J} + \varepsilon_0 \mu_0 \frac{\partial \boldsymbol{E}}{\partial t}. \qquad \boxed{\text{SI}} \qquad (2.1.4)$$

The first equation is called Gauss' law, the third Faraday's law and the fourth either Ampère's law or the Ampère–Maxwell law. Notice the presence of the two constants $\varepsilon_0 = 8.854 \times 10^{-12}\,\mathrm{F\,m^{-1}}$ and $\mu_0 = 1.257 \times 10^{-6}\,\mathrm{H\,m^{-1}}$.

Apart from when explicitly stated otherwise the Einstein summation convention is in place. This means that we sum over repeated indices. We follow the convention of Latin indices, such as i and j, running over 1, 2, and 3, and Greek indices, such as μ and ν, running over 0, 1, 2, and 3.

From these, using various vector calculus identities, such as the fact that $\nabla \times (\nabla \times \boldsymbol{V}) = \nabla(\nabla \cdot \boldsymbol{V}) - \nabla^2 \boldsymbol{V}$, we can derive the wave equations

$$\varepsilon_0 \mu_0 \frac{\partial^2 \boldsymbol{E}}{\partial t^2} - \nabla^2 \boldsymbol{E} = 0, \qquad \text{and} \qquad \varepsilon_0 \mu_0 \frac{\partial^2 \boldsymbol{B}}{\partial t^2} - \nabla^2 \boldsymbol{B} = 0. \qquad (2.1.5)$$

Identifying electromagnetic waves as light we see that the speed of light is given by

$$c = \frac{1}{\sqrt{\varepsilon_0 \mu_0}}. \qquad (2.1.6)$$

In order to see the effect of the electric and magnetic fields on matter we need the **Lorentz force law**:

$$\boldsymbol{f} = \rho \boldsymbol{E} + \boldsymbol{J} \times \boldsymbol{B}. \qquad \boxed{\text{SI}} \qquad (2.1.7)$$

This gives the force density, $\boldsymbol{f}$, which is the force per unit mass, on a charge density, ρ, and current density, $\boldsymbol{J}$. In order to use this we need to know how matter reacts

to a force. For now we can simply consider the non-relativistic response given by Newton's second law

$$\boldsymbol{F} = m\boldsymbol{a}, \tag{2.1.8}$$

where

$$\boldsymbol{F} = \int_V \boldsymbol{f}\, \mathrm{d}^3 x. \tag{2.1.9}$$

2.1.1 Lorentz–Heaviside Units

Notice that the electric and magnetic fields have different dimensions in SI units. The can be seen from, for example Faraday's law, which shows that

$$[\nabla \times \boldsymbol{E}] = \frac{[\boldsymbol{E}]}{L} = \left[\frac{\partial \boldsymbol{B}}{\partial t}\right] = \frac{[\boldsymbol{B}]}{T}. \tag{2.1.10}$$

Rearranging this we have that

$$[\boldsymbol{B}] = [\boldsymbol{E}]L^{-1}T. \tag{2.1.11}$$

In order to combine the electric field and magnetic field into a single object to allow us to study their connection it is best to use a unit system where the electric and magnetic field have the same dimensions. This can be done by slightly redefining the fields and sources to absorb some factors of ε_0 and μ_0:

$$\boldsymbol{E} \to \frac{1}{\sqrt{\varepsilon_0}}\boldsymbol{E}, \qquad \boldsymbol{B} \to \sqrt{\mu_0}\boldsymbol{B}, \qquad \rho \to \sqrt{\varepsilon_0}\rho, \qquad \text{and} \qquad \boldsymbol{J} \to \sqrt{\varepsilon_0}\boldsymbol{J}. \tag{2.1.12}$$

This defines a system of units called **Lorentz–Heaviside units**. Alternatively to scaling the charge and current densities we could scale the charge, $q \to \sqrt{\varepsilon_0}q$, which then carries on through to the charge and current densities.

Maxwell's equations take a slightly different form in these units. To see this we take the SI versions and make the above changes:

$$\nabla \cdot \boldsymbol{E} = \frac{\rho}{\varepsilon_0} \to \nabla \cdot \frac{1}{\sqrt{\varepsilon_0}}\boldsymbol{E} = \frac{\sqrt{\varepsilon_0}\rho}{\varepsilon_0} \implies \nabla \cdot \boldsymbol{E} = \rho, \tag{2.1.13}$$

$$\nabla \cdot \boldsymbol{B} = 0 \to \nabla \cdot \sqrt{\mu_0}\boldsymbol{B} \implies \nabla \cdot \boldsymbol{B} = 0, \tag{2.1.14}$$

$$\nabla \times \boldsymbol{E} = -\frac{\partial \boldsymbol{B}}{\partial t} \to \nabla \times \frac{1}{\sqrt{\varepsilon_0}}\boldsymbol{E} = -\frac{\partial}{\partial t}\sqrt{\mu_0}\boldsymbol{B}$$

$$\implies \nabla \times \boldsymbol{E} = -\sqrt{\varepsilon_0\mu_0}\frac{\partial \boldsymbol{B}}{\partial t} = -\frac{1}{c}\frac{\partial \boldsymbol{B}}{\partial t}, \tag{2.1.15}$$

$$\nabla \times \boldsymbol{B} = \mu_0\boldsymbol{J} + \mu_0\varepsilon_0\frac{\partial \boldsymbol{E}}{\partial t} \to \nabla \times \sqrt{\mu_0}\boldsymbol{B} = \mu_0\sqrt{\varepsilon_0}\boldsymbol{J} + \varepsilon_0\mu_0\frac{\partial}{\partial t}\sqrt{\varepsilon_0}\boldsymbol{E}$$

$$\implies \nabla \times \boldsymbol{B} = \sqrt{\mu_0\varepsilon_0}\boldsymbol{J} + \sqrt{\mu_0\varepsilon_0}\frac{\partial \boldsymbol{E}}{\partial t} = \frac{1}{c}\boldsymbol{J} + \frac{1}{c}\frac{\partial \boldsymbol{E}}{\partial t}. \tag{2.1.16}$$

Another way to derive these is to start with the form of the SI equations but dropping all factors of ε_0 and μ_0 and inserting factors of c to enforce the requirement that $\boldsymbol{E}$ and $\boldsymbol{B}$ have the same units. For example, Faraday's law is

$$\nabla \times \boldsymbol{E} = -x\frac{\partial \boldsymbol{B}}{\partial t} \tag{2.1.17}$$

where x is some power of c to be determined. Looking at the units and requiring $[\boldsymbol{E}] = [\boldsymbol{B}]$ we have

$$[\boldsymbol{E}]L^{-1} = [x][\boldsymbol{B}]T^{-1} \implies [x] = L^{-1}T, \tag{2.1.18}$$

so we take $x = 1/c$ giving us the same result as before.

Either way we work it out we find that in Lorentz–Heaviside units Maxwell's equations are

$$\nabla \cdot \boldsymbol{E} = \rho, \tag{2.1.19}$$

$$\nabla \cdot \boldsymbol{B} = 0, \tag{2.1.20}$$

$$\nabla \times \boldsymbol{E} = -\frac{1}{c}\frac{\partial \boldsymbol{B}}{\partial t}, \tag{2.1.21}$$

$$\nabla \times \boldsymbol{B} = \frac{1}{c}\boldsymbol{J} + \frac{1}{c}\frac{\partial \boldsymbol{E}}{\partial t}. \tag{2.1.22}$$

The Lorentz force law also changes in these units. First notice that the factor of ε_0 in $\rho \to \sqrt{\varepsilon_0}$ and $\boldsymbol{E} \to \boldsymbol{E}/\sqrt{\varepsilon_0}$ cancel, so only the second term changes, picking up a factor of $\sqrt{\varepsilon_0}$ from $\boldsymbol{J}$ and $\sqrt{\mu_0}$ from $\boldsymbol{B}$ giving

$$\boldsymbol{f} = \rho\boldsymbol{E} + \frac{1}{c}\boldsymbol{J} \times \boldsymbol{B}. \tag{2.1.23}$$

Again, we can check that $\boldsymbol{E}$ and $\boldsymbol{B}$ have the same dimensions by using the fact that for a charge of fixed shape moving at velocity $\boldsymbol{v}$ we have $\boldsymbol{J} = \rho\boldsymbol{v}$, and so $\boldsymbol{J}/c$ and ρ have the same units, meaning $\boldsymbol{E}$ and $\boldsymbol{B}$ must also have the same units.

2.1.1.1 Gaussian Units

There is a third set of units also in common use in electromagnetism, known as **Gaussian units**. These are related to Lorentz–Heaviside units by introducing a factor of 4π to the charge: $q \to 4\pi q$. This introduces factors of 4π in Maxwell's equations but removes them from solutions. For this reason these units are often preferred for practical use with electromagnetism where the solutions are of interest, rather than in theoretical use where the equations themselves are the interesting thing. In Gaussian units Maxwell's equations are

$$\nabla \cdot \boldsymbol{E} = 4\pi\rho, \qquad \boxed{\text{Gaussian}} \tag{2.1.24}$$

$$\nabla \cdot \boldsymbol{B} = 0, \qquad \boxed{\text{Gaussian}} \tag{2.1.25}$$

$$\nabla \times \boldsymbol{E} = -\frac{1}{c}\frac{\partial \boldsymbol{B}}{\partial t}, \qquad \boxed{\text{Gaussian}} \tag{2.1.26}$$

$$\nabla \times \boldsymbol{B} = \frac{4\pi}{c}\boldsymbol{J} + \frac{1}{c}\frac{\partial \boldsymbol{E}}{\partial t}. \qquad \boxed{\text{Gaussian}} \tag{2.1.27}$$

From now on we will work almost exclusively with Lorentz–Heaviside units.

2.2 Conservation Laws

2.2.1 Current Conservation

> **Lemma 2.2.1** For a smooth vector field $V \colon \mathbb{R}^3 \to \mathbb{R}^3$ we have
>
> $$\nabla \cdot (\nabla \times V) = 0. \tag{2.2.2}$$
>
> ---
>
> *Proof.* In index notation $(\nabla \times V)_i = \varepsilon_{ijk} \partial_j V_k$. Hence,
>
> $$\nabla \cdot (\nabla \times V) = \partial_i \varepsilon_{ijk} \partial_j V_k. \tag{2.2.3}$$
>
> Notice that $\partial_i \partial_j$ is symmetric in i and j whereas ε_{ijk} is antisymmetric so this vanishes. $\qquad\square$

Using this identity we can take any equation for the curl of a field and reduce it to zero. Doing so for Faraday's law we have

$$0 = \nabla \cdot (\nabla \times E) = -\frac{1}{c} \nabla \cdot \frac{\partial B}{\partial t} = -\frac{1}{c} \frac{\partial}{\partial t} (\nabla \cdot B) = 0. \tag{2.2.4}$$

This is consistent, but doesn't give us any new information.

To get new information we have to apply this identity to Ampère's law:

$$0 = \nabla \cdot (\nabla \times E) = \frac{1}{c} \nabla \cdot J + \frac{1}{c} \frac{\partial}{\partial t} (\nabla \cdot E) = \frac{1}{c} \nabla \cdot J + \frac{1}{c} \frac{\partial \rho}{\partial t}. \tag{2.2.5}$$

We can identify this as the **continuity equation**. What it says is that not only is charge conserved globally, meaning that the total charge is constant, but it is conserved locally, meaning that to change the charge at a point we have to have a current of charge flow into or out of that point. That is, if ρ is to decrease at a point, so $\partial_t \rho < 0$, we must have that $\nabla \cdot J > 0$, which is to say there is a current coming out of that point.

2.2.2 Conservation of Energy

The equation for conservation of energy is slightly more complicated because we have to consider different forms of energy. We will consider energy densities, which can then be integrated over a volume to get the energy. Start by considering the mechanical work, w, done by the Lorentz force. For an infinitesimal displacement, dr, the work done is

$$dw = f \cdot dr. \tag{2.2.6}$$

Using $v = dr/dt$ we have $dr = v\,dt$. Hence, $dw = f \cdot v\,dt$. We now write the current density as $J = \rho v$, from which it is clear that $J \times B$ is perpendicular to the velocity, and so its dot product with the velocity vanishes:

$$dw = f \cdot v\,dt = \left(\rho E + \frac{1}{c} J \times B \right) \cdot v\,dt = \rho E \cdot v\,dt = E \cdot J\,dt. \tag{2.2.7}$$

From this we can identify the rate at which work is done:

$$\frac{\partial w}{\partial t} = E \cdot J. \tag{2.2.8}$$

The problem with this is that it involves the current density, but we want a statement about the energy stored in the fields. We can use Ampère's law to eliminate J:

$$J = c\nabla \times B - \frac{\partial E}{\partial t}. \tag{2.2.9}$$

We then have

$$\boldsymbol{E} \cdot \boldsymbol{J} = c\boldsymbol{E} \cdot (\nabla \times \boldsymbol{B}) - \boldsymbol{E} \cdot \frac{\partial \boldsymbol{E}}{\partial t}. \tag{2.2.10}$$

We now use the product rule to identify

$$\frac{\partial}{\partial t}\boldsymbol{E}^2 = 2\boldsymbol{E} \cdot \frac{\partial \boldsymbol{E}}{\partial t} \tag{2.2.11}$$

and so

$$\boldsymbol{E} \cdot \boldsymbol{J} = c\boldsymbol{E} \cdot (\nabla \times \boldsymbol{B}) - \frac{1}{2}\frac{\partial}{\partial t}\boldsymbol{E}^2. \tag{2.2.12}$$

> **Lemma 2.2.13** For smooth vector fields $\boldsymbol{a}, \boldsymbol{b} : \mathbb{R}^3 \to \mathbb{R}^3$ we have
>
> $$\nabla \cdot (\boldsymbol{a} \times \boldsymbol{b}) = \boldsymbol{a} \cdot (\nabla \times \boldsymbol{b}) - \boldsymbol{b} \cdot (\nabla \times \boldsymbol{a}). \tag{2.2.14}$$
>
> ---
>
> *Proof.* Consider the left hand side in index notation:
>
> $$\begin{aligned}
\nabla \cdot (\boldsymbol{a} \times \boldsymbol{b}) &= \partial_i(\boldsymbol{a} \times \boldsymbol{b})_i & \text{(2.2.15)} \\
&= \partial_i \varepsilon_{ijk} a_j b_k & \text{(2.2.16)} \\
&= \varepsilon_{ijk} b_k \partial_i a_j + \varepsilon_{ijk} a_j \partial_i b_k & \text{(2.2.17)} \\
&= b_k \varepsilon_{kij} \partial_i a_j - a_j \varepsilon_{jik} \partial_i b_k & \text{(2.2.18)} \\
&= b_k (\nabla \times \boldsymbol{a})_k - a_k (\nabla \times \boldsymbol{b})_j & \text{(2.2.19)} \\
&= \boldsymbol{b} \cdot (\nabla \times \boldsymbol{a}) - \boldsymbol{a} \cdot (\nabla \times \boldsymbol{b}). & \square
\end{aligned}$$

Using this identity we have

$$\nabla\cdot(\boldsymbol{E}\times\boldsymbol{B}) = \boldsymbol{B}\cdot(\nabla\times\boldsymbol{E})-\boldsymbol{E}\cdot(\nabla\times\boldsymbol{B}) \implies c\boldsymbol{E}\cdot(\nabla\times\boldsymbol{B}) = -c\nabla\cdot(\boldsymbol{E}\times\boldsymbol{B})-c\boldsymbol{B}\cdot(\nabla\times\boldsymbol{E}). \tag{2.2.20}$$

Using Faraday's law we can replace $\nabla \times \boldsymbol{E}$ to get

$$c\boldsymbol{E} \cdot (\nabla \times \boldsymbol{B}) = -c\nabla \cdot (\boldsymbol{E} \times \boldsymbol{B}) - \boldsymbol{B} \cdot \frac{\partial \boldsymbol{B}}{\partial t}. \tag{2.2.21}$$

Again, using the product rule trick we can write the last term as half the time derivative of $\boldsymbol{B}^2$ giving

$$c\boldsymbol{E} \cdot (\nabla \times \boldsymbol{B}) = -c\nabla \cdot (\boldsymbol{E} \times \boldsymbol{B}) - \frac{1}{2}\frac{\partial}{\partial t}\boldsymbol{B}^2. \tag{2.2.22}$$

Hence, we have

$$\boldsymbol{E} \cdot \boldsymbol{J} = -c\nabla \cdot (\boldsymbol{E} \times \boldsymbol{B}) - \frac{1}{2}\frac{\partial}{\partial t}(\boldsymbol{E}^2 + \boldsymbol{B}^2). \tag{2.2.23}$$

We now define two quantities, first, the **electromagnetic energy density**,

$$u := \frac{1}{2}(\boldsymbol{E}^2 + \boldsymbol{B}^2), \tag{2.2.24}$$

$$u := \frac{1}{2}\left(\varepsilon_0 \boldsymbol{E}^2 + \frac{1}{\mu_0}\boldsymbol{B}^2\right). \qquad \boxed{\text{SI}} \tag{2.2.25}$$

Second, the **Poynting vector**,

$$S := c\boldsymbol{E} \times \boldsymbol{B}, \tag{2.2.26}$$

$$S := \frac{1}{\mu_0} \boldsymbol{E} \times \boldsymbol{B}. \qquad \boxed{\text{SI}} \tag{2.2.27}$$

We then have that the power density is

$$\frac{\partial w}{\partial t} = \frac{\partial u}{\partial t} + \nabla \cdot \boldsymbol{S}. \tag{2.2.28}$$

For the case where an electromagnetic field does no work we can identify this as the continuity equation

$$\frac{\partial u}{\partial t} + \nabla \cdot \boldsymbol{S} = 0. \tag{2.2.29}$$

This means that the only way for the energy stored in a field at a point to change is for energy to flow in or out, which allows us to identify $\boldsymbol{S}$ as the energy flux.

For the case where an electromagnetic field does do work we get **Poynting's theorem**

$$\frac{\partial}{\partial t}(u + w) + \nabla \cdot \boldsymbol{S} = 0. \tag{2.2.30}$$

This says that for an electromagnetic process to increase the mechanical energy, by doing work, there must be a corresponding decrease in the electromagnetic energy, and vice versa.

2.2.3 Conservation of Momentum

Lemma 2.2.31 Given a smooth vector field, $\boldsymbol{V} \colon \mathbb{R}^3 \to \mathbb{R}^3$ we have

$$[(\nabla \cdot \boldsymbol{V})\boldsymbol{V} - \boldsymbol{V} \times (\nabla \times \boldsymbol{V})]_i = \partial_j\left(V_i V_j - \frac{1}{2}V^2\right). \tag{2.2.32}$$

Proof. We start by writing the left hand side in index notation:

$$[(\nabla \cdot \boldsymbol{V})\boldsymbol{V} - \boldsymbol{V} \times (\nabla \times \boldsymbol{V})]_i = (\partial_j V_j)V_i - \varepsilon_{ijk}V_j(\nabla \times \boldsymbol{V})_k \tag{2.2.33}$$

$$= (\partial_j V_j)V_i - \varepsilon_{ijk}\varepsilon_{klm}V_j\partial_l V_m \tag{2.2.34}$$

$$= V_i\partial_j V_j - (\delta_{il}\delta_{jm} - \delta_{im}\delta_{jl})V_j\partial_l V_m \tag{2.2.35}$$

$$= V_i\partial_j V_j - V_j\partial_i V_j + V_j\partial_j V_i. \tag{2.2.36}$$

Now notice that

$$\partial_j(V_i V_j) = V_i\partial_j V_j + V_j\partial_j V_i \tag{2.2.37}$$

and

$$\partial_i V^2 = \partial_i(V_j V_j) = 2V_j\partial_i V_j. \tag{2.2.38}$$

We can change the index on this partial derivative using δ_{ij}:

$$\partial_i V^2 = \delta_{ij}\partial_j V^2. \tag{2.2.39}$$

Hence, we have

$$[(\nabla \cdot \boldsymbol{V})\boldsymbol{V} - \boldsymbol{V} \times (\nabla \times \boldsymbol{V})]_i = \partial_j \left(V_i V_j - \frac{1}{2}\delta_{ij}V^2 \right). \tag{2.2.40}$$

$\square$

Consider the Lorentz force law

$$\boldsymbol{f} = \rho\boldsymbol{E} + \frac{1}{c}\boldsymbol{J} \times \boldsymbol{B}. \tag{2.2.41}$$

We can use Maxwell's equations to eliminate the sources. First, Gauss' law gives $\rho = \nabla \cdot \boldsymbol{E}$, and Ampère's law gives

$$\frac{1}{c}\boldsymbol{J} = \nabla \times \boldsymbol{B} - \frac{1}{c}\frac{\partial \boldsymbol{E}}{\partial t}. \tag{2.2.42}$$

Inserting this the Lorentz force law becomes

$$\boldsymbol{f} = (\nabla \cdot \boldsymbol{E})\boldsymbol{E} + \left[\nabla \times \boldsymbol{B} - \frac{1}{c}\frac{\partial \boldsymbol{E}}{\partial t} \right] \times \boldsymbol{B}. \tag{2.2.43}$$

Now consider the product rule applied to $\boldsymbol{E} \times \boldsymbol{B}$:

$$\frac{\partial}{\partial t}(\boldsymbol{E} \times \boldsymbol{B}) = \frac{\partial \boldsymbol{E}}{\partial t} \times \boldsymbol{B} + \boldsymbol{E} \times \frac{\partial \boldsymbol{B}}{\partial t} \implies -\frac{1}{c}\frac{\partial \boldsymbol{E}}{\partial t} \times \boldsymbol{B} = \frac{1}{c}\boldsymbol{E} \times \frac{\partial \boldsymbol{B}}{\partial t} - \frac{\partial}{\partial t}\left(\frac{1}{c}\boldsymbol{E} \times \boldsymbol{B} \right). \tag{2.2.44}$$

We can further eliminate the time derivative of $\boldsymbol{B}$ using Faraday's law:

$$\frac{\partial \boldsymbol{B}}{\partial t} = -c\nabla \times \boldsymbol{E}. \tag{2.2.45}$$

Hence

$$-\frac{1}{c}\frac{\partial \boldsymbol{E}}{\partial t} \times \boldsymbol{B} = -\boldsymbol{E} \times (\nabla \times \boldsymbol{E}) - \frac{\partial}{\partial t}\left(\frac{1}{c}\boldsymbol{E} \times \boldsymbol{B} \right). \tag{2.2.46}$$

Putting this into the Lorentz force law, and using the fact that the cross product is antisymmetric we get

$$\boldsymbol{f} = (\nabla \cdot \boldsymbol{E})\boldsymbol{E} - \boldsymbol{E} \times (\nabla \times \boldsymbol{E}) - \boldsymbol{B} \times (\nabla \times \boldsymbol{B}) - \frac{\partial}{\partial t}\left(\frac{1}{c}\boldsymbol{E} \times \boldsymbol{B} \right). \tag{2.2.47}$$

Finally we can use the fact that $\nabla \cdot \boldsymbol{B} = 0$ to insert a term proportional to $\boldsymbol{B}$ to make the equation more symmetric in $\boldsymbol{E}$ and $\boldsymbol{B}$. We also identify $\boldsymbol{E} \times \boldsymbol{B} = \boldsymbol{S}/c$ giving

$$\boldsymbol{f} = [(\nabla \cdot \boldsymbol{E})\boldsymbol{E} - \boldsymbol{E} \times (\nabla \times \boldsymbol{E})] + [(\nabla \cdot \boldsymbol{B})\boldsymbol{B} - \boldsymbol{B} \times (\nabla \times \boldsymbol{B})] - \frac{1}{c^2}\frac{\partial \boldsymbol{S}}{\partial t}. \tag{2.2.48}$$

We now use Newton's second law to identify that the rate of change of momentum is given by

$$\frac{\partial \boldsymbol{P}}{\partial t} = \boldsymbol{F} = \int_V \boldsymbol{f}\,\mathrm{d}^3 x. \tag{2.2.49}$$

Considering just a single component of this equation we have

$$\frac{\partial P_i}{\partial t} = \int_V f_i \, \mathrm{d}^3 x \tag{2.2.50}$$

$$= \int_V \Big([(\nabla \cdot \boldsymbol{E})\boldsymbol{E} - \boldsymbol{E} \times (\nabla \times \boldsymbol{E})]_i$$

$$+ [(\nabla \cdot \boldsymbol{B})\boldsymbol{B} - \boldsymbol{B} \times (\nabla \times \boldsymbol{B})]_i - \frac{1}{c^2} \frac{\partial S_i}{\partial t} \Big) \mathrm{d}^3 x. \tag{2.2.51}$$

We can now apply the identity from [Lemma 2.2.31](#) giving

$$\frac{\partial P_i}{\partial t} = \int_V \partial_j \left[\left(E_i E_j - \frac{1}{2} \delta_{ij} E^2 \right) + \left(B_i B_j - \frac{1}{2} \delta_{ij} B^2 \right) \right] \mathrm{d}^3 x - \frac{1}{c^2} \int_V \frac{\partial S_i}{\partial t} \, \mathrm{d}^3 x. \tag{2.2.52}$$

Now consider the divergence theorem:

$$\int_V \nabla \cdot \boldsymbol{U} \mathrm{d}^3 x = \int_A \boldsymbol{U} \cdot \mathrm{d}\boldsymbol{A} \iff \int_V \partial_j U_j \, \mathrm{d}^3 x = \int_A U_j \, \mathrm{d}A_j. \tag{2.2.53}$$

Here V is some closed volume with boundary A. Using this index form of the divergence theorem we get

$$\frac{\partial P_i}{\partial t} = \int_A \left[\left(E_i E_j - \frac{1}{2} \delta_{ij} \right) + \left(B_i B_j - \frac{1}{2} \delta_{ij} B^2 \right) \right] \mathrm{d}A_j - \frac{1}{c^2} \frac{\mathrm{d}}{\mathrm{d}t} \int_V S_i \, \mathrm{d}^3 x. \tag{2.2.54}$$

Now consider the continuity equation

$$0 = \nabla \cdot \boldsymbol{J} + \frac{\partial \rho}{\partial t}. \tag{2.2.55}$$

Integrating this over some volume, V, with boundary A we get

$$0 = \int_V \nabla \cdot \boldsymbol{J} \mathrm{d}^3 x + \int_V \frac{\partial \rho}{\partial t} \, \mathrm{d}^3 x \tag{2.2.56}$$

$$= \int_A \boldsymbol{J} \cdot \mathrm{d}\boldsymbol{A} + \int_V \frac{\partial \rho}{\partial t} \, \mathrm{d}^3 x \tag{2.2.57}$$

$$= \int_A J_j \cdot \mathrm{d}A_j + \int_V \frac{\partial \rho}{\partial t} \, \mathrm{d}^3 x. \tag{2.2.58}$$

Now compare this to [Equation (2.2.54)](#) and we can identify this as an integral form of a continuity equation. Notice, however, that there are two indices, i and j, and so we have a tensor equation. In particular we can identify the momentum flux tensor, T_{ij}, as

$$T_{ij} := - \left[E_i E_j + B_i B_j - \frac{1}{2} \delta_{ij} (E^2 + B^2) \right] \tag{2.2.59}$$

and the electromagnetic momentum density

$$p_i := \frac{1}{c^2} S_i. \tag{2.2.60}$$

The interpretation of T_{ij} is the flux of momentum along the i direction in the j direction. We then have the continuity equation

$$\partial_j T_{ij} + \frac{\partial p_i}{\partial t} = 0, \tag{2.2.61}$$

or in a more familiar form

$$\nabla \cdot \underline{\mathbf{T}} + \frac{\partial \boldsymbol{p}}{\partial t} = \mathbf{0} \tag{2.2.62}$$

where we use upright bold underlined for second rank tensors and define the divergence of such a tensor to be the vector with components $[\nabla \cdot \underline{\mathbf{T}}]_i := \partial_j T_i j$.

2.3 Electromagnetic Potentials

2.3.1 Static Case

We always have that

$$\nabla \cdot \boldsymbol{B} = 0. \tag{2.3.1}$$

It can be shown that if this is the case then there exists some **vector potential**, $\boldsymbol{A}$, such that

$$\boldsymbol{B} = \nabla \times \boldsymbol{A}. \tag{2.3.2}$$

Thus if we can find $\boldsymbol{A}$ we will have found $\boldsymbol{B}$. Inserting this into Faraday's law, setting time derivatives to zero for the static case, we have

$$\nabla \times \boldsymbol{B} = \nabla \times (\nabla \times \boldsymbol{A}) = \frac{1}{c}\boldsymbol{J}. \tag{2.3.3}$$

Lemma 2.3.4 Given a smooth vector field $\boldsymbol{V} \colon \mathbb{R}^3 \to \mathbb{R}^3$ we have

$$\nabla^2 \boldsymbol{V} = \nabla(\nabla \cdot \boldsymbol{V}) - \nabla \times (\nabla \times \boldsymbol{V}). \tag{2.3.5}$$

Proof. Consider the right hand side in index notation:

$$[\nabla(\nabla \cdot \boldsymbol{V}) - \nabla \times (\nabla \times \boldsymbol{V})]_i = \partial_i(\nabla \cdot \boldsymbol{V}) - \varepsilon_{ijk}\partial_j(\nabla \times \boldsymbol{V})_k \tag{2.3.6}$$

$$= \partial_i\partial_j V_j - \varepsilon_{ijk}\partial_j\varepsilon_{klm}\partial_l V_m. \tag{2.3.7}$$

Now we use the identity

$$\varepsilon_{ijk}\varepsilon_{klm} = \delta_{il}\delta_{jm} - \delta_{im}\delta_{jl}. \tag{2.3.8}$$

This gives

$$[\nabla(\nabla \cdot \boldsymbol{V}) - \nabla \times (\nabla \times \boldsymbol{V})]_i = \partial_i\partial_j V_j - (\delta_{il}\delta_{jm} - \delta_{im}\delta_{jl})\partial_j\partial_l V_m \tag{2.3.9}$$

$$= \partial_i\partial_j V_j - \partial_j\partial_i V_j + \partial_j\partial_j V_i \tag{2.3.10}$$

$$= \partial_j\partial_j V_i \tag{2.3.11}$$

$$= [\nabla^2 \boldsymbol{V}]_i. \qquad \square$$

Using this identity we have

$$\nabla \times \boldsymbol{B} = \nabla \times (\nabla \times \boldsymbol{A}) = \nabla(\nabla \cdot \boldsymbol{A}) - \nabla^2 \boldsymbol{A} = \frac{1}{c}\boldsymbol{J}. \tag{2.3.12}$$

This gives us an equation that can, in theory, be solved for the vector potential given the current density.

In the static case Faraday's law becomes $\nabla \times \boldsymbol{E} = \boldsymbol{0}$. It can be shown that whenever the curl vanishes there is a **scalar potential**, φ, such that $\boldsymbol{E} = -\nabla\varphi$. So if we find φ we have solved Faraday's law. Substituting this definition into Gauss' law we have

$$\nabla \cdot \boldsymbol{E} = \nabla \cdot (-\nabla\varphi) = \rho \implies \nabla^2\varphi = -\rho. \tag{2.3.13}$$

This gives us an equation that we can solve for φ given a charge density, ρ.

2.3.2 Dynamic Case

The dynamic case is similar to the static case but the non-vanishing time derivatives make things a bit more complicated. The process is the same, we take the homogeneous Maxwell equations, that is the equations without sources, and we introduce potentials to solve these. We then take the inhomogeneous Maxwell equations, with sources, and we derive equations for the potentials in terms of the sources, which we can then, in theory, solve for the potentials.

Since we still have $\nabla \cdot \boldsymbol{B} = 0$ we still have some vector potential, $\boldsymbol{A}$, such that $\boldsymbol{B} = \nabla \times \boldsymbol{A}$. However, the equation for this vector potential will be different since we now have a time derivative term in Ampère's law.

Next consider Faraday's law:

$$\nabla \times \boldsymbol{E} = -\frac{1}{c}\frac{\partial \boldsymbol{B}}{\partial t} \neq \boldsymbol{0}. \tag{2.3.14}$$

This means that there does *not* exist a scalar potential, φ, such that $\boldsymbol{E} = -\nabla\varphi$. Instead, we can insert the vector potential giving

$$\nabla \times \boldsymbol{E} = -\frac{1}{c}\frac{\partial}{\partial t}(\nabla \times \boldsymbol{A}) \implies \nabla \times \left(\boldsymbol{E} + \frac{1}{c}\frac{\partial \boldsymbol{A}}{\partial t}\right) = 0. \tag{2.3.15}$$

So there *is* a scalar potential, φ, such that

$$\boldsymbol{E} + \frac{1}{c}\frac{\partial \boldsymbol{A}}{\partial t} = -\nabla\varphi \implies \boldsymbol{E} = -\nabla\varphi - \frac{1}{c}\frac{\partial \boldsymbol{A}}{\partial t}. \tag{2.3.16}$$

We have now solved the two homogenous Maxwell equations, given that we can find the two potentials, φ and $\boldsymbol{A}$. In order to find equations for these we consider the inhomogeneous equations. First, since $\nabla \cdot \boldsymbol{E} = \rho$ we have

$$\nabla \cdot \boldsymbol{E} = \nabla \cdot \left(-\nabla\varphi - \frac{1}{c}\frac{\partial \boldsymbol{A}}{\partial t}\right) \implies \nabla^2\varphi + \frac{1}{c}\frac{\partial}{\partial t}\nabla \cdot \boldsymbol{A} = -\rho. \tag{2.3.17}$$

This equation si complicated, but in theory solvable. Similarly starting with Ampère's law we have

$$\nabla \times \boldsymbol{B} = \nabla \times (\nabla \times \boldsymbol{A}) = \frac{1}{c}\boldsymbol{J} + \frac{1}{c}\frac{\partial \boldsymbol{E}}{\partial t}. \tag{2.3.18}$$

Using our identity for the curl of the curl again we have

$$\nabla(\nabla \cdot \boldsymbol{A}) - \nabla^2 \boldsymbol{A} = \frac{1}{c}\boldsymbol{J} + \frac{1}{c}\frac{\partial \boldsymbol{E}}{\partial t}. \tag{2.3.19}$$

Rearranging and substituting in the potentials for $\boldsymbol{E}$ we get

$$\nabla^2 \boldsymbol{A} - \frac{1}{c^2}\frac{\partial^2 \boldsymbol{A}}{\partial t^2} - \nabla\left(\nabla \cdot \boldsymbol{A} + \frac{1}{c}\frac{\partial \varphi}{\partial t}\right) = -\frac{1}{c}\boldsymbol{J}. \tag{2.3.20}$$

This is again complicated but solvable. Fortunately we have a certain freedom in the definitions of the potentials which allows us to simplify these equations.

2.4 Gauge Transformations

Suppose we replace the vector potential, $\boldsymbol{A}$, with another vector potential, $\boldsymbol{A}'$, according to

$$\boldsymbol{A} \to \boldsymbol{A}' = \boldsymbol{A} + \nabla\chi, \tag{2.4.1}$$

for some arbitrary smooth scalar field χ. Then we have

$$\nabla \times \boldsymbol{A}' = \nabla \times \boldsymbol{A} + \nabla \times (\nabla\chi). \tag{2.4.2}$$

Lemma 2.4.3 For a smooth scalar field, $f : \mathbb{R}^3 \to \mathbb{R}$, we have

$$\nabla \times (\nabla f) = 0. \tag{2.4.4}$$

Proof. Consider this equation in index notation:

$$[\nabla \times (\nabla f)]_i = \varepsilon_{ijk}\partial_j(\nabla f)_k = \varepsilon_{ijk}\partial_j\partial_k f. \tag{2.4.5}$$

Notice that $\partial_j\partial_k$ is symmetric in j and k, whereas ε_{ijk} is antisymmetric. Hence this vanishes proving our lemma. $\square$

Using this we have that

$$\nabla \times \boldsymbol{A}' = \nabla \times \boldsymbol{A} + \nabla \times (\nabla\chi) = \nabla \times \boldsymbol{A} = \boldsymbol{B}. \tag{2.4.6}$$

So the important thing, which is the $\boldsymbol{B}$ field, is unchanged by this transformation.

Now, consider what happens if we change the scalar potential from φ to φ' according to

$$\varphi \to \varphi' = \varphi - \frac{1}{c}\frac{\partial \chi}{\partial t}. \tag{2.4.7}$$

We then have

$$-\nabla\varphi' - \frac{1}{c}\frac{\partial \boldsymbol{A}'}{\partial t} = -\nabla\varphi + \frac{1}{c}\nabla\frac{\partial \chi}{\partial t} - \frac{1}{c}\frac{\partial \boldsymbol{A}}{\partial t} - \frac{1}{c}\frac{\partial}{\partial t}\nabla\chi = -\nabla\varphi - \frac{1}{c}\frac{\partial \boldsymbol{A}}{\partial t} = \boldsymbol{E}. \tag{2.4.8}$$

So, again, the important thing, now $\boldsymbol{E}$, is not changed by this transformation.

We call these transformations **gauge transformations**, and the fact that $\boldsymbol{E}$ and $\boldsymbol{B}$ don't change under these transformations is known as **gauge invariance**. Essentially what we have is a freedom in choice, both pairs, $(\varphi, \boldsymbol{A})$ and $(\varphi', \boldsymbol{A}')$, describe the electromagnetic fields equally well. We exploit this freedom to simplify equations.

2.4.1 Lorenz Gauge

Suppose that we choose to define χ such that

$$\nabla \cdot \boldsymbol{A} + \frac{1}{c}\frac{\partial \varphi}{\partial t} = 0. \tag{2.4.9}$$

This is called the **Lorenz gauge**. If this is the case then we have

$$\nabla \cdot \boldsymbol{A} = -\frac{1}{c}\frac{\partial \varphi}{\partial t}, \tag{2.4.10}$$

which when substituted into Equation (2.3.17) gives

$$\nabla^2\varphi + \frac{1}{c}\frac{\partial}{\partial t}\left(-\frac{1}{c}\frac{\partial \varphi}{\partial t}\right) = \nabla^2\varphi - \frac{1}{c^2}\frac{\partial^2\varphi}{\partial t^2} = -\rho. \tag{2.4.11}$$

Similarly inserting this in Equation (2.3.20) the third term vanishes giving

$$\nabla^2\boldsymbol{A} - \frac{1}{c^2}\frac{\partial^2\boldsymbol{A}}{\partial t^2} = -\frac{1}{c}\boldsymbol{J}. \tag{2.4.12}$$

These are wave equations with sources which can be solved much more easily than before we imposed the Lorenz gauge condition.

Notice that in the static case these wave equations reduce to Poisson's equation:

$$\nabla^2\varphi = -\rho, \qquad \text{and} \qquad \nabla^2\boldsymbol{A} = -\frac{1}{c}\boldsymbol{J}. \tag{2.4.13}$$

2.4.1.1 Imposing the Lorenz Gauge

Is it always possible to impose the Lorenz gauge condition? Suppose we have potentials $(\varphi, \boldsymbol{A})$ such that

$$\nabla \cdot \boldsymbol{A} + \frac{1}{c}\frac{\partial \varphi}{\partial t} = \psi \neq 0. \tag{2.4.14}$$

We then have

$$\nabla \cdot (\boldsymbol{A}' - \nabla\chi) + \frac{1}{c}\frac{\partial}{\partial t}\left(\varphi' + \frac{1}{c}\frac{\partial \chi}{\partial t}\right) = \psi. \tag{2.4.15}$$

Hence, we have

$$\nabla \cdot \boldsymbol{A}' + \frac{1}{c}\frac{\partial \varphi'}{\partial t} = \psi + \nabla^2\chi - \frac{1}{c^2}\frac{\partial \chi}{\partial t}. \tag{2.4.16}$$

So, in order to impose the Lorenz gauge condition on $(\varphi', \boldsymbol{A}')$ we need to find χ such that the right hand side vanishes. This can be done easily by choosing to have χ satisfy

$$\nabla^2\chi - \frac{1}{c^2}\frac{\partial \chi}{\partial t} = -\psi, \tag{2.4.17}$$

which is again a wave equation with a source and can be solved so we can find potentials satisfying the Lorenz gauge conditions.

Three

Relativity

Special relativity is based on two postulates:

- The speed of light is the same in all inertial frames.

- The laws of physics take the same mathematical form in all inertial frames.

It is the second of these which we shall focus on. In special relativity we transform between inertial reference frames with Lorentz transformations. This means we should look for laws of physics which are invariant under Lorentz transformations. Consider rotations. We know that $\boldsymbol{x} \cdot \boldsymbol{y}$ is invariant under rotation, therefore any equation that we see made entirely of scalar products is clearly invariant under rotations. We will now introduce a special relativity equivalent of this notation which makes it immediately clear when an equation is Lorentz invariant. We will be motivated by two examples.

First, we ask if the current conservation law,

$$\nabla \cdot \boldsymbol{J} + \frac{\partial \varphi}{\partial t} = 0, \tag{3.0.1}$$

is Lorentz invariant. Clearly it is rotationally invariant, since it is built from dot products and scalars, since time is unaffected by rotations. It is not immediately clear but it is also invariant under Lorentz rotations.

Second, we ask if the Lorentz force law,

$$\boldsymbol{f} = \rho \boldsymbol{E} + \frac{1}{c} \boldsymbol{J} \times \boldsymbol{B}, \tag{3.0.2}$$

is Lorentz invariant. Again, it is clearly invariant under rotations since the left and right hand sides are both vectors and therefore vary in the same way under rotations. We say that this equation is **covariant** under rotations.

16

3.1 Lorentz Transformations

If people are confused we're going to have an even more confusing thing in a moment. Which is not *my fault.*

Donal O'Connell

Lengths and time differences aren't preserved by Lorentz transformations, one of the first things that one learns about special relativity is the existence of time dilation and length contractions. Instead the invariant spacetime interval between two events is conserved. In order to define this we need to combine the position, $\boldsymbol{x} = (x, y, z)$, and the time, t, into a single object, known as the position **four-vector**:

$$x^\mu := (ct, x, y, z) = (ct, \boldsymbol{x}) = (x^0, x^1, x^2, x^3) = (x^0, x^i). \tag{3.1.1}$$

The space time interval between and event, x^μ, and the origin, $(0,0,0,0)$, is defined to be

$$s^2 := (ct)^2 - x^2 - y^2 - z^2 = \begin{pmatrix} ct & x & y & z \end{pmatrix} \begin{pmatrix} 1 & & & \\ & -1 & & \\ & & -1 & \\ & & & -1 \end{pmatrix} \begin{pmatrix} ct \\ x \\ y \\ z \end{pmatrix}. \tag{3.1.2}$$

We use this to define the **Minkowski metric**

$$\eta_{\mu\nu} := \begin{pmatrix} 1 & & & \\ & -1 & & \\ & & -1 & \\ & & & -1 \end{pmatrix}. \tag{3.1.3}$$

We then define the scalar product according to

$$x \cdot x := x^\mu \eta_{\mu\nu} x^\nu = x_\mu x^\mu \tag{3.1.4}$$

where we define

$$x_\mu := \eta_{\mu\nu} x^\nu. \tag{3.1.5}$$

From this we see that

$$x_\mu = (ct, -x, -y, -z) = (ct, -\boldsymbol{x}) = (x^0, -x^1, -x^2, -x^3) = (x^0, -x^i). \tag{3.1.6}$$

So $x_0 = x^0$ and $x_i = -x^i$. We call vectors with upper indices **contravariant** and vectors with lower indices **covariant**.

We define the **Kronecker delta** to be the identity matrix

$$\delta^\mu{}_\nu := \begin{pmatrix} 1 & & & \\ & 1 & & \\ & & 1 & \\ & & & 1 \end{pmatrix} = \mathbb{1}_4, \tag{3.1.7}$$

and the **inverse metric** to be

$$\eta^{\mu\nu} = \begin{pmatrix} 1 & & & \\ & -1 & & \\ & & -1 & \\ & & & -1 \end{pmatrix}. \tag{3.1.8}$$

Notice that numerically this is the same as $\eta_{\mu\nu}$, but it plays a different role in calculations, and the numerical equivalence only holds in a flat spacetime. The inverse metric is of course defined such that

$$\eta_{\mu\nu}\eta^{\nu\rho} = \delta^\rho{}_\mu \iff \eta^{-1}\eta = \mathbb{1}_4. \tag{3.1.9}$$

Suppose we have two frames, S and S', such that the origin of S' is travelling along the x-axis of S at constant speed v and at time $t = 0$ the origins of the frames coincide. This is called the **standard configuration**, and the transformation between these frames is the **standard Lorentz transformation**, and is given by

$$ct' = x'^0 = \gamma(x^0 - \beta x^1), \tag{3.1.10}$$
$$x'^1 = \gamma(x^1 - \beta x^0), \tag{3.1.11}$$
$$x'^2 = x^2, \tag{3.1.12}$$
$$x'^3 = x^3. \tag{3.1.13}$$

We can write this more compactly as

$$x'^\mu = \Lambda^\mu{}_\nu x^\nu \tag{3.1.14}$$

where $\Lambda^\mu{}_\nu$ is the matrix

$$\Lambda^\mu{}_\nu = \begin{pmatrix} \gamma & -\gamma\beta & 0 & 0 \\ -\gamma\beta & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{3.1.15}$$

We want the scalar product, $x \cdot x := x^\mu \eta_{\mu\nu} x^\nu$ to be Lorentz invariant, meaning that $x'^\mu \eta_{\mu\nu} x'^\nu = x^\mu \eta_{\mu\nu} x^\nu$. Using the Lorentz transformation for x'^μ this gives

$$x'^\mu \eta_{\mu\nu} x'^\nu = (\Lambda^\mu{}_\rho x^\rho)\eta_{\mu\nu}(\Lambda^\nu{}_\sigma x^\sigma) \tag{3.1.16}$$
$$= x^\rho(\Lambda^\mu{}_\rho \eta_{\mu\nu} \Lambda^\nu{}_\sigma)x^\sigma \tag{3.1.17}$$
$$= x^\rho \eta_{\rho\sigma} x^\sigma. \tag{3.1.18}$$

For this to hold we must have that

$$\Lambda^\mu{}_\rho \eta_{\mu\nu} \Lambda^\nu{}_\sigma = \eta_{\rho\sigma} \iff \Lambda^\mathsf{T} \eta \Lambda = \eta. \tag{3.1.19}$$

This property defines Lorentz transformations.

Now that we know how a contravariant vector, x^μ, transforms we can work out how a covariant vector transforms, since we have $x_\mu = \eta_{\mu\nu} x^\nu$. We therefore have

$$x'_\mu = \eta_{\mu\nu} x'^\nu = \eta_{\mu\nu} \Lambda^\nu{}_\rho x^\rho = \eta_{\mu\nu} \Lambda^\nu{}_\rho \eta^{\rho\sigma} x_\sigma. \tag{3.1.20}$$

Hence, the transformation law for a covariant vector is

$$x'_\mu = \Lambda_\mu{}^\nu x_\nu \tag{3.1.21}$$

where we define

$$\Lambda_\mu{}^\nu := \eta_{\mu\sigma}\Lambda^\sigma{}_\rho\eta^{\rho\nu}. \tag{3.1.22}$$

We can find out how $\Lambda^\mu{}_\nu$ and $\Lambda_\mu{}^\nu$ are related by considering the invariant spacetime interval:

$$x_\mu x^\mu = x'_\mu x'^\mu = (\Lambda_\mu{}^\rho x_\rho)(\Lambda^\mu{}_\sigma x^\sigma) = (\Lambda_\mu{}^\rho \Lambda^\mu{}_\sigma)x_\rho x^\sigma. \tag{3.1.23}$$

In order for this to hold we must have that

$$\Lambda_\mu{}^\rho\Lambda^\mu{}_\sigma = \delta^\rho{}_\sigma. \tag{3.1.24}$$

From this we see that $\Lambda_\mu{}^\rho$ is the transpose of the inverse of $\Lambda^\mu{}_\rho$.

We can readily extend these definitions to tensors. First, and simplest, we have scalars, f, which are the same in all frames, that is

$$f' = f. \tag{3.1.25}$$

Then we have contravariant vectors, v^μ, and covariant vectors, v_μ, which transform as

$$v'^\mu = \Lambda^\mu{}_\nu v^\nu, \qquad \text{and} \qquad v'_\mu = \Lambda_\mu{}^\nu v_\nu. \tag{3.1.26}$$

For second rank tensors there are three options, contravariant, $T^{\mu\nu}$, covariant, $T_{\mu\nu}$, and mixed, $T^\mu{}_\nu$. These transform according to

$$T'^{\mu\nu} = \Lambda^\mu{}_\rho\Lambda^\nu{}_\sigma T^{\rho\sigma}, \tag{3.1.27}$$

$$T'_{\mu\nu} = \Lambda_\mu{}^\rho\Lambda_\nu{}^\sigma T_{\rho\sigma}, \tag{3.1.28}$$

$$T'^\mu{}_\nu = \Lambda^\mu{}_\rho\Lambda_\nu{}^\sigma T^\rho{}_\sigma. \tag{3.1.29}$$

3.2 Derivatives

Consider the following derivative:

$$\frac{\partial}{\partial x^\mu}x^\mu = \frac{\partial}{\partial x^0}x^0 + \frac{\partial}{\partial x^1}x^1 + \frac{\partial}{\partial x^2}x^2 + \frac{\partial}{\partial x^3}x^3 = 4. \tag{3.2.1}$$

Clearly this is a scalar, which means that $\partial/\partial x^\mu$ must transform as a covariant vector. We therefore define

$$\partial_\mu := \frac{\partial}{\partial x^\mu} = \left(\frac{\partial}{\partial x^0}, \frac{\partial}{\partial x^1}, \frac{\partial}{\partial x^2}, \frac{\partial}{\partial x^3}\right) = (\partial_0, \partial_1, \partial_2, \partial_3) = \left(\frac{1}{c}\frac{\partial}{\partial t}, \nabla\right). \tag{3.2.2}$$

We can then also define a contravariant derivative

$$\partial^\mu = \eta^{\mu\nu}\partial_\nu = \left(\frac{1}{c}\frac{\partial}{\partial t}, -\nabla\right). \tag{3.2.3}$$

Notice that

$$\frac{\partial x'^{\mu}}{\partial x^{\nu}} = \frac{\partial}{\partial x^{\nu}} \Lambda^{\mu}{}_{\sigma} x^{\sigma} = \Lambda^{\mu}{}_{\nu}. \tag{3.2.4}$$

We can build a scalar operator as the scalar product of derivative operators:

$$\partial^2 := \partial_{\mu}\partial^{\mu} = \eta^{\mu\nu}\partial_{\mu}\partial_{\nu} = \frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \nabla^2. \tag{3.2.5}$$

This same operator is sometimes denoted $\square$ or $\square^2$, and is a four-dimensional analogue of the Laplacian. Notice that this allows us to compactly write the wave equation as

$$\partial^2 f = 0. \tag{3.2.6}$$

In particular this also shows that the wave equation is unchanged under Lorentz transformations.

3.3   Lorentz Group

Consider two contravariant vectors, x and y. We know that $x^{\mathsf{T}}\eta y$ is invariant, so

$$x'^{\mathsf{T}}\eta y' = (\Lambda x)^{\mathsf{T}}\eta \Lambda y = x^{\mathsf{T}}\Lambda^{\mathsf{T}}\eta \Lambda y = x\eta y \tag{3.3.1}$$

from which we see that

$$\Lambda^{\mathsf{T}}\eta\Lambda = \eta. \tag{3.3.2}$$

In fact we can take this as the defining condition of a Lorentz transformation, and define the **Lorentz group** to be

$$O(1,3) := \{\Lambda \in \mathcal{M}_4(\mathbb{R}) \mid \Lambda^{\mathsf{T}}\eta\Lambda = \eta\}. \tag{3.3.3}$$

Applying the determinant to this defining condition we have

$$\det\eta = \det(\Lambda^{\mathsf{T}}\eta\Lambda) = \det(\Lambda^{\mathsf{T}})\det(\eta)\det(\Lambda). \tag{3.3.4}$$

Using the fact that $\det\eta = -1 \neq 0$ and $\det M^{\mathsf{T}} = \det M$ for all matrices M we have that $(\det\Lambda)^2 = 1$, which means that $\det\Lambda = \pm 1$. We define a **proper Lorentz transform** to be one with $\det\Lambda = +1$, and an improper Lorentz transform has $\det\Lambda = -1$. The proper Lorentz transformations also form a group:

$$SO(1,3) := \{\Lambda \in \mathcal{M}_4(\mathbb{R}) \mid \Lambda^{\mathsf{T}}\eta\Lambda = \eta \text{ and } \det\Lambda = +1\}. \tag{3.3.5}$$

Note that the Lorentz group and proper Lorentz group contain the rotation groups and proper rotation groups, $O(3)$ and $SO(3)$, respectively as subgroups, leaving the time unaffected.

3.4 Special Tensors

Consider the **Levi–Civita symbol** in four dimensions. We can define a contravariant tensor, $\varepsilon^{\mu\nu\rho\sigma}$. Applying the transformation law for a rank four tensor we have

$$\varepsilon'^{\mu\nu\rho\sigma} = \Lambda^{\mu}{}_{\alpha}\Lambda^{\nu}{}_{\beta}\Lambda^{\rho}{}_{\gamma}\Lambda^{\sigma}{}_{\delta}\varepsilon^{\alpha\beta\gamma\delta}. \tag{3.4.1}$$

We can now use a property of the determinant,

$$\Lambda^{\mu}{}_{\alpha}\Lambda^{\nu}{}_{\beta}\Lambda^{\rho}{}_{\gamma}\Lambda^{\sigma}{}_{\delta}\varepsilon^{\alpha\beta\gamma\delta} = \varepsilon^{\mu\nu\rho\sigma} \det\Lambda. \tag{3.4.2}$$

If Λ is a proper Lorentz transformation we therefore have

$$\varepsilon'^{\mu\nu\rho\sigma} = \varepsilon^{\mu\nu\rho\sigma}, \tag{3.4.3}$$

so the Levi–Civita symbol is invariant under Lorentz transforms. If instead Λ is an improper Lorentz transformation then we pick up a minus sign,

$$\varepsilon'^{\mu\nu\rho\sigma} = -\varepsilon^{\mu\nu\rho\sigma}, \tag{3.4.4}$$

so we see that the Levi–Civita symbol is really a pseudo tensor.

We choose to define $\varepsilon_{0123} = +1$. Notice that this then means that $\varepsilon^{0123} = -1$, since we can raise the 0 for free and pick up a minus sign raising each of 1, 2, and 3.

More care is needed when finding the value of various index combinations for the Levi–Civita symbol in four dimensions since it is no longer sufficient to consider cyclic permutations, anti-cyclic permutations, and repeating indices like in three dimensions. For example, 1230 is a cyclic permutation of 0123 but it is an odd permutation, since we need 3 swaps, $1230 \to 1203 \to 1023 \to 0123$, to get back to 0123. Hence $\varepsilon_{1230} = -1$. Of course, if an index is repeated we still get zero like in three dimensions.

One of the main uses of the Levi–Civita symbol is in defining the **volume element**, which in three dimensions is

$$\mathrm{d}^3 x := \frac{1}{3!}\varepsilon_{ijk}\,\mathrm{d}x^i\,\mathrm{d}x^j\,\mathrm{d}x^k = \mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z, \tag{3.4.5}$$

and in four-dimensions is

$$\mathrm{d}^4 x := \frac{1}{4!}\varepsilon_{\mu\nu\rho\sigma}\,\mathrm{d}x^{\mu}\,\mathrm{d}x^{\nu}\,\mathrm{d}x^{\rho}\,\mathrm{d}x^{\sigma} = c\,\mathrm{d}t\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z. \tag{3.4.6}$$

It should be noted that the three dimensional volume form is *not* Lorentz invariant. Choose the x direction to be in the direction of motion and length contraction gives us $\mathrm{d}x \to \mathrm{d}x/\gamma$, so

$$\mathrm{d}^3 x \to \frac{1}{\gamma}\mathrm{d}^3 x. \tag{3.4.7}$$

Another common special tensor is the **Kronecker delta**, $\delta^{\mu}{}_{\nu}$, which is defined to be 1 if $\mu = \nu$ and 0 otherwise. Consider how this transforms, given that it is a mixed tensor:

$$\delta^{\mu}{}_{\nu} = \Lambda^{\mu}{}_{\rho}\Lambda_{\nu}{}^{\sigma}\delta^{\rho}{}_{\sigma} = \Lambda^{\mu}{}_{\rho}\Lambda_{\nu}{}^{\rho} = \delta^{\mu}{}_{\nu}, \tag{3.4.8}$$

where we have used the fact that the contravariant and covariant Lorentz transformations are inverses. So $\delta^{\mu}{}_{\nu}$ is a tensor and is invariant under Lorentz transformations.

Four

Dynamics

4.1 Proper Time

So far we have considered single points in spacetime. Now consider the path of some particle. We can specify the path as $(t, x(t), y(t), z(t))$, so given a time, t, in a given frame we can work out where the particle is. Given a path like this we can find a scalar value that everyone will agree upon, called the **proper time**, τ. This is defined by

$$d\tau^2 := \frac{1}{c^2} ds^2 = dt^2 - \frac{1}{c^2} d\boldsymbol{x}^2. \tag{4.1.1}$$

We can then use the fact that x is parametrised as a function of t to write dx^i as

$$dx^i = dx^i(t) = \frac{dx^i}{dt} dt = v^i dt, \tag{4.1.2}$$

where we define the (coordinate) velocity to be $\boldsymbol{v} := d\boldsymbol{x}/dt$. We then have that

$$d\tau^2 = dt^2 - \frac{1}{c^2} dx^i dx^i = \left(1 - \frac{\boldsymbol{v}^2}{c^2}\right) dt^2. \tag{4.1.3}$$

Choosing to take the positive square root, so that τ increases as t increases, we get

$$d\tau = \frac{1}{\gamma} dt, \qquad \text{where} \qquad \gamma := \frac{1}{\sqrt{1 - \boldsymbol{v}^2/c^2}}. \tag{4.1.4}$$

Dividing through by dt we get and ordinary differential equation for τ:

$$\frac{d\tau}{dt} = \gamma \implies \tau = \int \gamma \, dt. \tag{4.1.5}$$

In general the velocity, and hence γ, will be a function of time, and this integral may be non-trivial.

We interpret the proper time as the time elapsed along the path in the frame of the particle, since in this frame $dx^i = 0$ and so $ds^2 = c^2 dt^2 = c^2 d\tau^2$, with $\gamma = 1$ in this case. The nice thing about the proper time is that while the coordinate time, t, is frame dependent the proper time is not and all observers will agree on its value. This makes it much more useful for computing values. For this reason we prefer to use the proper time to parametrise our path, $(t(\tau), x(\tau), y(\tau), z(\tau))$, where we interpret $x(\tau)$ as $x(t(\tau))$, and similarly for y and z. We find the function $t(\tau)$

by simply inverting whatever equation we have found for τ by integrating γ with respect to t.

Compare this process to considering a circle. We could in theory give a point on the circle as $(x, y(x))$, but this is awkward to work with. It is much better to find a third value, say the angle from the x-axis, ϑ, and parametrise the circle using this, so $(x(\vartheta), y(\vartheta))$. This gives a much more natural interpretation to quantities like $\mathrm{d}\boldsymbol{r}/\mathrm{d}\vartheta = (x'(\vartheta), y'(\vartheta))$, as the velocity, rather than $\mathrm{d}\boldsymbol{r}/\mathrm{d}x = (1, y'(x))$ with the old parametrisation. We are used to using the time to parametrise the position in Newtonian mechanics, it is only in relativity when time and space are treated on a more equal footing where this starts to fail.

4.2 Proper Velocity and Acceleration

4.2.1 Proper Velocity

Now that we have parametrised the path of a particle in times of the proper time it makes sense to define the **proper velocity**, u^μ, as the proper time derivative of the position:

$$u^\mu := \frac{\mathrm{d}x^\mu}{\mathrm{d}\tau} = \frac{\mathrm{d}t}{\mathrm{d}\tau}\frac{\mathrm{d}x^\mu}{\mathrm{d}t} = \gamma\frac{\mathrm{d}x^\mu}{\mathrm{d}t} = \gamma v^\mu. \tag{4.2.1}$$

Here we are defining the **coordinate velocity**, v^μ, as the coordinate time derivative of the position. These two velocities are therefore related by

$$u^\mu = \gamma v^\mu = (\gamma c, \gamma\boldsymbol{v}). \tag{4.2.2}$$

Note that the proper velocity, unlike the coordinate velocity, is the derivative of a four-vector by a scalar quantity, and hence it is a four-vector.

Going back to the circle example we see that the proper velocity is like $\mathrm{d}\boldsymbol{x}/\mathrm{d}\vartheta$ and the coordinate velocity is like $\mathrm{d}\boldsymbol{x}/\mathrm{d}x$.

Consider the square of the proper velocity:

$$u^2 = u \cdot u = \gamma^2(c^2 - \boldsymbol{v}^2) = \frac{1}{1 - \boldsymbol{v}^2/c^2}(c^2 - \boldsymbol{v}^2) = c^2. \tag{4.2.3}$$

We see that thee proper velocity is always normalised to c. Another way to see this is to consider the definition of the proper time:

$$u^2 = \frac{\mathrm{d}x^\mu}{\mathrm{d}\tau}\frac{\mathrm{d}x_\mu}{\mathrm{d}\tau} = \left(\frac{\mathrm{d}s}{\mathrm{d}\tau}\right)^2 = c^2, \tag{4.2.4}$$

since $\mathrm{d}\tau^2 = \mathrm{d}s^2/c^2$. A third way to see this is to consider the proper velocity in the rest frame of the particle, in which case $u^\mu = (c, \boldsymbol{0})$, since $\gamma = 1$ in this frame, and we then trivially have $u^2 = c^2$. This then holds in any frame since u^2 is a Lorentz scalar.

4.2.2 Proper Acceleration

We can keep going taking higher order derivatives. We define the **proper acceleration**, a^μ, to be

$$a^\mu := \frac{\mathrm{d}u^\mu}{\mathrm{d}\tau} = \frac{\mathrm{d}^2 x^\mu}{\mathrm{d}\tau^2}. \tag{4.2.5}$$

This is a again a four-vector, since it is the derivative of a four-vector with respect to a scalar[1].

Now consider the following:

$$\frac{\mathrm{d}}{\mathrm{d}\tau}(u \cdot u) = 2u \cdot \frac{\mathrm{d}u}{\mathrm{d}\tau} = 2u \cdot a, \qquad \text{and} \qquad \frac{\mathrm{d}}{\mathrm{d}\tau}(u \cdot u) = \frac{\mathrm{d}}{\mathrm{d}\tau}c^2 = 0. \qquad (4.2.6)$$

This shows that the proper acceleration is always orthogonal to the proper acceleration. This is a bit like circular motion, where we have centripetal acceleration towards the centre and the velocity is tangential. The way we interpret this is that there are only really three degrees of freedom here and the time component of the proper acceleration doesn't really mean much, its just chosen to make u and a perpendicular.

4.3   Momentum

We make the assumption that each particle has an associated Lorentz scalar mass. That is the mass is the same in all frames. This is different to the more old fashioned idea of "relativistic mass" which scales with γ. We further assume that this is constant, which basically means we don't consider cases where particles decay or combine to form other particles. We want to define a relativistic analogue of the momentum. It seems sensible that this should be done by replacing the coordinate velocity in three dimensions with the proper velocity in four dimensions:

$$p^\mu := mu^\mu = m\frac{\mathrm{d}x^\mu}{\mathrm{d}\tau} = (\gamma mc, \gamma m\boldsymbol{v}) = (p^0, \boldsymbol{p}). \qquad (4.3.1)$$

Note that $\boldsymbol{p} := \gamma m\boldsymbol{v} \neq m\boldsymbol{v}$, this last term is the non-relativistic momentum.

For the interpretation of the first term consider the case when $v \ll c$. We can Taylor expand γ giving

$$\gamma = (1 - v^2/c^2)^{-1/2} \approx 1 + \frac{1}{2}\frac{v^2}{c^2}. \qquad (4.3.2)$$

We then have that

$$p^0 = \gamma mc = \frac{1}{c}\gamma mc^2 \approx \frac{1}{c}\left(1 + \frac{1}{2}\frac{v^2}{c^2}\right)mc^2 = \frac{mc^2}{c} + \frac{1}{2}mv^2. \qquad (4.3.3)$$

We can identify the first term as the rest energy and the second as the non-relativistic kinetic energy, this suggests we interpret p^0 as the energy, E/c:

$$p^\mu = (E/c, \boldsymbol{p}). \qquad (4.3.4)$$

Notice that keeping terms only to first order in v/c we also recover the non-relativistic momentum, $\boldsymbol{p} = \gamma m\boldsymbol{v} \approx m\boldsymbol{v}$.

Now consider the square of the four-momentum:

$$p^2 = p \cdot p = \frac{E^2}{c^2} - \boldsymbol{p}^2. \qquad (4.3.5)$$

In the rest frame of the particle $\boldsymbol{p} = \boldsymbol{0}$ and $\gamma = 1$ so we have $p^\mu = (mc, \boldsymbol{0})$ giving $p^2 = m^2c^2$. This must hold in all frames since p^2 is a scalar and so after a bit of rearranging we have that

$$E^2 = \boldsymbol{p}^2c^2 + m^2c^4. \qquad (4.3.6)$$

[1] Note that this fails in general relativity, since the transformation coefficients are not constant there and so derivatives of four-vectors aren't necessarily four vectors.

This is of course the famous energy-momentum relation, which reduces to the even more famous mass-energy equivalence $E = mc^2$ in the rest frame of the particle. Notice that for a massless particle, such as a photon, we have $E^2 = p^2c^2$, and $p^2 = 0$, in which case we can write

$$p^\mu = \left(\frac{E}{c}, \frac{E}{c}\hat{n}\right), \tag{4.3.7}$$

where $\hat{n}$ is a unit vector, since this gives

$$p^2 = \frac{E^2}{c^2} - \frac{E^2}{c^2}e_i \cdot \hat{n} = 0. \tag{4.3.8}$$

4.3.1 Force

Newton's second law in the non-relativistic case is most generally written as

$$f = \frac{\mathrm{d}p}{\mathrm{d}t}. \tag{4.3.9}$$

The obvious generalisation to special relativity is to define the force to be

$$f^\mu := \frac{\mathrm{d}p^\mu}{\mathrm{d}\tau} = m\frac{\mathrm{d}u^\mu}{\mathrm{d}\tau} = ma^\mu. \tag{4.3.10}$$

Since f^μ is just proportional to the acceleration, assuming constant mass, we have that $f \cdot u = 0$, so the force must be perpendicular to the velocity. This reduces the degrees of freedom so there is no need for an interpretation of f^0.

Part II

Relativistic Electrodynamics

Five

Relativistic Electrodynamic Quantities

5.1 Four Current

Recall that the continuity equation for charge density/current is

$$\frac{\partial \rho}{\partial t} + \nabla \cdot \boldsymbol{\rho} = 0. \tag{5.1.1}$$

The right hand side of this is a Lorentz scalar. We want to write this in a way that makes it relativistic, so the entire equation is invariant under transformations. First we notice that we have time and space derivatives, which in relativity we should combine into a single derivative, ∂_μ. This is a covariant vector and we want to combine it with an object to give a Lorentz scalar, so said object should be contravariant. Call this object $J^\mu = (J^0, \boldsymbol{J})$. We then look to define J^μ such that

$$0 = \partial_\mu J^\mu = \frac{1}{c} \frac{\partial J^0}{\partial t} + \nabla \cdot \boldsymbol{J}. \tag{5.1.2}$$

From this we see that we should interpret J^0 as ρc and $\boldsymbol{J}$ is indeed the current density. We call J^μ the **four-current**, and it follows the **relativistic continuity equation**

$$\partial_\mu J^\mu = 0. \tag{5.1.3}$$

We know that J^μ is contravariant, so it transforms as

$$J'^\mu = \Lambda^\mu{}_\nu J^\nu. \tag{5.1.4}$$

For a standard Lorentz transformation we then have

$$\begin{pmatrix} \rho'c \\ J'^1 \\ J'^2 \\ J'^3 \end{pmatrix} = \begin{pmatrix} \gamma & -\gamma\beta & 0 & 0 \\ -\gamma\beta & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \rho c \\ J^1 \\ J^2 \\ J^3 \end{pmatrix} = \begin{pmatrix} \gamma\rho c - \gamma\beta J^1 \\ \gamma J^1 - \gamma\beta\rho c \\ J^2 \\ J^3 \end{pmatrix}. \tag{5.1.5}$$

Notice that this clearly shows that ρ is *not* a scalar. We can interpret the fact that ρ changes between frames by understanding that ρ is a density, and hence changes when the volume element changes, which happens in special relativity due to Lorentz contraction, in particular $\mathrm{d}^3 x \to \mathrm{d}^3 x / \gamma$. This means that if there is no current then we have $\rho' = \gamma\rho$. Of course, in the limit of $v \ll c$ we have $\gamma \approx 1$ and $\beta/c = v/c^2 \approx 0$ so we recover $\rho \approx \rho'$.

27
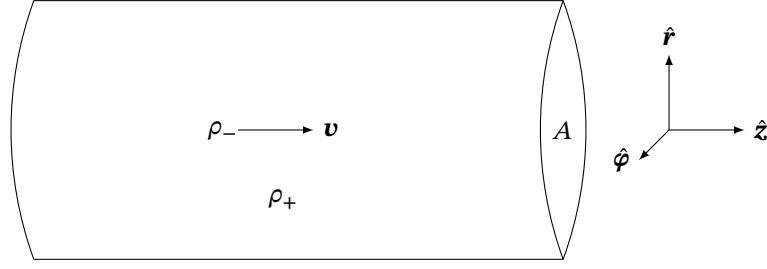
Figure 5.1: Feynman's wire with cross sectional area A carrying current $\boldsymbol{J} = \rho_- \boldsymbol{v}$.

5.1.1 Feynman's Wire

Consider a neutral wire of cross sectional area A. We can consider the wire to be made of two charge densities, one positive, ρ_+, and one negative, ρ_-, such that the total charge density, $\rho = \rho_+ + \rho_-$, is zero. For a metal we think of ρ_+ as the ions and ρ_- as the sea of delocalised electrons. In order for this to hold we must have that $\rho_\pm = \pm\rho_0$ for some constant density ρ_0. When there is a current in the wire this results in the negative charges moving with velocity $\boldsymbol{v}$, giving a current density $\boldsymbol{J} = \rho_- \boldsymbol{v}$.

We can compute the magnetic field outside of the wire using Ampère's law,

$$\nabla \times \boldsymbol{B} = \frac{1}{c}\boldsymbol{J}, \tag{5.1.6}$$

since $\partial \boldsymbol{E}/\partial t = 0$ in this scenario. In order to be useful in this situation we need the integral form of Ampère's law. For this we introduce an Ampèrian loop encircling the wire at a distance r, greater than the radius of the wire, and we integrate Ampère's law over a surface bounded by this loop. The left hand side then becomes

$$\int_S (\nabla \times \boldsymbol{B}) \cdot \mathrm{d}\boldsymbol{S} = \oint \boldsymbol{B} \cdot \mathrm{d}\boldsymbol{r} \tag{5.1.7}$$

through an application of Stoke's theorem, and the right hand side becomes

$$\frac{1}{c}\int_S \boldsymbol{J} \cdot \mathrm{d}\boldsymbol{S} = \frac{1}{c}\rho_- v A, \tag{5.1.8}$$

where we have used the fact that $\boldsymbol{J}$ is $\rho_- \boldsymbol{v}$ inside the wire and zero outside the wire.

By symmetry we have that $\boldsymbol{B}$ depends only on r, the distance from the centre of the wire. The right hand rule tells us that $\boldsymbol{B}$ will be in the $\hat{\boldsymbol{\varphi}}$ direction. So, $\boldsymbol{B}(\boldsymbol{r}) = B(r)\hat{\boldsymbol{\varphi}}$. Hence we have

$$\oint \boldsymbol{B} \cdot \mathrm{d}\boldsymbol{r} = 2\pi r B(r). \tag{5.1.9}$$

Equating these terms and solving for $\boldsymbol{B}$ we have that

$$\boldsymbol{B}(r) = \frac{\rho_- v A}{2\pi r c}\hat{\boldsymbol{\varphi}}. \tag{5.1.10}$$

Note that inside the wire the only thing that changes is we replace A with πr^2. For a neutral wire we have that $\rho_+ = \rho_0$ and $\rho_- = -\rho_0$, which allows us to write the magnetic field as

$$\boldsymbol{B} = -\frac{\rho_0 v A}{2\pi r c}\hat{\boldsymbol{\varphi}}. \tag{5.1.11}$$

Suppose we place a small test charge, q, at a distance r from the wire, such that the test charge also moves with velocity $\boldsymbol{v}$. The force on this test charge will be

$$\boldsymbol{F} = \frac{q}{c}\boldsymbol{v} \times \boldsymbol{B} = \frac{q\rho_0 A}{2\pi r}\frac{v^2}{c^2}\hat{\boldsymbol{r}}. \tag{5.1.12}$$

Here we have used the fact that $\boldsymbol{v} = v\hat{\boldsymbol{z}}$ and $\hat{\boldsymbol{z}} \times \hat{\boldsymbol{\varphi}} = -\hat{\boldsymbol{r}}$. The result is that the test charge will accelerate radially away from the wire, at least initially.

Suppose now, that we transform to a frame in which the negative charges are stationary. In this frame the positive charges now move with velocity $-\boldsymbol{v}$. In the first frame there was a net force on the charge, so we expect there to be a force in this frame also. However, now that q is stationary in this frame the magnetic field cannot be the source of this force. Instead the force is due to the electric field in this frame. To see this we have to transform the current into our frame.

In the first frame we had

$$J^\mu = (0, \rho_-\hat{\boldsymbol{z}}). \tag{5.1.13}$$

Transforming to the new frame with a boost along the z-axis we get that

$$c\rho' = \gamma(c\rho - \beta J^z) = -\beta\rho_- v \implies \rho' = \gamma\beta^2\rho_0. \tag{5.1.14}$$

So, despite being neutral in the first frame we see that the wire is charged in the new frame. This is because densities are not Lorentz invariant, since volumes are subject to Lorentz contraction. The result is that there is an electric field in this frame given by

$$\boldsymbol{E}' = \frac{\rho_0 \gamma A}{2\pi r}\frac{v^2}{c^2}\hat{\boldsymbol{r}}. \tag{5.1.15}$$

This is just the standard electric field for a charged rod,

$$\boldsymbol{E} = \frac{\rho A}{2\pi r}\hat{\boldsymbol{r}}, \tag{5.1.16}$$

which can be derived from Gauss' law. The force on the test charge is then

$$\boldsymbol{F}' = q\boldsymbol{E}' = \gamma\frac{q\rho_0 A}{2\pi r}\frac{v^2}{c^2}\hat{\boldsymbol{r}}. \tag{5.1.17}$$

Note that this force is the same as in the first frame, but with an extra factor of γ, since we are using the force here as defined by $\boldsymbol{F} = \mathrm{d}\boldsymbol{p}/\mathrm{d}t$, rather than $\boldsymbol{F} = \mathrm{d}\boldsymbol{p}/\mathrm{d}\tau$, so the force also changes between frames due to time dilation. The resulting motion of the test charge will of course be the same regardless of the frame we perform the calculations in.

This example shows the rather remarkable fact that we can turn electric fields into magnetic fields and vice versa by transforming to different frames. It should be noted that there is still a magnetic field in the second frame, it just has no effect on the particle since it is stationary. This idea will shortly lead to us combining the electric and magnetic fields into a single object.

5.2   Electromagnetic Potentials

Lorentz and Lorenz. They're different guys. Not that it matters, they're both dead.

Donal O'Connell

Recall that in the Lorenz gauge we have

$$\left(\frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \nabla^2\right)\varphi = \rho, \qquad \text{and} \qquad \left(\frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \nabla^2\right)\boldsymbol{A} = \frac{1}{c}\boldsymbol{J}. \tag{5.2.1}$$

These are simply Equations (2.4.11) and (2.4.12). We can write these more compactly using the d'Alembertian, ∂^2:

$$\partial^2\varphi = \rho = \frac{1}{c}J^0, \qquad \text{and} \qquad \partial^2\boldsymbol{A} = \frac{1}{c}\boldsymbol{J} = \frac{1}{c}J^i\boldsymbol{e_i}. \tag{5.2.2}$$

This suggests that we should define a new quantity $A^\mu = (\varphi, \boldsymbol{A})$, such that

$$\partial^2 A^\mu = \frac{1}{c}J^\mu. \tag{5.2.3}$$

Since the right hand side is a contravariant Lorentz vector and the d'Alembertian is a scalar operator the quotient theorem means that A^μ is a contravariant Lorentz vector also. We call A^μ the **four-potential**, **gauge potential**, or **gauge field**.

5.2.1   Gauge Transformations

Consider the Lorenz gauge condition,

$$\frac{1}{c}\frac{\partial\varphi}{\partial t} + \nabla \cdot \boldsymbol{A} = 0. \tag{5.2.4}$$

We can rewrite this in relativistic notation as

$$\partial_\mu A^\mu = 0, \tag{5.2.5}$$

which shows that this gauge condition is Lorentz invariant. This makes this gauge condition a good choice for our use. Compare this to the Coulomb gauge condition, $\nabla \cdot \boldsymbol{A} = 0$, which is *not* Lorentz invariant, and so isn't of much use when doing relativistic electrodynamics.

Consider the generic gauge transformations

$$\varphi \to \varphi' = \varphi - \frac{1}{c}\frac{\partial\chi}{\partial t}, \qquad \text{and} \qquad \boldsymbol{A} \to \boldsymbol{A}' = \boldsymbol{A} + \nabla\chi. \tag{5.2.6}$$

We can write these in terms of the four-potential as

$$A^0 \to A'^0 = A^0 - \partial_0\chi, \qquad \text{and} \qquad A^i \to A'^i = A^i + \partial_i\chi. \tag{5.2.7}$$

The mixture of up and down indices is not good here, we can lower the time indices freely and the spatial indices at the cost of a minus sign giving

$$A_0 \to A'_0 = A_0 - \partial_0\chi, \quad \text{and} \quad -A_i \to -A'_i = -A_i + \partial_i\chi \implies A_i \to A'_i = A_i - \partial_i\chi.$$

We can then combine these into a single Lorentz invariant statement:

$$A_\mu \to A'_\mu - \partial_\mu\chi. \tag{5.2.8}$$

Six

The Field Tensor

6.1 Electromagnetic Fields

The electric field is given in terms of the potentials by

$$\boldsymbol{E} = -\frac{1}{c}\frac{\partial \boldsymbol{A}}{\partial t} - \nabla\varphi. \tag{6.1.1}$$

We can write this in terms of components as

$$E^i = -\frac{1}{c}\frac{\partial A^i}{\partial t} - \partial_i\varphi = -\partial_0 A^i - \partial_i A^0. \tag{6.1.2}$$

Again, we have mixed up and down indices which we should get rid of:

$$E^i = -\partial^0 A^i + \partial^i A^0. \tag{6.1.3}$$

The magnetic field is given by

$$\boldsymbol{B} = \nabla \times \boldsymbol{A}. \tag{6.1.4}$$

The components of $\boldsymbol{B}$ are then given by

$$B^1 = \partial_2 A^3 - \partial_3 A^2 = -\partial^2 A^3 + \partial^3 A^2, \tag{6.1.5}$$
$$B^2 = \partial_3 A^1 - \partial_1 A^3 = -\partial^3 A^1 + \partial^1 A^3, \tag{6.1.6}$$
$$B^3 = \partial_1 A^2 - \partial_2 A^1 = -\partial^1 A^2 + \partial^2 A^1. \tag{6.1.7}$$
$$\tag{6.1.8}$$

Notice how similar the components of both $\boldsymbol{E}$ and $\boldsymbol{B}$ are when written in terms of A^μ. Following this similarity we define the quantity

$$F^{\mu\nu} := \partial^\mu A^\mu - \partial^\nu A^\mu, \tag{6.1.9}$$

which we call the **electromagnetic field strength tensor**. First, notice that $F^{\mu\nu}$ is an antisymmetric rank 2 tensor, which as given above transforms contravariantly. Since it is antisymmetric it has at most 6 independent, non-zero components, which just so happen to correspond to the components of $\boldsymbol{E}$ and $\boldsymbol{B}$. We can easily see from the definition that

$$F^{i0} = -F^{0i} = E^i. \tag{6.1.10}$$

A bit more work shows us that the components of $\boldsymbol{B}$ appear in F through

$$\varepsilon^{ijk}B^k = \varepsilon^{ijk}(\nabla \times \boldsymbol{A})^k \tag{6.1.11}$$

$$= \varepsilon^{ijk}\varepsilon^{klm}\partial_l A^m \tag{6.1.12}$$

$$= -\varepsilon^{ijk}\varepsilon^{klm}\partial^l A^m \tag{6.1.13}$$

$$= -(\delta^{il}\delta^{jm} - \delta^{im}\delta^{jl})\partial^l A^m \tag{6.1.14}$$

$$= -\partial^i A^j + \partial^j A^i \tag{6.1.15}$$

$$= -F^{ij}. \tag{6.1.16}$$

Since F is antisymmetric its diagonal vanishes. Writing out F in terms of E^i and B^i we have

$$F^{\mu\nu} = \begin{pmatrix} 0 & -E^1 & -E^2 & -E^3 \\ E^1 & 0 & -B^3 & B^2 \\ E^2 & B^3 & 0 & -B^1 \\ E^3 & -B^2 & B^1 & 0 \end{pmatrix}. \tag{6.1.17}$$

We can lower indices on $F^{\mu\nu}$ giving

$$F_\mu{}^\nu = \begin{pmatrix} 0 & -E_x & -E_y & -E_z \\ -E_x & 0 & B_z & -B_y \\ -E_y & -B_z & 0 & B_x \\ -E_z & B_y & -B_x & 0 \end{pmatrix}, \tag{6.1.18}$$

$$F^\mu{}_\nu = \begin{pmatrix} 0 & E_x & E_y & E_z \\ E_x & 0 & B_z & -B_y \\ E_y & -B_z & 0 & B_x \\ E_z & B_y & -B_x & 0 \end{pmatrix}, \tag{6.1.19}$$

$$F_{\mu\nu} = \begin{pmatrix} 0 & E_x & E_y & E_z \\ -E_x & 0 & -B_z & B_y \\ -E_y & B_z & 0 & -B_x \\ -E_z & -B_y & B_x & 0 \end{pmatrix}. \tag{6.1.20}$$

A handy trick for filling in the indices on the B components of the field strength tensor is that the indices follow sudoku like rules. For example consider F^{21}. This cannot be B^1 since $F^{01} = -E^1$, and it cannot be B^2 since $F^{20} = E^2$. So we are left with B^3. For the sign we still have to consider $F^{ij} = -\varepsilon^{ijk}B^k$.

6.2 Gauge Invariance

The field strength tensor inherits gauge invariance from the fields. Consider the gauge transformation $A_\mu \to A'_\mu - \partial_\chi$, we have

$$F_{\mu\nu} \to F'_{\mu\nu} = \partial_\mu(A_\nu - \partial_\nu\chi) - \partial_\nu(A_\mu - \partial_\mu\chi) \tag{6.2.1}$$

$$= \partial_\mu A_\nu - \partial_\nu A_\mu - \partial_\mu\partial_\nu\chi + \partial_\nu\partial_\mu\chi \tag{6.2.2}$$

$$= F_{\mu\nu}, \tag{6.2.3}$$

assuming that χ is sufficiently smooth for partial derivatives to commute.

6.3 Inhomogeneous Maxwell Equations

Consider the divergence of $F^{\mu\nu}$:

$$\partial_\mu F^{\mu\nu} = \partial_\mu(\partial^\mu A^\nu - \partial^\nu A^\mu) = \partial^2 A^\nu - \partial^\nu \partial_\mu A^\mu. \tag{6.3.1}$$

The final term contains $\partial_\mu A^\mu$, which vanishes in the Lorenz gauge, so in the Lorenz gauge we have that

$$\partial_\mu F^{\mu\nu} = \partial^2 A^\nu = \frac{1}{c}J^\nu. \tag{6.3.2}$$

This last equality is Equation (5.2.3), which is the defining property of A^μ.

Now, notice that the left hand side is gauge invariant, since $F^{\mu\nu}$ is gauge invariant. The right hand side is also gauge invariant, since the gauge transformations effect only the potentials, not the sources, which have physical meaning. Whenever we have a start and end point that are gauge invariant it doesn't matter if we made a gauge choice in between, and so we have in *all* gauges that

$$\partial_\mu F^{\mu\nu} = \frac{1}{c}J^\nu. \tag{6.3.3}$$

This equation immediately implies conservation of current. Recall that in relativistic notation the continuity equation is $\partial_\nu J^\nu = 0$, so we take the divergence of $\partial_\mu F^{\mu\nu}$ to give us

$$\partial_\nu \partial_\mu F^{\mu\nu} = \frac{1}{c}\partial_\nu J^\nu. \tag{6.3.4}$$

Now, the left hand side of this is something symmetric in μ and ν, $\partial_\nu \partial_\mu$, and something antisymmetric in μ and ν, $F^{\mu\nu}$. The product therefore vanishes. This is a general rule, that the product of symmetric and antisymmetric components in the same indices vanishes. In this case we can prove it by first commuting the derivatives, then swapping the indices on the field tensor giving a negative sign, and then relabelling $\mu \leftrightarrow \nu$:

$$\partial_\nu \partial_\mu F^{\mu\nu} = \partial_\mu \partial_\nu F^{\mu\nu} = -\partial_\mu \partial_\nu F^{\nu\mu} \overset{\mu \leftrightarrow \nu}{=} -\partial_\nu \partial_\mu F^{\mu\nu}. \tag{6.3.5}$$

Finally, the only thing equal to its negative is 0 so we must have

$$0 = \partial_\nu \partial_\mu F^{\mu\nu} = \frac{1}{c}\partial_\nu J^\nu \implies \partial_\nu J^\nu = 0. \tag{6.3.6}$$

So, $F^{\mu\nu}$ automatically encodes the conservation laws for charge and current.

Now consider again the divergence $\partial_\mu F^{\mu\nu} = J^\nu/c$. In order to recover our pre-relativistic-notation equations we should treat time and space separately. First, consider the case of $\nu = 0$. We then have $J^0/c = \rho$, giving

$$\partial_\mu F^{\mu 0} = \partial_0 F^{00} + \partial_i F^{i0} = \partial_i F^{i0} = \partial_i E^i = \nabla \cdot \boldsymbol{E}. \tag{6.3.7}$$

Here we have used the fact that $F^{00} = 0$ and $F^{i0} = E^i$. Combining these results we have

$$\nabla \cdot \boldsymbol{E} = \rho, \tag{6.3.8}$$

which is Gauss' law.

This time consider $\nu = i$. We then have

$$\partial_i F^{\mu i} = \partial_0 F^{0i} + \partial_j F^{ji} \tag{6.3.9}$$

$$= -\partial_0 F^{i0} - \partial_j F^{ij} \tag{6.3.10}$$

$$= -\partial_0 E^i + \partial_j \varepsilon^{ijk} B^k \tag{6.3.11}$$

$$= -\frac{1}{c}\frac{\partial E^i}{\partial t} + (\nabla \times \boldsymbol{B})^i. \tag{6.3.12}$$

Here we have used antisymmetry: $F^{\mu\nu} = -F^{\nu\mu}$, and identified the components $F^{i0} = E^i$ and $F^{ij} = -\varepsilon^{ijk}B^k$. We also have $\partial_i F^{\mu i} = J^i/c$. Combining these results and rearranging for $\nabla \times \boldsymbol{B}$ we get

$$\nabla \times \boldsymbol{B} = \frac{1}{c}\boldsymbol{J} + \frac{1}{c}\frac{\partial \boldsymbol{E}}{\partial t}. \tag{6.3.13}$$

This is Ampère's law.

6.4 Homogeneous Maxwell Equations

When we introduced the potentials we did so as the solutions to the two homogeneous (source free) Maxwell equations, $\nabla \cdot \boldsymbol{B} = 0$ and $\nabla \times \boldsymbol{E} + c^{-1}\partial\boldsymbol{B}/\partial t = \boldsymbol{0}$. We therefore look for an equation for which $F^{\mu\nu}$ is a solution causing the equation to vanish identically.

If we consider the term $\partial_\mu F^{\mu\nu}$ appearing in the previous section to be the divergence then the fact that both divergences and curls appear in Maxwell's equations suggests that we should look for some curl-like term. Since the curl has components $\varepsilon^{ijk}\partial_j V^k$ we will try something similar, say $\varepsilon^{\mu\nu\alpha\beta}\partial_\mu F_{\alpha\beta}$. The question is, does this vanish?

Before we answer this question we introduce a new quantity, called the **dual field strength tensor**, and defined by

$$F^*_{\mu\nu} := -\frac{1}{2}\varepsilon_{\mu\nu\alpha\beta}F^{\alpha\beta}. \tag{6.4.1}$$

> **Notation 6.4.2** There is not a consistent notation for this quantity. Some sources call it $F^*_{\mu\nu}$, like we have, others use $G_{\mu\nu}$, or $\tilde{F}_{\mu\nu}$.
> Note that the star is nothing to do with complex conjugates, the field strength tensor is a real quantity. Instead we can think of the star as relating to the Hodge star operator, since in the language of differential forms the potential is a one-form, A, the field strength is a two-form, $F = \mathrm{d}A$, and the dual field strength is its Hodge dual, $\star F$.

We can easily derive the components of the dual field tensor in terms of the

potential:

$$F^*_{\mu\nu} = -\frac{1}{2}\varepsilon_{\mu\nu\alpha\beta}F^{\alpha\beta} \tag{6.4.3}$$

$$= -\frac{1}{2}\varepsilon_{\mu\nu\alpha\beta}(\partial^\alpha A^\beta - \partial^\beta A^\alpha) \tag{6.4.4}$$

$$= -\frac{1}{2}\varepsilon_{\mu\nu\alpha\beta}\partial^\alpha A^\beta + \frac{1}{2}\varepsilon_{\mu\nu\alpha\beta}\partial^\beta A^\alpha \tag{6.4.5}$$

$$= -\frac{1}{2}\varepsilon_{\mu\nu\alpha\beta}\partial^\alpha A^\beta - \frac{1}{2}\varepsilon_{\mu\nu\beta\alpha}\partial^\beta A^\alpha \tag{6.4.6}$$

$$= -\varepsilon_{\mu\nu\alpha\beta}\partial^\alpha A^\beta \tag{6.4.7}$$

where in the last step we relabel indices in the second term, $\alpha \leftrightarrow \beta$, and in the step before that we used the antisymmetry of the Levi-Civita symbol.

Now consider the divergence of the dual field tensor:

$$\partial_\mu F^{*\mu\nu} = -\frac{1}{2}\partial_\mu\varepsilon^{\mu\nu\alpha\beta}F_{\alpha\beta} \tag{6.4.8}$$

$$= -\frac{1}{2}\varepsilon^{\mu\nu\alpha\beta}\partial_\mu F_{\alpha\beta} \tag{6.4.9}$$

$$= -\frac{1}{2}\varepsilon^{\mu\nu\alpha\beta}\partial_\mu(\partial_\alpha A_\beta - \partial_\beta A_\alpha) \tag{6.4.10}$$

$$= -\frac{1}{2}\varepsilon^{\mu\nu\alpha\beta}\partial_\mu\partial_\alpha A_\beta + \frac{1}{2}\varepsilon^{\mu\nu\alpha\beta}\partial_\mu\partial_\beta A_\alpha \tag{6.4.11}$$

$$= 0. \tag{6.4.12}$$

This last equality holds since the partial derivatives are symmetric and the Levi-Civita symbol is antisymmetric, and so their product vanishes. Hence,

$$\partial_\mu F^{*\mu\nu} = -\frac{1}{2}\varepsilon^{\mu\nu\alpha\beta}\partial_\mu F_{\alpha\beta} = 0. \tag{6.4.13}$$

We can fairly easily work out what the components of F^* are in terms of the electric and magnetic fields. First, consider the case of F^{*i0}:

$$F^{*i0} = -\frac{1}{2}\varepsilon^{i0\alpha\beta}F_{\alpha\beta}. \tag{6.4.14}$$

The non-vanishing terms of this cannot have $\alpha = 0$ or $\beta = 0$, so we have that α and β are spatial indices, and so

$$F^{*i0} = -\frac{1}{2}\varepsilon^{i0jk}F_{jk}. \tag{6.4.15}$$

now consider the Levi-Civita symbol. Swapping the first two indices gives us a negative sign,

$$F^{*i0} = \frac{1}{2}\varepsilon^{0ijk}F_{jk}. \tag{6.4.16}$$

Suppose $ijk = 123$. Then we have $\varepsilon^{0123} = -1$, following our convention of $\varepsilon_{0123} := 1$. This means that $\varepsilon^{0123} = -\varepsilon^{123}$. The Levi-Civita symbol is entirely defined by a single value and requiring that it is totally antisymmetric. This allows

us to conclude from this that $\varepsilon^{0ijk} = -\varepsilon^{ijk}$. Hence,

$$F^{*0i} = -\frac{1}{2}\varepsilon^{ijk}F_{jk} \tag{6.4.17}$$

$$= \frac{1}{2}\varepsilon^{ijk}\varepsilon^{jkl}B^l, \tag{6.4.18}$$

where we have used the fact that $F^{jk} = -\varepsilon^{jkl}B^l$. We now have two options, we can apply an identity for the product of two Levi-Civita symbols, first using the antisymmetry to get the indices to line up with the usual form for the identity:

$$F^{*0i} = -\frac{1}{2}\varepsilon^{ijk}\varepsilon^{kjl}B^l \tag{6.4.19}$$

$$= -\frac{1}{2}(\delta^{ij}\delta^{jl} - \delta^{il}\delta^{jj})B^l \tag{6.4.20}$$

$$= -\frac{1}{2}(\delta^{il} - 3\delta^{il})B^l \tag{6.4.21}$$

$$= \delta^{il}B^l \tag{6.4.22}$$

$$= B^l. \tag{6.4.23}$$

We could also have gotten to the same result by considering $\varepsilon^{ijk}\varepsilon^{jkl}$ with $i = 1$, which means that for a non-vanishing term we must have either $j = 2$ and $k = 3$, or $j = 3$ and $k = 2$. Either way a non-vanishing result then requires that $l = 1$. Doing the same with $i = 2, 3$ we see that $i = l$ and the values of j and k aren't important for non-vanishing terms and so $\varepsilon^{ijk}\varepsilon^{jkl} = 2\delta^{il}$.

now consider the components F^{*ij}:

$$F^{*ij} = -\frac{1}{2}\varepsilon^{ij\alpha\beta}F_{\alpha\beta}. \tag{6.4.24}$$

In order for this not to vanish either $\alpha = 0$ or $\beta = 0$. Suppose $\beta = 0$, we can then swap α and β in both $\varepsilon^{ij\alpha\beta}$ and $F_{\alpha\beta}$. Each swap gives a minus sign, and there are two swaps, so there is no change. We can then relabel $\alpha \leftrightarrow \beta$. Therefore we have two non-zero terms, one with $\alpha = 0$ and one with $\beta = 0$, but with the previously mentioned index reshuffling these two terms are actually the same, so we just cancel the factor of 1/2. This gives

$$F^{*ij} = -\varepsilon^{ij0\beta}F_{0\beta}. \tag{6.4.25}$$

In order for this not to vanish $\beta \neq 0$, and so β can be replaced with a spatial index:

$$F^{*ij} = -\varepsilon^{ij0k}F_{0k}. \tag{6.4.26}$$

Swapping 0 and j on the Levi-Civita gives a minus sign, which then cancels if we further swap 0 and i, that is $\varepsilon^{ij0k} = -\varepsilon^{i0jk} = \varepsilon^{0ijk} = -\varepsilon^{ijk}$, with the last equality following from the previous logic about four and three index Levi-Civita symbols. Hence,

$$F^{*ij} = \varepsilon^{ijk}F_{0k}. \tag{6.4.27}$$

Now raise the two indices on F_{0k}. We can raise the 0 for free, and raiding the k introduces a minus sign:

$$F^{*ij} = -\varepsilon^{ijk}F^{0k}. \tag{6.4.28}$$

Now, we know $F^{k0} = E^k$, and $F^{0k} = -F^{k0}$, so

$$F^{*ij} = \varepsilon^{ijk} E^k. \tag{6.4.29}$$

Comparing the components of the two field tensors we have

$$F^{*i0} = B^i, \qquad F^{*ij} = \varepsilon^{ijk} E^k, \tag{6.4.30}$$
$$F^{i0} = E^i, \qquad F^{*ij} = -\varepsilon^{ijk} B^k. \tag{6.4.31}$$

Notice that we can obtain the dual field tensor from the field tensor, $F \to F^*$, via $\boldsymbol{E} \to \boldsymbol{B}$ and $\boldsymbol{B} \to -\boldsymbol{E}$. This is called **electromagnetic duality**. We can view it as a result of Maxwell's equations in a vacuum not distinguishing between the electric and magnetic fields, up to a sign, since we have

$$\nabla \cdot \boldsymbol{E} = 0, \qquad \nabla \cdot \boldsymbol{B} = 0, \tag{6.4.32}$$
$$\nabla \times \boldsymbol{E} = -\frac{1}{c}\frac{\partial \boldsymbol{B}}{\partial t}, \qquad \nabla \times \boldsymbol{B} = \frac{1}{c}\frac{\partial \boldsymbol{E}}{\partial t}. \tag{6.4.33}$$

> **Application 6.4.34 — Complex Electromagnetic Field** One way to make the electromagnetic duality more obvious is to define a complex electromagnetic field
>
> $$\mathcal{E} := \boldsymbol{E} + i\boldsymbol{B}. \tag{6.4.35}$$
>
> Then the vacuum Maxwell equations can be written compactly as two equations:
>
> $$\nabla \cdot \mathcal{E} = 0, \qquad \text{and} \qquad \nabla \times \mathcal{E} = \frac{i}{c}\frac{\partial \mathcal{E}}{\partial t}. \tag{6.4.36}$$
>
> Electromagnetic duality is then expressed by multiplying this quantity by $-i$:
>
> $$-i\mathcal{E} = \boldsymbol{B} - i\boldsymbol{E}. \tag{6.4.37}$$
>
> This is the same result as replacing $\boldsymbol{E}$ with $\boldsymbol{B}$ and $\boldsymbol{B}$ with $-\boldsymbol{E}$.

We can now finally recover the homogenous Maxwell equations. To do so we consider the divergence of the dual field tensor, which is known to vanish, $\partial_\mu F^{*\mu\nu} = 0$. First, consider the case of $\nu = 0$:

$$0 = \partial_\mu F^{*\mu 0} \tag{6.4.38}$$
$$= \partial_0 F^{*00} + \partial_i F^{*i0} \tag{6.4.39}$$
$$= \partial_i F^{*i0} \tag{6.4.40}$$
$$= \partial_i B^i \tag{6.4.41}$$
$$= \nabla \cdot \boldsymbol{B}. \tag{6.4.42}$$

Second, consider the case of $\nu = i$:

$$0 = \partial_\mu F^{*\mu i} \tag{6.4.43}$$

$$= \partial_0 F^{*0i} + \partial_j F^{*ji} \tag{6.4.44}$$

$$= -\partial_0 F^{*i0} - \partial_j F^{*ij} \tag{6.4.45}$$

$$= -\partial_0 B^i - \partial_j \varepsilon^{ijk} E^k \tag{6.4.46}$$

$$= -\frac{1}{c}\frac{\partial B^i}{\partial t} - (\nabla \times \boldsymbol{E})^i, \tag{6.4.47}$$

which gives Faraday's law:

$$\nabla \times \boldsymbol{E} = -\frac{1}{c}\frac{\partial \boldsymbol{B}}{\partial t}. \tag{6.4.48}$$

To summarise, the electromagnetic field tensor allows us to write the four Maxwell equations in vector calculus notation as two equations in relativistic notation:

$$\partial_\mu F^{\mu\nu} = \frac{1}{c}J^\nu \qquad \begin{cases} \nabla \cdot \boldsymbol{E} = \rho, \\ \nabla \times \boldsymbol{B} = \frac{1}{c}\boldsymbol{J} + \frac{1}{c}\frac{\partial \boldsymbol{E}}{\partial t}, \end{cases} \tag{6.4.49}$$

$$\partial_\mu F^{*\mu\nu} = 0 \qquad \begin{cases} \nabla \cdot \boldsymbol{B} = 0, \\ \nabla \times \boldsymbol{E} = -\frac{1}{c}\frac{\partial \boldsymbol{B}}{\partial t}. \end{cases} \tag{6.4.50}$$

Notice that in the absence of sources electromagnetic duality is more obvious:

$$\partial_\mu F^{\mu\nu} = 0 = \partial_\mu F^{*\mu\nu}. \tag{6.4.51}$$

6.5 Lorentz Transform of the Field Tensor

The electric and magnetic field are *not* Lorentz vectors, this is because they are really just a neat packaging of components of the electromagnetic field tensor, appropriate for non-relativistic electrodynamics. The field tensor is a rank 2 contravariant tensor, and so transforms as

$$F^{\mu\nu} \to F'^{\mu\nu} = \Lambda^\mu{}_\alpha \Lambda^\nu{}_\beta F^{\alpha\beta}. \tag{6.5.1}$$

Consider a standard Lorentz transformation

$$\Lambda^\mu{}_\nu = \begin{pmatrix} \gamma & -\gamma\beta & 0 & 0 \\ -\gamma\beta & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{6.5.2}$$

Consider how the component $E^1 = F^{10}$ transforms under this Lorentz transformation, in particular we can use the fact that the zeroth and first rows of the standard Lorentz transform are non-zero only in the zeroth and first column. Hence,

$$E^1 = F^{10} \to F'^{10} = \Lambda^1{}_\alpha \Lambda^0{}_\beta F^{\alpha\beta} \tag{6.5.3}$$

$$= \Lambda^1{}_0 \Lambda^0{}_0 F^{00} + \Lambda^1{}_0 \Lambda^0{}_1 F^{01} + \Lambda^1{}_1 \Lambda^0{}_0 F^{10} + \Lambda^1{}_1 \Lambda^0{}_1 F^{11}.$$

We can now use the fact that $F^{00} = F^{11} = 0$ to further simplify this giving

$$E^1 = F^{10} \to F'^{10} = \Lambda^1{}_0 \Lambda^0{}_1 F^{01} + \Lambda^1{}_1 \Lambda^0{}_0 F^{10} \tag{6.5.4}$$
$$= -\Lambda^1{}_0 \Lambda^0{}_1 F^{10} + \Lambda^1{}_1 \Lambda^0{}_0 F^{10} \tag{6.5.5}$$
$$= \gamma^2(1 - \beta^2) F^{10} \tag{6.5.6}$$
$$= E^1, \tag{6.5.7}$$

since $\gamma^2 = 1/(1 - \beta^2)$.

Now consider $E^2 = F^{20}$, the transformation coefficient $\Lambda^2{}_\alpha$ is only non-zero if $\alpha = 2$, in which case $\Lambda^2{}_2 = 1$ so we have

$$E^2 = F^{20} \to F'^{20} = \Lambda^2{}_\alpha \Lambda^0{}_\beta F^{\alpha\beta} \tag{6.5.8}$$
$$= \Lambda^2{}_2 \Lambda^0{}_\beta F^{2\beta} \tag{6.5.9}$$
$$= \Lambda^0{}_\beta F^{2\beta} \tag{6.5.10}$$
$$= \Lambda^0{}_0 F^{20} + \Lambda^0{}_1 F^{21} \tag{6.5.11}$$
$$= \Lambda^0{}_0 E^2 - \Lambda^0{}_1 B^3 \tag{6.5.12}$$
$$= \gamma E^2 - \gamma\beta B^3. \tag{6.5.13}$$

Similarly

$$E^3 = F^{03} \to F'^{03} = \gamma(E^3 + \beta B^3). \tag{6.5.14}$$

We can derive the transformation of the magnetic field in the same way, but it is easier to use electromagnetic duality, replacing $E^i \to B^i$ and $B^i \to -E^i$, giving

$$B^1 \to B^1, \tag{6.5.15}$$
$$B^2 \to \gamma(B^2 + \beta E^3), \tag{6.5.16}$$
$$B^3 \to \gamma(B^3 - \beta E^2). \tag{6.5.17}$$

6.5.1 Point Charge

Consider a point charge, q. We know that in the rest frame of this charge, S', the electric field is

$$\boldsymbol{E}'(\boldsymbol{r}') = \frac{q}{4\pi r'^3}\boldsymbol{r}' \tag{6.5.18}$$

and the magnetic field is $\boldsymbol{B}'(\boldsymbol{r}') = \boldsymbol{0}$. Suppose that in a frame stationary with respect to some observer this particle has velocity $\boldsymbol{v}$. We want to calculate the values that the observer would measure for the electric and magnetic fields. We are free to set up the frames such that $\boldsymbol{v} = v\hat{\boldsymbol{x}}$ and at time $t = 0$ we have the origins of the two frames coincide, this allows us to use standard Lorentz transforms, with the slight change that if the particle has velocity $\boldsymbol{v}$, then the S frame has velocity $-\boldsymbol{v}$ in frame S', and so we replace β with $-\beta$ in our transformations. For simplicity we only consider the fields at time $t = 0$.

There are two steps required to calculate the fields the observer measures:

1. First, we must Lorentz transform the fields, so $\boldsymbol{E}'(\boldsymbol{r}') \to \boldsymbol{E}(\boldsymbol{r}')$ and $\boldsymbol{B}'(\boldsymbol{r}') \to \boldsymbol{B}(\boldsymbol{r}')$.

2. Second, we must perform a coordinate transformation to relate positions in the particles rest frame, S', to positions in the frame at rest relative to the observer, S, so $\boldsymbol{r}' \to \boldsymbol{r}$.

First we take the Lorentz transform of the fields. We'll start with the electric field and come back to the magnetic field later. We have already seen how the electric field transforms under a standard Lorentz transformation and so we can use this, we just need to be careful to swap β for $-\beta$. This gives the components

$$E^1 = E'^1, \tag{6.5.19}$$
$$E^2 = \gamma(E'^2 + \beta B'^3) = \gamma E'^2, \tag{6.5.20}$$
$$E^3 = \gamma(E'^3 - \beta B'^2) = \gamma E'^3. \tag{6.5.21}$$

Here we have used the fact that $B'^i = 0$ for all i. From this we can fairly easily see that the electric field at position $\boldsymbol{r}'$, as measured in the rest frame of the particle, is

$$\boldsymbol{E}(\boldsymbol{r}') = \frac{q}{4\pi r'^3}(x', \gamma y', \gamma z'). \tag{6.5.22}$$

We now need to make a coordinate transformation to express this same value in un-primed coordinates. This is simple enough for a standard Lorentz transformation, using the fact that we are choosing to have $t = 0$, we simply have $x' = \gamma x$, $y' = y$, and $z' = z$. Now consider the value r'^2:

$$r'^2 = x'^2 + y'^2 + z'^2 \tag{6.5.23}$$
$$= \gamma^2 x^2 + y^2 + z^2 \tag{6.5.24}$$
$$= \gamma^2(x^2 + y^2 + z^2) + (1 - \gamma^2)(y^2 + z^2). \tag{6.5.25}$$

To see why this last equality holds just expand it out and it quickly reduces to the previous line. We can now use

$$1 - \gamma^2 = 1 - \frac{1}{1 - \beta^2} = \frac{1 - \beta^2}{1 - \beta^2} - \frac{1}{1 - \beta^2} = \frac{1 - \beta^2 - 1}{1 - \beta^2} = -\frac{\beta^2}{1 - \beta^2} = -\beta^2 \gamma^2. \tag{6.5.26}$$

This gives

$$r'^2 = \gamma^2(x^2 + y^2 + z^2) - \gamma^2 \beta^2(y^2 + z^2) \tag{6.5.27}$$
$$= \gamma^2 r^2 - \gamma^2 \beta^2(y^2 + z^2) \tag{6.5.28}$$
$$= \gamma^2 r^2 \left(1 - \beta^2 \frac{y^2 + z^2}{r^2}\right). \tag{6.5.29}$$

Now consider the angle ϑ defined to be the angle between the position vector, $\boldsymbol{r}$, and the x-axis. This is shown in Figure 6.1. We can see that the side opposite ϑ is $\sqrt{y^2 + z^2}$, and the hypotenuse is r^2. Hence,

$$\sin^2 \vartheta = \frac{y^2 + z^2}{r^2}. \tag{6.5.30}$$

This allows us to write r'^2 in terms of ϑ:

$$r'^2 = \gamma^2 r^2(1 - \beta^2 \sin^2 \vartheta). \tag{6.5.31}$$
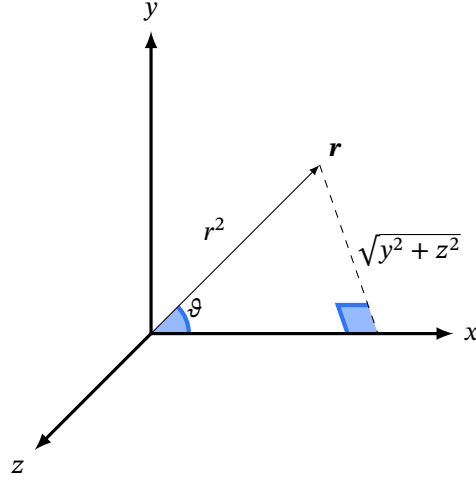
Figure 6.1: Coordinates for describing a point charge.

For the electric field we actually need the value $r'^3 = (r'^2)^{3/2}$:

$$r'^3 = \gamma^3 r^3 (1 - \beta^2 \sin^2 \vartheta)^{3/2}. \tag{6.5.32}$$

The electric field as measured by the observer at position $\boldsymbol{r}$ in their frame is

$$\boldsymbol{E(r)} = \frac{q}{4\pi} \frac{1}{\gamma^3 r^3 (1 - \beta^2 \sin^2 \vartheta)^{3/2}} (\gamma x, \gamma y, \gamma z) = \frac{q}{4\pi} \frac{1}{\gamma^2 r^3 (1 - \beta^2 \sin^2 \vartheta)} \boldsymbol{r}. \tag{6.5.33}$$

Note that the factors of γ for x comes from the coordinate transformation and the factors of γ for y and z come from transforming the field.

Notice that the electric field is still radial, meaning along the $\boldsymbol{r}$ direction. However, it is no longer spherically symmetric. For example, if $\vartheta = 0$ then

$$\boldsymbol{E(r)} = \frac{1}{\gamma^2} \frac{q}{4\pi r^3} \boldsymbol{r}, \tag{6.5.34}$$

which, for $\boldsymbol{v} \neq \boldsymbol{0}$, is lower than the static case since $\gamma \geq 1$. On the other hand, for $\vartheta = \pi/2$ we get a factor of

$$\gamma^2 (1 - \beta^2)^{3/2} = \gamma^2 \gamma^{-3} = \gamma^{-1}. \tag{6.5.35}$$

So, the electric field is

$$\boldsymbol{E(r)} = \gamma \frac{q}{4\pi r^3} \boldsymbol{r}, \tag{6.5.36}$$

which is generally greater than in the static case. The electric field is shown in Figure 6.2.

We can also compute the magnetic field in the observers frame, since in their frame the charge is moving we expect that the magnetic field will not vanish. To do so we use the fact that $\boldsymbol{B'} = \boldsymbol{0}$. This allows us to use the relation

$$\boldsymbol{B'} = -\boldsymbol{\beta} \times \boldsymbol{E'}, \tag{6.5.37}$$

Figure 6.2: The electric field of a moving point charge. The velocity of the point charge increases to the right, and the velocity is to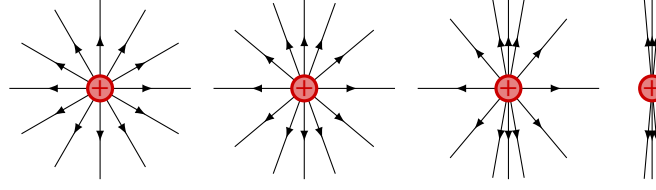 the right. The left most point charge is stationary. As velocity increases the field becomes stronger above and below the charge and weaker in front and behind. In the extreme case of $v = c$ we get a shock wave, much like a sonic boom.

where $\boldsymbol{\beta} = \boldsymbol{v}/c$. This relation holds for a standard Lorentz transformation whenever $\boldsymbol{B} = \boldsymbol{0}$. This is easily shown for a standard Lorentz transformation, simply compute the cross product and show that we recover the known transformation laws for the components of $\boldsymbol{B}$. In our case we are actually transforming from the primed frame to the un-primed frame, and we have $-\boldsymbol{\beta}$, instead of $\boldsymbol{\beta}$, so

$$\boldsymbol{B} = \boldsymbol{\beta} \times \boldsymbol{E}. \tag{6.5.38}$$

We know that $\boldsymbol{E} \propto \boldsymbol{r}$, and $\boldsymbol{\beta} \propto \boldsymbol{v}$, and so in full we have

$$\boldsymbol{B(r)} = \boldsymbol{\beta} \times \boldsymbol{E} = \frac{q}{4\pi r^2} \frac{1}{\gamma^2(1 - \beta^2 \sin^2 \vartheta)^{3/2}} \frac{1}{c} \boldsymbol{v} \times \boldsymbol{r}. \tag{6.5.39}$$

6.6 Field Invariants

Given a field strength tensor and its dual there are two invariants we can form, $F^{\mu\nu}F_{\mu\nu}$ and $F^{\mu\nu}F^*_{\mu\nu}$. All other combinations, such as $F^{*\mu\nu}F_{*\mu\nu}$, can be written in terms of these two invariants. We can evaluate these two invariants fairly easily in terms of the electric and magnetic fields. First, consider $F^{\mu\nu}F_{\mu\nu}$:

$$F^{\mu\nu}F_{\mu\nu} = F^{i0}F_{i0} + F^{0i}F_{0i} + F^{ij}F_{ij} \tag{6.6.1}$$

$$= 2F^{i0}F_{i0} + F^{ij}F_{ij} \tag{6.6.2}$$

$$= -2F^{i0}F^{i0} + F^{ij}F^{ij} \tag{6.6.3}$$

$$= -2E^iE^i + \varepsilon^{ijk}B^k\varepsilon^{ijl}B^l \tag{6.6.4}$$

$$= -2\boldsymbol{E}^2 + 2\delta^{kl}B^kB^l \tag{6.6.5}$$

$$= -2\boldsymbol{E}^2 + 2\boldsymbol{B}^2. \tag{6.6.6}$$

Here we have used antisymmetry, $F^{0i}F_{0i} = -F^{i0}F_{0i} = F^{i0}F_{i0}$, raising spatial indices at the cost of a minus sign, $F^{i0}F_{i0} = -F^{i0}F^i_{\ 0} = -F^{i0}F^{i0}$ and $F^{ij}F_{ij} = -F^{ij}F^i_{\ j} = F^{ij}F^{ij}$, the relations $F^{i0} = E^i$ and $F^{ij} = -\varepsilon^{ijk}B^k$, and the identity $\varepsilon^{ijk}\varepsilon^{ijl} = 2\delta^{kl}$.

Similarly we find that

$$F^{\mu\nu}F^*_{\mu\nu} = F^{i0}F^*_{i0} + F^{0i}F^*_{0i} + F^{ij}F^*_{ij} \tag{6.6.7}$$

$$= 2F^{i0}F^*_{i0} + F^{ij}F^*_{ij} \tag{6.6.8}$$

$$= -2F^{i0}F^{*i0} + F^{ij}F^{*ij} \tag{6.6.9}$$

$$= -2E^iB^i - \varepsilon^{ijk}B^k\varepsilon^{ijl}E^l \tag{6.6.10}$$

$$= -2\boldsymbol{E}\cdot\boldsymbol{B} - 2\delta^{kl}B^kE^l \tag{6.6.11}$$

$$= -4\boldsymbol{E}\cdot\boldsymbol{B}. \tag{6.6.12}$$

These invariants can be used to shortcut having to do full calculations. For example, if $\boldsymbol{E}\cdot\boldsymbol{B} = 0$ in some frame, as is the case for electromagnetic waves, then $\boldsymbol{E}\cdot\boldsymbol{B} = 0$ in all frames. If $\boldsymbol{E}^2 = \boldsymbol{B}^2$ in some frame then $\boldsymbol{E}^2 = \boldsymbol{B}^2$ in all frames, again this is true for electromagnetic waves. Further, if $E < B$ in some frame then $E > B$ in all frames, and vice versa.

We can also use these invariants to guide which transformations we might consider. For example, often we want to make a transformation such that one field vanishes in the new frame. If $F^{\mu\nu}F_{\mu\nu} > 0$ then this implies that $B < E$, and so we might look for a frame such that $\boldsymbol{B}$ vanishes. On the other hand if $F^{\mu\nu}F_{\mu\nu} < 0$ then $E < B$ and so we look for a frame in which $\boldsymbol{E}$ vanishes.

6.7 Binachi Identity

The **Binachi identity**, is

$$\partial_\rho F_{\mu\nu} + \partial_\mu F_{\nu\rho} + \partial_\nu F_{\rho\mu} = 0. \tag{6.7.1}$$

This is easily shown by expanding $F_{\mu\nu}$ in terms of potentials and using the symmetry of partial derivatives:

$$\partial_\rho F_{\mu\nu} + \partial_\mu F_{\nu\rho} + \partial_\nu F_{\rho\mu} \tag{6.7.2}$$

$$= \partial_\rho(\partial_\mu A_\nu - \partial_\nu A_\mu) + \partial_\mu(\partial_\nu A_\rho - \partial_\rho A_\nu) + \partial_\nu(\partial_\rho A_\mu - \partial_\mu A_\rho) \tag{6.7.3}$$

$$= \partial_\rho\partial_\mu A_\nu - \partial_\rho\partial_\nu A_\mu + \partial_\mu\partial_\nu A_\rho - \partial_\mu\partial_\rho A_\nu + \partial_\nu\partial_\rho A_\mu - \partial_\nu\partial_\mu A_\rho \tag{6.7.4}$$

$$= 0. \tag{6.7.5}$$

If we consider the Binachi identity multiplied by $\varepsilon^{\mu\nu\rho\sigma}$ we find that this implies $\partial_\mu F^{*ms\nu} = 0$, which is what we started looking for to define $F^{*\mu\nu}$ in the first place, so we can replace this condition with the Binachi identity if we wish.

> **Application 6.7.6 — Binachi Identity** In general relativity there is a similar identity also known as the Binachi identity for the Riemann curvature tensor:
>
> $$\nabla_\varepsilon R^\alpha{}_{\beta\gamma\delta} + \nabla_\gamma R^\alpha{}_{\beta\delta\varepsilon} + \nabla_\delta R^\alpha{}_{\beta\varepsilon\gamma} = 0. \tag{6.7.7}$$
>
> Here ∇_μ is a covariant derivative, which reduces to the partial derivative ∂_μ in a locally inertial frame. The Riemann tensor is defined by
>
> $$R^\alpha{}_{\beta\gamma\delta} = \partial_\gamma\Gamma^\alpha{}_{\delta\beta} - \partial_\delta\Gamma^\alpha{}_{\gamma\beta} + \Gamma^\alpha{}_{\gamma\mu}\Gamma^\mu{}_{\beta\delta} - \Gamma^\alpha{}_{\delta\mu}\Gamma^\mu{}_{\beta\gamma}. \tag{6.7.8}$$

Here $\Gamma^{\alpha}{}_{\beta\gamma}$ are the affine connections. In a locally inertial frame this reduces to

$$R^{\alpha}{}_{\beta\gamma\delta} = \partial_{\gamma}\Gamma^{\alpha}{}_{\delta\beta} - \partial_{\delta}\Gamma^{\alpha}{}_{\gamma\beta}. \tag{6.7.9}$$

Notice then that if we replace $R^{\alpha}{}_{\beta\gamma\delta} \rightarrow F_{\alpha\beta}$ and $\Gamma^{\mu}{}_{\alpha\sigma} \rightarrow A_{\alpha}$ we get the definition of the field strength tensor.

These two Binachi identities are related by the fact that we can express both the Riemann tensor and the field strength tensor as exterior derivatives. For the field strength tensor we have $F = \mathrm{d}A$, and then we can use the identity $\mathrm{d}^2 = 0$ giving $\mathrm{d}F = \mathrm{d}^2 A = 0$. For the Riemann tensor the Riemann tensor can be expressed as the exterior derivative of the covariant derivative of a vector field.

Seven

Energy and Momentum

7.1 Energy–Momentum Tensor

The **energy–momentum tensor**, also called the **stress–energy tensor** is a central object of study in electrodynamics, as well as in wider relativity, including general relativity, where it appears as the source term in Einstein's field equations. For the purposes of this course we will simply define it and then probe its properties. One can also derive it as a logical quantity to study based on a few assumptions, or it is possible to derive it entirely from symmetry arguments.

> **Definition 7.1.1 — Energy–Momentum Tensor** The energy–momentum tensor, T, is a rank 2 Lorentz tensor with components
>
> $$T_{\mu\nu} := F_\mu{}^\alpha F_{\alpha\nu} + \frac{1}{4}\eta_{\mu\nu}F^{\alpha\beta}F_{\alpha\beta}. \tag{7.1.2}$$

We defined the energy–momentum tensor here in terms of covariant components, but we could equally define it in terms of the contravariant components

$$T^{\mu\nu} = F^{\mu\alpha}F_\alpha{}^\nu + \frac{1}{4}\eta^{\mu\nu}F^{\alpha\beta}F_{\alpha\beta}. \tag{7.1.3}$$

Now that we have defined the energy–momentum tensor we will investigate some of its properties.

7.1.1 Symmetry

The energy–momentum tensor is symmetric, that is $T^{\mu\nu} = T^{\nu\mu}$. Clearly the second term is symmetric, since the only appearance of μ and ν is in the metric, $\eta^{\mu\nu}$, which is symmetric. For the first term we have

$$F_\mu{}^\alpha F_{\alpha\nu} \rightarrow F_\nu{}^\alpha F_{\alpha\mu} = F_{\alpha\mu}F_\nu{}^\alpha = -F_{\mu\alpha}F_\nu{}^\alpha = -F_\mu{}^\alpha F_{\nu\alpha} = F_\mu{}^\alpha F_{\alpha\nu}. \tag{7.1.4}$$

Here we have used the antisymmetry of the field tensor, $F_{\mu\nu} = -F_{\nu\mu}$, as well as the freedom we have to raise an summed index, here α, as long as we lower its pair. So the first term is also symmetric, and therefore the energy–momentum tensor is symmetric.

7.1.2 Trace

The energy–momentum tensor is traceless, that is $T^\mu{}_\mu = 0$. This is simple enough to show, first note that we can write the energy–momentum tensor in a mixed form as

$$T^\mu{}_\nu = F^{\mu\alpha}F_{\alpha\nu} + \frac{1}{4}\eta^\mu{}_\nu F^{\alpha\beta}F_{\alpha\beta}\,. \tag{7.1.5}$$

Recall that $\eta^\mu{}_\nu = \delta^\mu{}_\nu$, and that $\delta^\mu{}_\mu = 4$ in 4 dimensions. Hence, the trace of the energy–momentum tensor is

$$T^\mu{}_\mu = F^{\mu\alpha}F_{\alpha\mu} + \frac{1}{4}\delta^\mu{}_\mu F^{\alpha\beta}F_{\alpha\beta} \tag{7.1.6}$$

$$= F^{\mu\alpha}F_{\alpha\mu} + \frac{1}{4}\cdot 4F^{\alpha\beta}F_{\alpha\beta} \tag{7.1.7}$$

$$= F^{\mu\alpha}F_{\alpha\mu} + F^{\alpha\beta}F_{\alpha\beta} \tag{7.1.8}$$

$$= -F^{\mu\alpha}F_{\mu\alpha} + F^{\alpha\beta}F_{\alpha\beta} \tag{7.1.9}$$

$$= 0. \tag{7.1.10}$$

In the last step we can simply relabel $\alpha \to \mu$ and $\beta \to \alpha$ in the last term and we see it is the same as the first but with the opposite sign.

7.1.3 Components

We can compute the components of the energy–momentum tensor in terms of the electric and magnetic fields. This will allow us to prescribe interpretations to these terms. We will start with the time-time terms, so $\mu = \nu = 0$. We have that $\eta^{00} = 1$, and so

$$T^{00} = F^{0\alpha}F_\alpha{}^0 + \frac{1}{4}F^{\alpha\beta}F_{\alpha\beta} \tag{7.1.11}$$

$$= F^{00}F_0{}^0 + F^{0i}F_i{}^0 - \frac{1}{2}(\boldsymbol{E}^2 - \boldsymbol{B}^2) \tag{7.1.12}$$

$$= -F^{0i}F^{i0} - \frac{1}{2}(\boldsymbol{E}^2 - \boldsymbol{B}^2) \tag{7.1.13}$$

$$= F^{i0}F^{i0} - \frac{1}{2}(\boldsymbol{E}^2 - \boldsymbol{B}^2) \tag{7.1.14}$$

$$= E^i E^i - \frac{1}{2}(\boldsymbol{E}^2 - \boldsymbol{B}^2) \tag{7.1.15}$$

$$= \frac{1}{2}(\boldsymbol{E}^2 + \boldsymbol{B}^2) \tag{7.1.16}$$

$$= u. \tag{7.1.17}$$

Here u is the energy density of the electromagnetic field. We have used here the antisymmetry of the electromagnetic field tensor, which implies $F^{00} = 0$, the fact that raising a spatial index gives a minus sign, the component $F^{i0} = E^i$, and the field tensor invariant

$$F^{\mu\nu}F_{\mu\nu} = 2(\boldsymbol{B}^2 - \boldsymbol{E}^2). \tag{7.1.18}$$

From this we can see that the interpretation of the time-time component of the energy–momentum tensor is as the energy density of the electromagnetic field.

Next consider the time-space component, so $\nu = 0$ and $\mu = i$, or equivalently $\mu = 0$ and $\nu = i$ since the energy–momentum tensor is symmetric. We have

$$T^{i0} = F^{i\alpha}F_{\alpha}{}^{0} + \frac{1}{4}\eta^{i0}F^{\alpha\beta}F_{\alpha\beta} \tag{7.1.19}$$

$$= F^{i0}F_{0}{}^{0} + F^{ij}F_{j}{}^{0} \tag{7.1.20}$$

$$= F^{i0}F^{00} - F^{ij}F^{j0} \tag{7.1.21}$$

$$= \varepsilon^{ijk}B^{k}E^{j} \tag{7.1.22}$$

$$= (\boldsymbol{E} \times \boldsymbol{B})^{i} \tag{7.1.23}$$

$$= \frac{1}{c}S^{i}. \tag{7.1.24}$$

Here $\boldsymbol{S} := c\boldsymbol{E} \times \boldsymbol{B}$ is the Poynting vector. As well as the previously mentioned properties, we have used the fact that η is diagonal, and the components $F^{ij} = -\varepsilon^{ijk}B^{k}$.

From this we can see that the interpretation of the time-space components of the energy–momentum tensor is as the energy density flux.

Finally consider the space-space components, so $\mu = i$ and $\nu = j$. We then have

$$T^{ij} = F^{i\alpha}F_{\alpha}{}^{j} + \frac{1}{4}\eta^{ij}F^{\alpha\beta}F_{\alpha\beta} \tag{7.1.25}$$

$$= F^{i0}F_{0}{}^{j} + F^{ik}F_{k}{}^{j} + \frac{1}{2}\delta^{ij}(\boldsymbol{E}^{2} - \boldsymbol{B}^{2}) \tag{7.1.26}$$

$$= F^{i0}F^{0j} - F^{ik}F^{kj} + \frac{1}{2}\delta^{ij}(\boldsymbol{E}^{2} - \boldsymbol{B}^{2}) \tag{7.1.27}$$

$$= -F^{i0}F^{j0} - \varepsilon^{ikl}B^{l}\varepsilon^{kjm}B^{m} + \frac{1}{2}\delta^{ij}(\boldsymbol{E}^{2} - \boldsymbol{B}^{2}). \tag{7.1.28}$$

As well as the previously mentioned properties, we have used the fact that $\eta^{ij} = -\delta^{ij}$. We now use

$$\varepsilon^{ikl}\varepsilon^{jkm} = \varepsilon^{ilk}\varepsilon^{kjm} = \delta^{ij}\delta^{lm} - \delta^{im}\delta^{lj} \tag{7.1.29}$$

and so

$$\varepsilon^{ikl}\varepsilon^{jkm}B^{l}B^{m} = \boldsymbol{B}^{2} - B^{i}B^{j}. \tag{7.1.30}$$

Hence,

$$T^{ij} = -E^{i}E^{j} - \boldsymbol{B}^{2} - B^{i}B^{j}\frac{1}{2}\delta^{ij}(\boldsymbol{E}^{2} - \boldsymbol{B}^{2}) \tag{7.1.31}$$

$$= -\left(E^{i}E^{j} + B^{i}B^{j} - \frac{1}{2}\delta^{ij}(\boldsymbol{E}^{2} + \boldsymbol{B}^{2})\right) \tag{7.1.32}$$

$$= -[\text{Maxwell Stress Tensor}]^{ij}. \tag{7.1.33}$$

Here we define the **Maxwell stress tensor** as the tensor with components

$$\sigma_{ij} = E^{i}E^{j} + B^{i}B^{j} - \frac{1}{2}\delta^{ij}(\boldsymbol{E}^{2} + \boldsymbol{B}^{2}). \tag{7.1.34}$$

We can think of this as the flux of three–momentum, which also gives us an interpretation of the space-space components of the energy–momentum tensor, and also explains the alternative name of the stress–energy tensor.

7.1.4 Divergence

The divergence of the energy–momentum tensor is fairly easy to compute:

$$\partial_\mu T^{\mu\nu} = \partial_\mu (F^{\mu\alpha} F_\alpha{}^\nu) + \frac{1}{4}\eta^{\mu\nu}\partial_\mu (F^{\alpha\beta}F_{\alpha\beta}) \tag{7.1.35}$$

$$= (\partial_\mu F^{\mu\alpha})F_\alpha{}^\nu + F^{\alpha\mu}\partial_\mu F_\alpha{}^\nu + \frac{1}{4}(\partial^\nu F^{\alpha\beta})F_{\alpha\beta} + \frac{1}{4}F^{\alpha\beta}\partial^\nu F_{\alpha\beta}. \tag{7.1.36}$$

So far this is just the product rule. In the first term we can identify $\partial_\mu F^{\mu\alpha} = J^\alpha/c$ as the inhomogeneous Maxwell equation. In the second term we can lower the α and μ on the first field tensor and raise the μ on the derivative and the α on the second field tensor, so the second term becomes $F_{\alpha\mu}\partial^\mu F^{\alpha\nu}$. In the final term if we lower the α and β on the first field tensor and raise them on the second then we see that the third and final terms are equal. Hence,

$$\partial_\mu T^{\mu\nu} = \frac{1}{c}J^\alpha F_\alpha{}^\nu + F_{\alpha\mu}\,\partial^\mu F^{\alpha\nu} + \frac{1}{2}F_{\alpha\beta}\,\partial^\nu F^{\alpha\beta} \tag{7.1.37}$$

$$= \frac{1}{c}J_\alpha F^{\alpha\nu} + F_{\alpha\mu}\,\partial^\mu F^{\alpha\nu} + \frac{1}{2}F_{\alpha\beta}\,\partial^\nu F^{\alpha\beta} \tag{7.1.38}$$

Now consider the final term of this. We can use the Binachi identity, $\partial^\nu F^{\alpha\beta} + \partial^\alpha F^{\beta\nu} + \partial^\beta F^{\nu\alpha} = 0$, to rewrite the derivative as $\partial^\nu F^{\alpha\beta} = -\partial^\alpha F^{\beta\nu} - \partial^\beta F^{\nu\alpha}$. The final term is therefore

$$\frac{1}{2}F_{\alpha\beta}\partial^\nu F^{\alpha\beta} = -\frac{1}{2}(F_{\alpha\beta}\partial^\alpha F^{\beta\nu} - F_{\alpha\beta}\partial^\beta F^{\nu\alpha}). \tag{7.1.39}$$

If we swap $\alpha \to \mu$ and $\beta \to \alpha$ in the first term and $\beta \to \mu$ in the second term then we see both terms here are the same and we are left with $-F_{\mu\alpha}\partial^\mu F^{\alpha\nu}$, which cancels with the second term in Equation (7.1.38). Hence,

$$\partial_\mu T^{\mu\nu} = \frac{1}{c}J_\alpha F^{\alpha\nu}. \tag{7.1.40}$$

From this we see that the divergence is non-zero when there is a current flow, and hence a flow fo energy/momentum.

Consider the time component of the divergence of the energy–momentum tensor, so $\nu = 0$. On the one side we have

$$\partial_\mu T^{\mu 0} = \partial_0 T^{00} + \partial_i T^{i0} \tag{7.1.41}$$

$$= \frac{1}{c}\frac{\partial u}{\partial t} + \frac{1}{c}\nabla \cdot \boldsymbol{S}. \tag{7.1.42}$$

On the other side we have

$$\frac{1}{c}J_\mu F^{\mu 0} = \frac{1}{c}J_0 F^{00} + \frac{1}{c}J_i F^{i0} = \frac{1}{c}J_i F^{i0} = -\frac{1}{c}J^i F^{i0} = -\frac{1}{c}J^i E^i = -\frac{1}{c}\boldsymbol{J}\cdot\boldsymbol{E}. \tag{7.1.43}$$

We can interpret this as the power density, and we have

$$\frac{\partial u}{\partial t} + \nabla \cdot \boldsymbol{S} = -\boldsymbol{J}\cdot\boldsymbol{E}. \tag{7.1.44}$$

Consider instead the space components of the divergence of the energy–momentum tensor, so $\nu = i$. We have

$$\frac{1}{c}J_\mu F^{\mu i} = \frac{1}{c}(J_0 F^{0i} + J_j F^{ji}) \tag{7.1.45}$$

$$= \frac{1}{c}(-J_0 F^{i0} - J_j F^{ij}) \tag{7.1.46}$$

$$= \frac{1}{c}(-J^0 F^{i0} + J^j F^{ij}) \tag{7.1.47}$$

$$= \frac{1}{c}(-J^0 E^i - J^j \varepsilon^{ijk} B^k) \tag{7.1.48}$$

$$= \frac{1}{c}(-\rho c E^i - (\boldsymbol{J} \times \boldsymbol{B})^i) \tag{7.1.49}$$

$$= -\left(\rho E^i + \frac{1}{c}(\boldsymbol{J} \times \boldsymbol{B})^i\right) \tag{7.1.50}$$

$$= -[\text{Lorentz Force Density}]^i. \tag{7.1.51}$$

So we can interpret this as the force.

This allows us to interpret $\partial_\mu T^{\mu\nu} = J_\mu F^{\mu\nu}/c$ as a conservation law. In particular,

$$\frac{\partial p_{\text{em}}^i}{\partial t} + \partial_j T^{ij} = -\frac{\partial p_{\text{mech}}^i}{\partial t}, \tag{7.1.52}$$

where p_{em}^i is the three-momentum associated with the electromagnetic fields and p_{mech}^i is the three-momentum associated with the movement of matter, whose rate of change is the force.

7.2 Relativistic Point Particle

Consider a point particle. We posit that each particle has two associated Lorentz scalars, its mass, m, and its charge, q. Further the current associated with such a particle is given by $j^\mu = qu^\mu$, where $u^\mu = \mathrm{d}x^\mu/\mathrm{d}\tau$ is the four-velocity of the particle. We therefore have

$$\frac{1}{c}F^{\mu\nu}j_\nu = \frac{q}{c}F^{\mu\nu}u_\nu. \tag{7.2.1}$$

The equation of motion is then

$$\frac{\mathrm{d}p^\mu}{\mathrm{d}\tau} = m\frac{\mathrm{d}u^\mu}{\mathrm{d}\tau} = ma^\mu = \frac{q}{c}F^{\mu\nu}u_\nu. \tag{7.2.2}$$

We know that $u \cdot u = c^2$ is constant and so $u \cdot a = 0$. Hence we have

$$0 = mu \cdot a = \frac{q}{c}u_\mu F^{\mu\nu}u_\nu. \tag{7.2.3}$$

This is guaranteed to be true by the fact that $u_\mu u_\nu$ is symmetric in μ and ν whereas $F^{\mu\nu}$ is antisymmetric in μ and ν. This is a good sign that we are going in the correct direction.

Recall that $u^\mu = (\gamma c, \gamma \boldsymbol{v})$, $\mathrm{d}/\mathrm{d}\tau = \gamma \mathrm{d}/\mathrm{d}t$, and $p^\mu = (E/c, \boldsymbol{p})$. Consider the $\mu = 0$ component, we have

$$\frac{\mathrm{d}p^0}{\mathrm{d}\tau} = \frac{\gamma}{c}\frac{\mathrm{d}E}{\mathrm{d}t} = \frac{q}{c}F^{0\nu}u_\nu = -\frac{q}{c}F^{\nu 0}u_\nu$$

$$= -\frac{q}{c}F^{i0}u_i = \frac{q}{c}F^{i0}u^i = \frac{q}{c}E^i u^i = \frac{q}{c}\boldsymbol{E}\cdot\boldsymbol{v}. \quad (7.2.4)$$

Now consider the $\mu = i$ components,

$$\frac{\mathrm{d}p^i}{\mathrm{d}\tau} = \gamma\frac{\mathrm{d}p^i}{\mathrm{d}t} \tag{7.2.5}$$

$$= \frac{q}{c}F^{i\nu}u_\nu \tag{7.2.6}$$

$$= \frac{q}{c}(F^{i0}u_0 + F^{ij}u_j) \tag{7.2.7}$$

$$= \frac{q}{c}(E^i u_0 + \varepsilon^{ijk}B^k \gamma v^j) \tag{7.2.8}$$

$$= \frac{q}{v}(\gamma c E^i + \gamma(\boldsymbol{v}\times\boldsymbol{B})^i) \tag{7.2.9}$$

$$= q\gamma E^i + \frac{q}{c}\gamma(\boldsymbol{v}\times\boldsymbol{B})^i. \tag{7.2.10}$$

Hence,

$$\frac{\partial \boldsymbol{p}}{\partial t} = q\boldsymbol{E} + \frac{q}{c}(\boldsymbol{v}\times\boldsymbol{B}). \tag{7.2.11}$$

This is exactly the Lorentz force law for a charged particle in an electromagnetic field. This shows us that we are considering the correct current with $j_\nu = qu_\nu$.

Part III

Action Principles

Eight

Non-Relativistic Actions

I got a career out of Taylor expanding.

Donal O'Connell

This section will be mostly a review of content covered in other courses, in particular in the Lagrangian dynamics course and the quantum theory course. For more on actions and the calculus of variations see the Lagrangian dynamics course and for more on the particular methods used here to find stationary actions see the quantum theory course.

We will introduce the key ideas of action principles through some, hopefully familiar, non-relativistic examples, before we discus relativistic generalisations. We start with a very simple example.

8.1 Non-Relativistic Point Particle in One Dimension in a Potential

Consider a non-relativistic point particle in one dimension in a potential, V, which depends on the position of the particle, x, and nothing else. The action for this particle along some path, $x(t)$, between times a and b is defined to be the functional

$$S[x(t)] = \int_a^b \left[\frac{1}{2} m \dot{x} - V(x) \right] \mathrm{d}t. \tag{8.1.1}$$

We consider a small variation, $x(t) \to x(t) + \delta x(t)$, where $\delta x(t)$ is small. The action on this varied path is

$$S[x(t) + \delta x(t)] = \int_a^b \left[\frac{1}{2} m \left(\frac{\mathrm{d}}{\mathrm{d}t} [x(t) + \delta x(t)] \right)^2 - V(x(t) + \delta x(t)) \right] \mathrm{d}t$$

$$= \int_a^b \left[\frac{1}{2} m \left(\frac{\mathrm{d}x}{\mathrm{d}t} + \frac{\mathrm{d}}{\mathrm{d}t} \delta x(t) \right)^2 - V(x(t) + \delta x(t)) \right] \mathrm{d}t. \tag{8.1.2}$$

Taylor expanding the potential to first order in $\delta x(t)$, and expanding the brackets, again keeping only terms up to linear order in $\delta x(t)$ we get

$$S[x(t) + \delta x(t)] = \int_a^b \left[\frac{1}{2} m \dot{x}(t)^2 + m \dot{x} \frac{\mathrm{d}}{\mathrm{d}t} \delta x(t) - V(x(t)) - V'(x(t)) \delta x(t) \right] \mathrm{d}t.$$

$$(8.1.3)$$

Carrying around all of this explicit time dependence is getting a bit bothersome, so we'll drop the (t) for x and δx most of the time:

$$S[x + \delta x] = \int_a^b \left[\frac{1}{2} m\dot{x} + \dot{x} \frac{\mathrm{d}}{\mathrm{d}t} \delta x - V(x) - V'(x)\delta x \right] \mathrm{d}t. \qquad (8.1.4)$$

Notice that we can write this in terms of the action on the unvaried path, plus some extra terms:

$$S[x + \delta x] = S[x] + \int_a^b \left[m\dot{x} \frac{\mathrm{d}}{\mathrm{d}t} \delta x - V'(x)\delta x \right] \mathrm{d}t. \qquad (8.1.5)$$

We therefore define the variation in S due to this variation in the path to be

$$\delta S[x(t)] := S[x(t) + \delta x(t)] - S[x(t)]. \qquad (8.1.6)$$

So,

$$\delta S[x(t)] = \int_a^b \left[m\dot{x} \frac{\mathrm{d}}{\mathrm{d}t} \delta x - V'(x)\delta x \right] \mathrm{d}t. \qquad (8.1.7)$$

In order to extract any physics from this variation we need the principle of least action.

> **Claim 8.1.8 — Principle of Least Action** The **principle of least action**, or more accurately the principle of stationary action, states that the physical trajectory of a particle is such as to minimise the appropriate action.

The phrase "appropriate action" is quite a lot here, since we usually define the action such that applying the principle of least action gives the correct results, so its a rather circular argument, but it works, and it generalises.

One argument for why the principle of least action applies comes from the path integral formulation of quantum mechanics. In this we have an integral of $\exp(iS[x(t)]/\hbar)$. This oscillates a lot, apart from near to where $S[x(t)]$ is stationary, and so only the stationary path has a meaningful contribution to the integral, with the rapid oscillations cancelling out.

Another thing we need to consider is what it means for a functional to be stationary. For a normal function we would simply say that $\mathrm{d}f/\mathrm{d}t = 0$ if f is stationary at some value of t. Similarly for a functional we say that $\delta f[x(t)] = 0$ if f is stationary on some path, $x(t)$. So, the principle of least action simply reduces to looking for $x(t)$ such that $\delta S[x(t)] = 0$.

Before we can do this we will need a theorem about integrals.

> **Theorem 8.1.9.** Let f be some sufficiently smooth function. If
>
> $$\int_a^b f(t)\delta x(t)\,\mathrm{d}t = 0 \qquad (8.1.10)$$
>
> for all distributions $\delta x(t)$ then $f(t) = 0$ for all $t \in [a, b]$.

"proof". Let $\delta x(t) = \delta(t - t_0)$ for some $t_0 \in [a, b]$. Then

$$0 = \int_a^b f(t)\delta x(t)\,dt = \int_a^b f(t)\delta(t - t_0)\,dt = f(t_0). \qquad (8.1.11)$$

Hence $f(t) = 0$ for all $t \in [a, b]$. $\qquad\qquad\qquad\qquad\qquad\qquad\square$

This theorem can almost be applied to Equation (8.1.7). The problem is that we have a derivative of δx, which doesn't work with this theorem. The solution is to either integrate by parts[1], or, equivalently, to massage our statement to allow us to integrate out this derivative. To do this we write

[1]as demonstrated multiple times in the quantum theory course

$$\frac{d}{dt}(\dot{x}\,\delta x) = \ddot{x}\,\delta x + \dot{x}\frac{d}{dt}\delta x. \qquad (8.1.12)$$

Hence,

$$\int_a^b \dot{x}\frac{d}{dt}\delta x\,dt = \int_a^b \left[\frac{d}{dt}(\dot{x}\,\delta x) - \ddot{x}\,\delta x\right]dt = [\dot{x}\,\delta x]_a^b - \int_a^b \ddot{x}\,\delta x\,dt. \quad (8.1.13)$$

Of course, this is exactly integration by parts, we've just done the middle steps that we usually miss out.

Using this we have

$$\delta S[x(t)] = \int_a^b \left[m\dot{x}\frac{d}{dt}\delta x - V'(x)\delta x\right]dt \qquad (8.1.14)$$

$$= [m\dot{x}\,\delta x(t)]_a^b - \int_a^b [m\ddot{x} + V'(x)]\delta x\,dt. \qquad (8.1.15)$$

Now restrict the variation to be zero on the boundary, so $\delta x(a) = \delta x(b) = 0$. This means that the first term vanishes and we have

$$\delta S[x(t)] = -\int_a^b [m\ddot{x} + V'(x)]\delta x\,dt. \qquad (8.1.16)$$

Imposing the principle of least action and requiring that this holds for any variation with $\delta x(a) = \delta x(b) = 0$ we have

$$0 = \int_a^b [m\ddot{x} + V'(x)]\delta x\,dt \implies m\ddot{x} + V'(x) = 0 \implies m\ddot{x} = -V'(x). \ (8.1.17)$$

Of course, this is just Newton's second law for a particle of mass m in a potential.

8.1.1   The Basic Idea

We have seen here all of the basic ideas behind using action principles. All that remains is to apply these ideas to more and more complex systems. The basic procedure is always the same:

1. Write out the action.

2. Write out the action on the varied path.

3. Taylor expand to first order in the variation.

4. Compute the variation in the action.

5. Set the variation in the action to zero.

6. Integrate by parts to get rid of derivatives.

7. We will be left with something of the form

$$\int f(t)\delta x(t)\,\mathrm{d}t = 0 \tag{8.1.18}$$

and requiring that this holds for any variation vanishing on the boundaries means that we must have $f(t) = 0$ as our equation of motion.

8.1.2 Calculus of Variations

We define the variation in something, say $f(x(t))$, depending on the path $x(t)$ to be

$$\delta f(x(t)) := f(x(t) + \delta x(t)) - f(x(t)). \tag{8.1.19}$$

One important example of this is the variation in $\mathrm{d}x/\mathrm{d}t$:

$$\delta\frac{\mathrm{d}x}{\mathrm{d}t} = \delta\frac{\mathrm{d}}{\mathrm{d}t}(x(t)) = \frac{\mathrm{d}}{\mathrm{d}t}(x(t) + \delta x(t)) - \frac{\mathrm{d}}{\mathrm{d}t}(x(t))$$

$$= \frac{\mathrm{d}}{\mathrm{d}t}x(t) + \frac{\mathrm{d}}{\mathrm{d}t}\delta x(t) - \frac{\mathrm{d}}{\mathrm{d}t}x(t) = \frac{\mathrm{d}}{\mathrm{d}t}\delta x(t). \tag{8.1.20}$$

We think of this as meaning that $\mathrm{d}/\mathrm{d}t$ and δ commute, so

$$\delta\frac{\mathrm{d}}{\mathrm{d}t} = \frac{\mathrm{d}}{\mathrm{d}t}\delta. \tag{8.1.21}$$

Similarly one can show that we can exchange the order of variations and integrals.

A fortunate consequence of this is that we can interpret $\delta\dot{x}$ in two ways, and both are equal. Either,

$$\delta\dot{x} = \delta(\dot{x}) = \delta\frac{\mathrm{d}x}{\mathrm{d}t} = \frac{\mathrm{d}}{\mathrm{d}t}\delta x, \qquad \text{or} \qquad \delta\dot{x} = \dot{\delta x} = \frac{\mathrm{d}}{\mathrm{d}t}\delta x = \delta\frac{\mathrm{d}x}{\mathrm{d}t}. \tag{8.1.22}$$

Since these are equal there is no ambiguity in the notation.

8.2 Non-Relativistic General Particle in One Dimension in a Potential

A slightly more generic action for a particle in one dimension in some potential is given by defining a **Lagrangian**, $L(x(t), \dot{x}(t)) = T - V$, where T is the kinetic energy of the particle and V is the potential energy. Note that we choose here for the Lagrangian to depend only on the position and velocity of the particle. This is not always the case, but often is. When this doesn't hold one has to take care with which of the following steps are still valid.

The Lagrangian is defined such that the action is given by its integral:

$$S[x(t)] = \int_a^b L(x(t), \dot{x}(t)) \, dt. \tag{8.2.1}$$

So, we can identify the Lagrangian from the previous example as

$$L = \frac{1}{2} m \dot{x}^2 - V(x), \tag{8.2.2}$$

which is exactly $T - V$.

The variation in the action is

$$\delta S[x(t)] = \delta \int_a^b L(x, \dot{x}) \, dt \tag{8.2.3}$$

$$= \int_a^b \delta L(x, \dot{x}) \, dt \tag{8.2.4}$$

$$= \int_a^b [L(x + \delta x, \dot{x} + \delta \dot{x}) - L(x, \dot{x})] \, dt. \tag{8.2.5}$$

To proceed we Taylor expand the Lagrangian to first order in δx and $\delta \dot{x}$. Note that we treat x and $\dot{x}$ as independent variables before we find the equations of motion. Therefore we simply have a Taylor expansion in two variables:

$$L(x + \delta x, \dot{x} + \delta \dot{x}) = L(x, \dot{x}) + \frac{\partial L}{\partial x} \delta x + \frac{\partial L}{\partial \dot{x}} \delta \dot{x}. \tag{8.2.6}$$

Hence, the variation in the Lagrangian is

$$\delta L(x, \dot{x}) = \frac{\partial L}{\partial x} \delta x + \frac{\partial L}{\partial \dot{x}} \delta \dot{x}, \tag{8.2.7}$$

and the variation in the action is

$$\delta S[x(t)] = \int_a^b \left[\frac{\partial L}{\partial x} \delta x + \frac{\partial L}{\partial \dot{x}} \delta \dot{x} \right] dt. \tag{8.2.8}$$

In order to deal with the derivative of δx we again integrate by parts, rewriting

$$\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{x}} \delta x \right) = \frac{\partial L}{\partial \dot{x}} \delta \dot{x} + \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{x}} \right) \delta x \tag{8.2.9}$$

we have

$$\delta S[x(t)] = \int_a^b \left[\frac{\partial L}{\partial x} \delta x + \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{x}} \delta x \right) - \frac{d}{dt} \left(\frac{\partial L}{\partial t} \right) \delta x \right] dt \tag{8.2.10}$$

$$= \left[\frac{\partial L}{\partial \dot{x}} \delta x \right]_a^b + \int_a^b \left[\frac{\partial L}{\partial x} \delta x - \frac{d}{dt} \left(\frac{\partial L}{\partial t} \right) \right] \delta x \, dt. \tag{8.2.11}$$

Again we impose that the variation vanishes at the boundary, so the first term vanishes. The principle of least action then gives us

$$\frac{\partial L}{\partial x} = \frac{d}{dt} \frac{\partial L}{\partial \dot{x}}. \tag{8.2.12}$$

This is the **Euler–Lagrange equation**.

8.2.1 Basic Idea

The Euler–Lagrange equation gives us another way to find the physical trajectory. If we know the action in the form

$$S = \int L \, \mathrm{d}t \tag{8.2.13}$$

then we can get the equations of motion by either

- varying S and setting the variation to zero, or

- solving the Euler–Lagrange equations.

The first will always work. The second requires that we have the right form of the Euler–Lagrange equations for the problem at hand, for example as derived here the standard Euler–Lagrange equation works if the Lagrangian depends on position and velocity, but not if it also depends on acceleration.

Nine

Relativistic Free Action

9.1 Motivation

In this section we will derive the action and Lagrangian for a free relativistic particle. We do so motivated by knowing the equations of motion for such a system are

$$m\frac{\mathrm{d}u^{\mu}}{\mathrm{d}\tau} = 0, \tag{9.1.1}$$

where u is the four-velocity of the particle.

For a free particle the motion, and hence action and Lagrangian, shouldn't depend on the location of the particle, x, which we now take to be the four-position. More formally put our system should have translation symmetry. Recall that this implies momentum conservation, by Noether's theorem[1], which we would expect since the equation of motion

$$0 = m\frac{\mathrm{d}u^{\mu}}{\mathrm{d}\tau} = \frac{\mathrm{d}p^{\mu}}{\mathrm{d}\tau} \tag{9.1.2}$$

can be seen as a statement of conservation of momentum. Taking x to be the four-position if our Lagrangian is independent of x then it is also time independent, which implies energy conservation, this also follows from the time component of the equation of motion.

Another requirement we impose is that the Lagrangian be Lorentz invariant. Strictly all we need is for the action to be Lorentz invariant, but this is almost always achieved by having a Lorentz invariant Lagrangian. This restriction is reasonable as the physics of the system should not depend on the frame and we want to derive Lorentz covariant equations of motion.

Recall that in the non-relativistic case a free particle has the Lagrangian

$$L = \frac{1}{2}m\dot{\boldsymbol{x}} \cdot \dot{\boldsymbol{x}}. \tag{9.1.3}$$

The obvious choice then is to replace the three-velocity, $\dot{\boldsymbol{x}}$ with the four-velocity, u. However, recall that $u \cdot u = c^2$, so this alone cannot give us the particle dynamics. The requirement that the Lagrangian is both Lorentz invariant and position independent is actually quite limiting. There are only really three Lorentz scalars from which we can build the Lagrangian. They are the mass, m, the speed of light, $c^2 = u \cdot u$, and the proper time experienced by the particle, τ. Motivated by the

[1]See the notes from the Lagrangian dynamics course for details.

non-relativistic case we look for an action of the form

$$S[x(t)] = \int_a^b L(x(t), x'(t)) \, d\tau \tag{9.1.4}$$

where x is the four-position, x' the four velocity, and we replace coordinate time, t, with proper time τ.

The action must have units of energy times time, and the Lagrangian units of energy. This restricts the possible ways we can combine these three Lorentz scalars. After some thinking one comes to the conclusion that the correct form for the action of a free relativistic particle is

$$S[x(t)] = -mc^2 \int_a^b d\tau \iff L = -mc^2. \tag{9.1.5}$$

This is motivated by the fact that it correctly reduces to the non-relativistic result in the non-relativistic limit, and also gives the expected equations of motion in the full relativistic case. We will demonstrate this now.

9.2 Non-Relativistic Limit

The proper time is given by $d\tau = dt/\gamma$. In the non-relativistic limit $v \ll c$, and so we can Taylor expand $1/\gamma$, giving

$$\frac{1}{\gamma} = \sqrt{1 - \frac{v^2}{c^2}} = \left(1 - \frac{v^2}{c^2}\right)^{1/2} \approx 1 - \frac{1}{2}\frac{v^2}{c^2}. \tag{9.2.1}$$

Hence,

$$S[x(t)] = -mc^2 \int_a^b d\tau \tag{9.2.2}$$

$$\approx -mc^2 \int_a^b \left(1 - \frac{1}{2}\frac{v^2}{c^2}\right) dt \tag{9.2.3}$$

$$= \int_a^b \frac{1}{2}mv^2 \, dt + \text{constant}. \tag{9.2.4}$$

The constant is unimportant for the dynamics of the system and can be neglected[2]. We see therefore that this choice of relativistic Lagrangian is equivalent to the non-relativistic choice of Lagrangian for a free particle.

9.3 Relativistic Equation of Motion

We will now derive the relativistic equation of motion by varying S. Doing so we have

$$\delta S = -mc^2 \delta \int d\tau. \tag{9.3.1}$$

The interpretation of $\int d\tau$ is as the total proper time along the path of the particle through spacetime (i.e., its world line). In order to have $\delta S = 0$, as required

[2]For a proof of this see the Lagrangian dynamics course. Essentially the Lagrangian is always differentiated when we use it so any constants vanish, this is similar to how we can add a constant to the potential energy by choosing a different place to define as zero.

by the principle of stationary action we need to extremise the total proper time

along the path. One way to do this is to take a straight line in spacetime[3]. This is a global property of the path. Equations of motion are typically differential equations, which are local, so we need another way to think about this.

Consider two paths, x_1 and x_2, both starting and ending at the same points. We can parametrise each path using the proper time along that path, τ_i. So $x_i = x_i(\tau_i)$ for $i = 1, 2$. The variation in the action between these two paths is

$$\delta S = -mc^2 \left[\int_{\text{path 1}} d\tau_1 - \int_{\text{path 2}} d\tau_2 \right]. \tag{9.3.2}$$

In order to derive the equations of motion we first consider path 2 as a deviation from path 1. We then choose to use τ_1 as a common parameter for both paths, which we can do since we can find consider τ_2 as a function of τ_1, $\tau_2(\tau_1)$. We then

have[4]

$$\frac{d\tau_2}{d\tau_1} = \frac{d\tau_2}{d\tau_1} \implies d\tau_2 = \frac{d\tau_2}{d\tau_1} d\tau_1. \tag{9.3.3}$$

This allows us to combine the two integrals over the paths into one, giving

$$\delta S = -mc^2 \int \left[1 - \frac{d\tau_2}{d\tau_1} \right] d\tau_1. \tag{9.3.4}$$

The problem with this approach is that it promotes path 1 above path 2. In order to avoid this we introduce some third parameter, λ, which parametrise both

paths[5], so we consider τ_i as a function of λ, $\tau_i(\lambda)$. We then have two paths, $x_1(\lambda)$ and $x_2(\lambda)$.

Next we need to find a way of writing $d\tau_i$ in terms of λ. To do so we consider the spacetime interval:

$$c^2 d\tau^2 = dx \cdot dx \implies d\tau = \frac{1}{c}\sqrt{dx \cdot dx}. \tag{9.3.5}$$

We can then write dx in terms of λ using

$$\frac{dx}{d\lambda} = \frac{dx}{d\lambda} \implies dx = \frac{dx}{d\lambda} d\lambda. \tag{9.3.6}$$

Factoring the factor of $d\lambda$ out of the square root we have

$$d\tau = \frac{1}{c}\sqrt{\frac{dx}{d\lambda} \cdot \frac{dx}{d\lambda}} d\lambda. \tag{9.3.7}$$

Hence, the action can be written in terms of the parameter λ as

$$S[x(t)] = -mc \int_{\lambda_a}^{\lambda_b} \sqrt{\frac{dx}{d\lambda} \cdot \frac{dx}{d\lambda}} d\lambda. \tag{9.3.8}$$

Here λ_a and λ_b are such that $\tau_i(\lambda_a) = a$ and $\tau_i(\lambda_b) = b$ fix the end points of the two paths. The Lagrangian parametrised by λ is then

$$L(x'(\lambda)) = -mc\sqrt{\frac{dx}{d\lambda} \cdot \frac{dx}{d\lambda}}. \tag{9.3.9}$$

9.3.1 Euler–Lagrange Equations

We can now consider a slightly more general Lagrangian with position dependence, which will be useful later, in order to derive the Euler–Lagrange equations for a relativistic particle:

$$S[x(t)] = \int_{\lambda_a}^{\lambda_b} L(x(\lambda), x'(\lambda)) \, d\lambda. \tag{9.3.10}$$

The variation in S is given by varying the Lagrangian, which is done by Taylor expanding, and then integrating by parts until we can identify the integrand as zero if δS is to vanish. Some care must be taken over this since x is a four-vector, but its pretty much identical to the one-dimensional non-relativistic case:

$$\delta S[x(t)] = \int_{\lambda_a}^{\lambda_b} \delta L(x(\lambda), x'(\lambda)) \, d\lambda \tag{9.3.11}$$

$$= \int_{\lambda_a}^{\lambda_b} \left[\frac{\partial L}{\partial x^\mu} \delta x^\mu + \frac{\partial L}{\partial x'^\mu} \delta x'^\mu \right] d\lambda \tag{9.3.12}$$

$$= \int_{\lambda_a}^{\lambda_b} \left[\frac{\partial L}{\partial x^\mu} \delta x^\mu + \frac{d}{d\lambda} \left(\frac{\partial L}{\partial x'^\mu} \delta x^\mu \right) - \frac{d}{d\lambda} \left(\frac{\partial L}{\partial x'^\mu} \right) \delta x^\mu \right] d\lambda \tag{9.3.13}$$

$$= \left[\frac{\partial L}{\partial x'^\mu} \delta x^\mu \right]_{\lambda_a}^{\lambda_b} + \int_{\lambda_a}^{\lambda_b} \left[\frac{\partial L}{\partial x^\mu} - \frac{d}{d\lambda} \left(\frac{\partial L}{\partial x'^\mu} \right) \right] \delta x^\mu \, d\lambda. \tag{9.3.14}$$

We require that δx^μ vanishes on the boundary, so the first term vanishes. We further require that this holds for arbitrary δx^μ, and so we have that each component of the square bracket in the integral must vanish independent of the other components giving the Euler Lagrange equations for a relativistic particle:

$$\frac{d}{d\lambda} \frac{\partial L}{\partial x'^\mu} = \frac{\partial L}{\partial x^\mu}. \tag{9.3.15}$$

9.3.2 Recovering Equations of Motion

Finally we are ready to derive the equations of motion for the relativistic particle. We go back to the position-independent Lagrangian, $L = -mc\sqrt{x' \cdot x'}$. We then have

$$\frac{dL}{dx^\mu} = 0 \implies \frac{d}{d\lambda} \frac{\partial L}{\partial x'^\mu} = 0. \tag{9.3.16}$$

So we need to compute $\partial L / \partial x'^\mu$. An application of the chain rule gives

$$\frac{\partial L}{\partial x'^\mu} = -mc \frac{\partial}{\partial x'^\mu} \sqrt{x'^\nu x'_\nu} = -\frac{1}{2} \frac{mc}{\sqrt{x' \cdot x'}} \frac{\partial}{\partial x'^\mu} (x'^\nu x'_\nu). \tag{9.3.17}$$

We can compute this derivative in a fairly straight forward way, we just have

to be careful about applying the product rule and matching indices:

$$\frac{\partial}{\partial x'^\mu}(x'^\nu x'_\nu) = \frac{\partial}{\partial x'^\mu}(x'^\nu \eta_{\nu\rho} x'^\rho) \tag{9.3.18}$$

$$= \frac{\partial x'^\nu}{\partial x'^\mu}\eta_{\nu\rho} x'^\rho + x'^\nu \eta_{\nu\rho}\frac{\partial x'^\rho}{\partial x'^\mu} \tag{9.3.19}$$

$$= \delta^\nu{}_\mu \eta_{\nu\rho} x'^\rho + x'^\nu \eta_{\nu\rho}\delta^\rho{}_\mu \tag{9.3.20}$$

$$= \eta_{\mu\rho} x'^\rho + x'^\nu \eta_{\nu\mu} \tag{9.3.21}$$

$$= 2\eta_{\mu\nu} x'^\nu \tag{9.3.22}$$

$$= 2\eta x'_\mu. \tag{9.3.23}$$

Here we renamed $\rho \to \mu$ and used the symmetry of the metric to combine the two terms into one. We therefore have

$$\frac{\partial L}{\partial x'^\mu}(x' \cdot x') = 2x'_\mu. \tag{9.3.24}$$

We could probably have guessed this by considering $\partial x^2/\partial x = 2x$ and matching indices, and indeed this is how people usually perform this calculation after some practice.

The result is that

$$\frac{\partial L}{\partial x'^\mu} = -\frac{mcx'_\mu}{\sqrt{x' \cdot x'}}. \tag{9.3.25}$$

Hence, the equation of motion is

$$\frac{\mathrm{d}}{\mathrm{d}\lambda}\left[\frac{mc}{\sqrt{\frac{\mathrm{d}x}{\mathrm{d}\lambda} \cdot \frac{\mathrm{d}x}{\mathrm{d}\lambda}}}\frac{\mathrm{d}x^\mu}{\mathrm{d}\lambda}\right] = 0 \tag{9.3.26}$$

where we have written the derivatives out in full, instead of using primes. We have also dropped the negative and raised the index on x'_μ, since the left hand side is equal to zero, and so the overall sign is not important.

While this looks quite complicated it turns out that it is just a straight line in spacetime. The complexity arises because λ may be a fairly complex parametrisation. The equation of motion gets significantly more simple if we choose a sensible, simple, parametrisation, such as a proper time.

Now consider a single path and choose to use the proper time along this path as the parameter, so $\lambda = \tau$. For a single path there is no reason not to use the proper time, it was only when comparing paths that this caused problems, since its different for each path. We then have that

$$\frac{\mathrm{d}x}{\mathrm{d}\lambda} = \frac{\mathrm{d}x}{\mathrm{d}\tau} = u \implies \frac{\mathrm{d}x}{\mathrm{d}\lambda} \cdot \frac{\mathrm{d}x}{\mathrm{d}\lambda} = u \cdot u = c^2. \tag{9.3.27}$$

Hence, the equation of motion simplifies to

$$\frac{\mathrm{d}}{\mathrm{d}\tau}\left[m\frac{\mathrm{d}x^\mu}{\mathrm{d}\tau}\right] = 0 \implies m\frac{\mathrm{d}u^\mu}{\mathrm{d}\tau} = 0, \tag{9.3.28}$$

which is exactly the equation of motion that we were looking to find, justifying our choice of action.

9.3.3 Reparametrisation Invariance

This use of parametrisations can be a bit worrying, after all parametrisation is an arbitrary choice, like a coordinate choice or a choice of units, and so it shouldn't effect the overall result. We can quite quickly show that this action is independent of how we parametrise the path.

Suppose we have two parameters, λ and κ, we want to show that choosing to parametrise our path with κ leaves the form of the action unchanged. To do so we consider κ as a function of λ, which must be possible since both parametrise the same path and we can therefore draw a correspondence between them. Now that we have $\kappa = \kappa(\lambda)$ we have

$$\frac{d\kappa}{d\lambda} = \kappa' \implies d\kappa = \kappa' \, d\lambda. \tag{9.3.29}$$

By the chain rule we also have

$$\frac{dx}{d\lambda} = \frac{dx}{d\kappa} \frac{d\kappa}{d\lambda}. \tag{9.3.30}$$

Hence, the action is

$$S = \int \left[-mc \sqrt{\frac{dx}{d\lambda} \cdot \frac{dx}{d\lambda}} \right] d\lambda \tag{9.3.31}$$

$$= \int \left[-mc \sqrt{\frac{dx}{d\kappa} \cdot \frac{dx}{d\kappa} \left(\frac{d\kappa}{d\lambda} \right)^2} \right] \frac{1}{\kappa'} \, d\kappa \tag{9.3.32}$$

$$= \int \left[-mc \sqrt{\frac{dx}{d\kappa} \cdot \frac{dx}{d\kappa}} \right] d\kappa. \tag{9.3.33}$$

So, the form of the action is independent of the choice of parametrisation. This is called **reparametrisation invariance**.

In particular this invariance means we may as well choose a simple parametrisation, such as proper time, leaving us with a particularly simple form of the action

$$S = -mc^2 \int d\tau. \tag{9.3.34}$$

Ten

Alternative Actions

The choice of action is not unique. We say that two actions are equivalent if they give the same equations of motion. Similarly two Lagrangians are equivalent if they give equivalent actions.

There are some obvious things we can do to Lagrangians to give trivially equivalent actions, first, we can add a constant. This has no effect on the equations of motion since we always differentiate the Lagrangian and so this term will vanish in the Euler–Lagrange equations. Second, we can add a total derivative term, this will not change the Euler–Lagrange equations either[1]. There are also non-trivial changes we can make to the action which still give the same results, including using completely different Lagrangians.

[1]For proof see the notes from the Lagrangian dynamics course.

10.1 Alternative Relativistic Action

For a relativistic particle we ruled out using $m\dot{x}^2/2$, since $\dot{x}^2 = c^2$ if we interpret this as a $(3 + 1)$-dimensional equation. However, it turns out that if we take this as the action for a free particle we do actually get the same equations of motion as with the action defined in the previous chapter. That is the action in Equation (9.3.8) is equivalent to the action

$$S_2[x(\tau)] := -\int \frac{m}{2} \frac{\mathrm{d}x^\mu}{\mathrm{d}\tau} \frac{\mathrm{d}x_\mu}{\mathrm{d}\tau} \, \mathrm{d}t, \tag{10.1.1}$$

we simply forget that $\dot{x}^2 = c^2$ until after we have done the variation. We will demonstrate this with an even more general choice of action.

Consider the action

$$\tilde{S}[x(\lambda)] := \int \left[-g(\lambda)^{-1} \frac{\mathrm{d}x^\mu}{\mathrm{d}\lambda} \frac{\mathrm{d}x_\mu}{\mathrm{d}\lambda} - \left(\frac{mc}{2}\right)^2 g(\lambda) \right] \mathrm{d}\lambda. \tag{10.1.2}$$

This has the undetermined quantity $g(\lambda)$. We can place restrictions on this quantity by requiring that this action has reparametrisation invariance. That is if we consider some new parameter, $\tilde{\lambda}$. We then have

$$\frac{\mathrm{d}x^\mu}{\mathrm{d}\lambda} \frac{\mathrm{d}x_\mu}{\mathrm{d}\lambda} = \frac{\mathrm{d}x^\mu}{\mathrm{d}\tilde{\lambda}} \frac{\mathrm{d}x_\mu}{\mathrm{d}\tilde{\lambda}} \left(\frac{\mathrm{d}\tilde{\lambda}}{\mathrm{d}\lambda}\right)^2, \tag{10.1.3}$$

and

$$\mathrm{d}\lambda = \frac{\mathrm{d}\lambda}{\mathrm{d}\tilde{\lambda}} \mathrm{d}\tilde{\lambda}. \tag{10.1.4}$$

Hence, the action can be written as

$$\tilde{S}[x(\tilde{\lambda})] = \int \left[-g(\lambda(\tilde{\lambda}))^{-1} \frac{\mathrm{d}x^\mu}{\mathrm{d}\tilde{\lambda}} \frac{\mathrm{d}x_\mu}{\mathrm{d}\tilde{\lambda}} \frac{\mathrm{d}\tilde{\lambda}}{\mathrm{d}\lambda} - \left(\frac{mc}{2}\right)^2 g(\lambda(\tilde{\lambda})) \frac{\mathrm{d}\lambda}{\mathrm{d}\tilde{\lambda}} \right] \mathrm{d}\tilde{\lambda}. \qquad (10.1.5)$$

Notice that one factor of $\mathrm{d}\tilde{\lambda}/\mathrm{d}\lambda$ from the chain rule applied to the derivatives cancels with a factor of $\mathrm{d}\lambda/\mathrm{d}\tilde{\lambda}$ from the change in the measure. Enforcing reparametrisation invariance is then equivalent to requiring that

$$\tilde{S}[x(\tilde{\lambda})] = \int \left[-g(\tilde{\lambda})^{-1} \frac{\mathrm{d}x^\mu}{\mathrm{d}\tilde{\lambda}} \frac{\mathrm{d}x_\mu}{\mathrm{d}\tilde{\lambda}} - \left(\frac{mc}{2}\right)^2 g(\tilde{\lambda}) \right] \mathrm{d}\tilde{\lambda}. \qquad (10.1.6)$$

We can see that this is true so long as

$$g(\tilde{\lambda}) = \frac{\mathrm{d}\lambda}{\mathrm{d}\tilde{\lambda}} g(\lambda(\tilde{\lambda})). \qquad (10.1.7)$$

Note that this means that g does not transform like a normal scalar function, which would be unchanged under reparametrisation, $f(\lambda) = f(\tilde{\lambda})$. It is perhaps clearer if we write this requirement as

$$g(\tilde{\lambda})\mathrm{d}\tilde{\lambda} = g(\lambda)\mathrm{d}\lambda. \qquad (10.1.8)$$

This requirement is equivalent to saying that $g(\lambda)\mathrm{d}\lambda$ is a one-form[2,3].

[2]A common abuse of terminology is to call $g(\lambda)$ alone a one-form.

[3]For more on one-forms see the notes from the general relativity course.

10.1.1 Equations of Motion

We can get equations of motion from $\tilde{S}$ by applying the Euler–Lagrange equations, and treating g as a degree of freedom, that is we take as the Lagrangian

$$L\left(x^\mu(\lambda), \frac{\mathrm{d}x^\mu}{\mathrm{d}\lambda}, g(\lambda), \frac{\mathrm{d}g}{\mathrm{d}\lambda}\right) = -g(\lambda)^{-1} \frac{\mathrm{d}x^\mu}{\mathrm{d}\lambda} \frac{\mathrm{d}x_\mu}{\mathrm{d}\lambda} - \left(\frac{mc}{2}\right)^2 g(\lambda). \qquad (10.1.9)$$

Doing so with g as our degree of freedom we have

$$\frac{\partial L}{\partial\left(\frac{\mathrm{d}g}{\mathrm{d}\lambda}\right)} = 0 \qquad (10.1.10)$$

and so one equation of motion is

$$0 = \frac{\partial L}{\partial g} = \frac{1}{g(\lambda)^2} \left(\frac{\mathrm{d}x}{\mathrm{d}\lambda}\right)^2 - \left(\frac{mc}{2}\right)^2. \qquad (10.1.11)$$

Now taking x as our degree of freedom we have

$$\frac{\mathrm{d}L}{\mathrm{d}x} = 0 \qquad (10.1.12)$$

and so

$$\frac{\mathrm{d}}{\mathrm{d}\lambda} \frac{\partial L}{\partial\left(\frac{\mathrm{d}x}{\mathrm{d}\lambda}\right)} = \frac{\mathrm{d}}{\mathrm{d}\lambda} \left[-\frac{1}{g(\lambda)} \frac{\mathrm{d}x^\mu}{\mathrm{d}\lambda} \right] = 0 \qquad (10.1.13)$$

is the other equation of motion.

Consider what happens if we choose $g(\lambda) = 2/m$. We then have that

$$\left(\frac{\mathrm{d}x}{\mathrm{d}\lambda}\right)^2 = c^2 \tag{10.1.14}$$

and our second equation of motion becomes

$$\frac{\mathrm{d}}{\mathrm{d}\lambda}\left[-\frac{m}{2}\frac{\mathrm{d}x^\mu}{\mathrm{d}\lambda}\right] = 0. \tag{10.1.15}$$

These are exactly the same equations of motion as we got by choosing to take $\lambda = \tau$ in our original action, hence $\tilde{S}$ and S are equivalent actions.

Setting $g(\lambda) = 2/m$ in $\tilde{S}$ we get

$$\tilde{S} = \int\left[-\frac{m}{2}\frac{\mathrm{d}x^\mu}{\mathrm{d}\lambda}\frac{\mathrm{d}x_\mu}{\mathrm{d}\lambda} + \text{constant}\right]\mathrm{d}\lambda \tag{10.1.16}$$

and since the constant can be ignored without changing the equations of motion we see that S_2 is an equivalent action to S also.

10.1.2   Which Action To Choose?

So, there are three actions, S, S_2, and $\tilde{S}$, which all produce the same equations of motion. Which should we use? The answer depends on what we are trying to do. The "correct" action, with an easy physical interpretation, is S, we simply extremise the length of the world line. However, it is often easier to work with S_2, after all it doesn't have a square root in it, we simply have to fix $u^2 = c^2$ after we've done all of the variation. The slightly more general action $\tilde{S}$ has the advantage that we can use it for massless particles also by taking the limit $m \to 0$, although we can't then take $g(\lambda) = 2/m$ (which is equivalent to $\lambda = \tau$) but instead pick $g(\lambda) = \text{constant}$ (which is equivalent to picking some alternative affine parameter instead of τ).

Eleven

Action for Electromagnetic Fields

In this chapter we will motivate the action for the electromagnetic field. First, we motivate the use of a Lagrangian density, then we will study which terms could possible occur in the action if it is to have the desired symmetry, and finally we will demonstrate that we recover the correct equations of motion, namely Maxwell's equations. We will start by doing so for a free field with no charges or currents

11.1 Lagrangian Density

When we considered a particle we found that its action was of the form

$$S = \int L \, d\lambda. \tag{11.1.1}$$

That is an integral of the Lagrangian. The Lagrangian in turn depends on $x^\mu(\lambda)$, the position of the particle, and $\dot{x}^\mu(\lambda)$, the velocity of the particle, which depend on the single parameter λ. Hence, there is a single integral.

Now consider a string, it was shown in a tutorial question that the position of the string is described by $\varphi(x, t)$, which gives the deviation from the equilibrium position of the string at position x along the string and time t. This is equivalent to $x^\mu(\lambda)$ above, with φ being x^μ and both x and t being parameters like λ. This value, φ, then evolves under the action

$$S[\varphi(x, t)] = \int \left[\frac{1}{2} \left(\frac{\partial \varphi}{\partial t} \right)^2 - \frac{Y}{2} \left(\frac{\partial \varphi}{\partial x} \right)^2 \right] dt \, dx \tag{11.1.2}$$

where μ is the mass per unit length and Y is the spring constant per unit length. The important thing here is that the "path", φ, depends on two parameters and we integrate some Lagrangian-like quantity over both of them to get the action.

Now, consider electromagnetism. It will be our general goal to see how A^μ, the four-potential, transforms. In general we will consider cases where A^μ depends on the position in spacetime, so we consider it to be a function $A^\mu(x^0, x^1, x^2, x^3)$. We treat x^μ as parameters here, so we expect a four-dimensional integral. We therefore look for an action of the form

$$S[A^\mu(x)] = \int \mathcal{L} \, d^4x. \tag{11.1.3}$$

We call $\mathcal{L}$ the **Lagrangian density**[1] since we integrate over space (hence density)

[1] It's very common to just refer to the Lagrangian density as the Lagrangian as well.

to get something (the Lagrangian) which we integrate over time to give the action, that is

$$S = \int \mathcal{L} \, \mathrm{d}^4 x = \int \mathcal{L} \, \mathrm{d}^3 x \, \mathrm{d}x^0 = c \int L \, \mathrm{d}t \tag{11.1.4}$$

where we define

$$L = \int \mathcal{L} \, \mathrm{d}^3 x. \tag{11.1.5}$$

11.2 Possible Terms

In order to narrow down possibilities for $\mathcal{L}$ we need to decide what symmetries it should satisfy. First, we would like for $\mathcal{L}$ to be *Lorentz invariant*, since this guarantees that the equations of motion will be Lorentz invariant. Second, we would like for $\mathcal{L}$ to be *gauge invariant*, since this guarantees that the equations of motion will be gauge invariant. Third, for a free electromagnetic field we would like for $\mathcal{L}$ to be *translation invariant*.

Imposing these symmetries we can narrow down the possible terms in the Lagrangian density. First, we expect to have the volume element, $\mathrm{d}^4 x$. This Lorentz invariant and is trivially gauge invariant since it doesn't depend on the potential. It is also translation invariant, since it doesn't depend on position in flat spacetime. Second, we expect the electromagnetic tensor to appear in some way. In order for this to happen in a Lorentz invariant way we must some how get a scalar, the most obvious way to do this is $F_{\mu\nu}F^{\mu\nu}$. Since the electromagnetic tensor is gauge invariant this product will also be gauge invariant, and for a free field there is no position dependence so it will be translation invariant. Note that $F_{\mu\nu}^* F^{*\mu\nu}$ is simply proportional to $F_{\mu\nu}F^{\mu\nu}$, so doesn't give us anything new.

There are other possible terms too, for example, $F_{\mu\nu}^* F^{\mu\nu}$ satisfies these required symmetries. However, it doesn't have parity invariance, which it turns out is a required symmetry in electromagnetism. Other terms with the required symmetry properties include powers of $F_{\mu\nu}F^{\mu\nu}$, such as $(F_{\mu\nu}F^{\mu\nu})^2$, and derivatives of the electromagnetic field tensor, such as $\partial_\rho F_{\mu\nu} \partial^\rho F^{\mu\nu}$. However, these terms turn out to be less important in classical electrodynamics and instead appear as higher order corrections in quantum electrodynamics.

11.3 Action

We need the action to have units of energy times time, or more generally units of $\hbar$. This is fundamentally since the action appears in the path integral

$$\int \exp\left[\frac{i}{\hbar} S\right] \mathcal{D}x \tag{11.3.1}$$

To which we can apply the stationary phase approximation giving the principle of least action in the first place[2].

[2]See the notes from the quantum theory course.

Recall that $F_{\mu\nu}F^{\mu\nu} = -2\boldsymbol{E}^2 + 2\boldsymbol{B}^2$, and that $u = (\boldsymbol{E}^2 + \boldsymbol{B}^2)/2$ is the energy density. We know therefore that $F_{\mu\nu}F^{\mu\nu}$ has dimensions of energy per volume, or

$$[F_{\mu\nu}F^{\mu\nu}] = [\text{energy density}] = \frac{[\text{energy}]}{[\text{length}]^3}. \tag{11.3.2}$$

Now, integrating over the three spatial dimensions gets rid of the per volume part of these dimensions, leaving us with dimensions of energy, which we then integrate over x^0, to get dimensions of length times density:

$$\left[\int F_{\mu\nu}F^{\mu\nu}\,\mathrm{d}^4x\right] = \left[\int \mathrm{d}^4x\right]\frac{[\text{energy}]}{[\text{length}]^3} = \int[\text{energy}]\,\mathrm{d}x^0 = [\text{energy}][\text{length}].$$

So, if we include a factor of $1/c$ then we will get the desired units:

$$\left[\frac{1}{c}\int F_{\mu\nu}F^{\mu\nu}\,\mathrm{d}^4x\right] = \frac{[\text{time}]}{[\text{length}]}[\text{energy}][\text{length}] = [\text{energy}][\text{time}] = [\hbar].$$

This isn't quite the form of the action, there is a conventional factor of $-1/4$, so the action is defined as

$$S[A^\mu(x)] := \frac{1}{c}\int_V\left[-\frac{1}{4}F_{\mu\nu}F^{\mu\nu}\right]\mathrm{d}^4x \tag{11.3.3}$$

where V is some four-dimensional volume. We then define

$$\mathcal{L}(\partial^\mu A^\nu) = -\frac{1}{4c}F_{\mu\nu}F^{\mu\nu} \tag{11.3.4}$$

for the Lagrangian density.

We can also consider a more general action

$$S = \frac{1}{c}\int_V\mathcal{L}(A^\mu,\partial^\mu A^\nu)\,\mathrm{d}^4x. \tag{11.3.5}$$

It should be noted that the indices appearing in the argument to $\mathcal{L}$ are not meant to represent a particular value, they are simply a reminder that A^μ is a four-vector and $\partial^\mu A^\nu$ is a rank 2 tensor. This is known as abstract index notation, as opposed to Ricci calculus which we have mostly been using so far, where each index is simply a place holder for a particular coordinate.

11.4 Equations of Motion

*I'm just some guy, you have to
wonder why you want to believe me.*

Donal O'Connell

Consider the more general action defined above. We will vary the potential in order to find the equations of motion:

$$A_\mu(x) \to A_\mu(x) + \delta A_\mu(x). \tag{11.4.1}$$

We choose the variation so that, as usual, it vanishes on the boundary of V, denoted ∂V. The variation in the action is then

$$\delta S = \frac{1}{c}\int_V\delta\mathcal{L}(A^\mu,\partial^\mu A^\nu)\,\mathrm{d}^4x \tag{11.4.2}$$

$$= \frac{1}{c}\int_V\left[\frac{\partial\mathcal{L}}{\partial A^\mu}\delta A^\mu + \frac{\partial\mathcal{L}}{\partial(\partial^\mu A^\nu)}\delta\partial^\mu A^\nu\right]\mathrm{d}^4x. \tag{11.4.3}$$

We get this result by simply Taylor expanding in the two variables, A^μ and $\partial^\mu A^\nu$, it just looks a bit weird since one is written as a derivative, and we're actually packaging 20 terms into two with sums over repeated indices.

We apply the usual manipulation of writing the second term, with the variation of a derivative, as the derivative of a product minus the extra term this produces:

$$S = \frac{1}{c} \int_V \left[\frac{\partial \mathcal{L}}{\partial A^\mu} \delta A^\mu + \partial^\mu \left(\frac{\partial \mathcal{L}}{\partial(\partial^\mu A^\nu)} \delta A^\nu \right) - \partial^\mu \frac{\partial \mathcal{L}}{\partial(\partial^\mu A^\nu)} \delta A^\nu \right] \mathrm{d}^4 x. \qquad (11.4.4)$$

Now consider the term with derivative of the product:

$$\int_V \partial^\mu \left(\frac{\partial \mathcal{L}}{\partial(\partial^\mu A^\nu)} \delta A^\nu \right) \mathrm{d}^4 x. \qquad (11.4.5)$$

This is simply the divergence theorem in four dimensions:

$$\int_V \nabla \cdot \boldsymbol{v} \, \mathrm{d}^d x = \int_{\partial V} \boldsymbol{v} \cdot \mathrm{d}\boldsymbol{S} \iff \int_V \partial^\mu v_\mu \, \mathrm{d}^d x = \int_{\partial V} v_\mu \, \mathrm{d}S^\mu. \qquad (11.4.6)$$

Hence, applying the divergence theorem we have

$$\int_V \partial^\mu \left(\frac{\partial \mathcal{L}}{\partial(\partial^\mu A^\nu)} \delta A^\nu \right) \mathrm{d}^4 x = \int_{\partial V} \frac{\partial \mathcal{L}}{\partial(\partial^\mu A^\nu)} \delta A^\nu \, \mathrm{d}S^\mu. \qquad (11.4.7)$$

However, since δA^ν vanishes on ∂V this integral is identically zero and so this term does not contribute to the action.

The variation in the action is therefore

$$\delta S = \frac{1}{c} \int_V \left[\frac{\partial \mathcal{L}}{\partial A^\mu} \delta A^\mu - \partial^\mu \frac{\partial \mathcal{L}}{\partial(\partial^\mu A^\nu)} \delta A^\nu \right] \mathrm{d}^4 x. \qquad (11.4.8)$$

We can reindex the first term, replacing μ with ν, to get a factor of δA^ν in both terms which we then factor out:

$$\delta S = \frac{1}{c} \int_V \left[\frac{\partial \mathcal{L}}{\partial A^\nu} - \partial^\mu \frac{\partial \mathcal{L}}{\partial(\partial^\mu A^\nu)} \right] \delta A^\nu \, \mathrm{d}^4 x. \qquad (11.4.9)$$

Hence, requiring that $\delta S = 0$ according to the principle of least action and taking δA^ν to be arbitrary, other than the requirement of vanishing on ∂V, we must have that the square brackets vanish, giving the Euler–Lagrange equations for this form of Lagrangian:

$$\frac{\partial \mathcal{L}}{\partial A^\nu} = \partial^\mu \frac{\partial \mathcal{L}}{\partial(\partial^\mu A^\nu)}. \qquad (11.4.10)$$

Finally, we return to the case of

$$\mathcal{L} = -\frac{1}{4} F^{\mu\nu} F_{\mu\nu}. \qquad (11.4.11)$$

Recalling that $F^{\mu\nu}$ is defined in terms of $\partial^\mu A^\nu$, and doesn't depend on A^μ alone we have

$$\frac{\partial \mathcal{L}}{\partial A^\nu} = 0. \qquad (11.4.12)$$

We can also calculate the other side of the Euler–Lagrange equations fairly easily, we just have to be careful to avoid index clash. We start with

$$\frac{\partial \mathcal{L}}{\partial(\partial^\mu A^\nu)} = -\frac{1}{4}\frac{\partial}{\partial(\partial^\mu A^\nu)}(F_{\alpha\beta}F^{\alpha\beta}) = -\frac{1}{2}F_{\alpha\beta}\frac{\partial}{\partial(\partial^\mu A^\nu)}F^{\alpha\beta} \tag{11.4.13}$$

where we have used the freedom in raising and lowering paired indices to notice that both terms in the product rule will give the same result. We need only then compute the derivative once, which we do by plugging in the definition for $F^{\alpha\beta}$:

$$\frac{\partial}{\partial(\partial^\mu A^\nu)}F^{\alpha\beta} = \frac{\partial}{\partial(\partial^\mu A^\nu)}[\partial^\alpha A^\beta - \partial^\beta A^\alpha]. \tag{11.4.14}$$

Now notice that the derivative will vanish on each term, apart from when the indices match up correctly across both the derivatives and the potentials, in which case we get unity. Hence,

$$\frac{\partial}{\partial(\partial^\mu A^\nu)}F^{\alpha\beta} = \delta^\alpha{}_\mu\delta^\beta{}_\nu - \delta^\beta{}_\mu\delta^\alpha{}_\nu. \tag{11.4.15}$$

Going back to one side of the Euler–Lagrange equations we have

$$\frac{\partial \mathcal{L}}{\partial(\partial^\mu A^\nu)} = -\frac{1}{2}F_{\alpha\beta}[\delta^\alpha{}_\mu\delta^\beta{}_\nu - \delta^\beta{}_\mu\delta^\alpha{}_\nu] \tag{11.4.16}$$

$$= -\frac{1}{2}[F_{\mu\nu} - F_{\nu\mu}] \tag{11.4.17}$$

$$= -F_{\mu\nu}, \tag{11.4.18}$$

where we have used the antisymmetry of the electromagnetic tensor in the last step.

Hence, the full Euler–Lagrange equations give

$$0 = \partial^\mu\frac{\partial \mathcal{L}}{\partial(\partial^\mu A^\nu)} = -\partial^\mu F_{\mu\nu}. \tag{11.4.19}$$

This is exactly the Maxwell equation for a free electromagnetic field: $\partial^\mu F_{\mu\nu} = 0$. This justifies our choice of the action, including neglecting higher order terms.

Twelve

Interactions

12.1 Action for A Current

Suppose we have a current density, $J^\mu(x)$, we can modify our action for the free electromagnetic field by adding an extra interaction term. The resulting action is

$$S = \frac{1}{c} \int \left[-\frac{1}{4} F^{\mu\nu} F_{\mu\nu} - \frac{1}{c} J^\mu A_\mu \right] d^4 x \tag{12.1.1}$$

$$= \frac{1}{c} \int \left[-\frac{1}{4} F^{\mu\nu} F_{\mu\nu} \right] d^4 x + \frac{1}{c} \int \left[-\frac{1}{c} J^\mu A_\mu \right] d^4 x \tag{12.1.2}$$

$$= S_{\text{free}} + S_{\text{int}}. \tag{12.1.3}$$

In terms of the Lagrangian density we have

$$\mathcal{L} = -\frac{1}{4} F^{\mu\nu} F_{\mu\nu} - \frac{1}{c} J^\mu A_\mu \tag{12.1.4}$$

$$= \mathcal{L}_{\text{free}} + \mathcal{L}_{\text{int}}. \tag{12.1.5}$$

12.2 Dimensions

The first step in justifying this choice is to check it works dimensionally. In order to do this it is sufficient to show that $\mathcal{L}_{\text{int}}$ has the same dimensions as $\mathcal{L}_{\text{free}}$. We can also ignore any indices, since they don't change the dimensions. That is, we want to show that $[FF] = [JA/c]$. To do this we use the fact that $F^{\mu\nu}$ is defined in terms of derivatives of A^μ, so $[F] = [\partial A] = L^{-1}[A]$. Hence, $[FF] = L^{-1}[FA] = L^{-1}[A][F]$. We also know that c is a speed, so $[JA/c] = L^{-1}T[JA] = L^{-1}[A]T[J]$. So, things work out dimensionally as long as $T[J] = [F]$. Recall that Maxwell's equation gives us $\partial^\mu F_{\mu\nu} = J_\nu/c$, and so we have $[\partial F] = L^{-1}[F] = [J/c] = L^{-1}T[J]$. Thus, $[F] = T[J]$, and so the dimensions work.

12.3 Equations of Motion

The next step in justifying this choice is to show that we get the correct equations of motion, both for the fields and for particles. We'll start with the field.

12.3.1 Equations of Motion for the Field

The Euler–Lagrange equations for a Lagrangian density of the form $\mathcal{L}(A^\mu, \partial^\mu A^\nu)$ are

$$\partial^\mu \frac{\partial \mathcal{L}}{\partial(\partial^\mu A^\nu)} = \frac{\partial \mathcal{L}}{\partial A^\nu}. \tag{12.3.1}$$

Since the interaction term doesn't introduce any derivatives of A^μ we know from Equation (11.4.19) that

$$\partial^\mu \frac{\partial \mathcal{L}}{\partial(\partial^\mu A^\nu)} = -\partial^\mu F_{\mu\nu}. \tag{12.3.2}$$

However, the interaction term introduces a term with A^μ, so the other side of the Euler–Lagrange equations no longer vanishes. Instead,

$$\frac{\partial \mathcal{L}}{\partial A^\nu} = \frac{\partial}{\partial A^\nu}\left[-\frac{1}{c}J^\mu A_\mu\right] \tag{12.3.3}$$

$$= -\frac{1}{c}\frac{\partial}{\partial A^\nu}J^\mu A_\mu \tag{12.3.4}$$

$$= -\frac{1}{c}\frac{\partial}{\partial A^\nu}J_\mu A^\mu \tag{12.3.5}$$

$$= -\frac{1}{c}J_\mu \frac{\partial A^\mu}{\partial A^\nu} \tag{12.3.6}$$

$$= -\frac{1}{c}J_\mu \delta^\mu{}_\nu \tag{12.3.7}$$

$$= -\frac{1}{c}J_\nu. \tag{12.3.8}$$

Hence, we have the equation of motion for the fields given by

$$\partial^\mu F_{\mu\nu} = \frac{1}{c}J_\nu, \tag{12.3.9}$$

which is exactly the inhomogeneous Maxwell equation.

At this point one might wonder how to derive the homogeneous Maxwell equation, $\partial^\mu F^*_{\mu\nu} = 0$. The answer is that we don't really need to do this because we have used $F^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu$ as a definition and this makes the homogeneous Maxwell equation trivially true. The two Maxwell equations aren't on quite the same footing in terms of importance, the inhomogeneous one contains most of the physics. This is similar to how the divergence equations in the non-relativistic version are not strictly necessary, since they follow from the curl equations and the identity $\nabla \cdot \nabla \times \boldsymbol{V} = 0$. On the physics side we can view this as being a consequence of the fact that it is A^μ, not $F^{\mu\nu}$, which is the fundamental variable appearing, for example, in the path integrals of QFT.

12.3.2 Gauge Invariance

> *It's actually very interesting, well, I say very interesting. I very quickly get interested in things like this. That's why I'm here.*
>
> Donal O'Connel

When justifying the choice of the free action we said that we were looking for something gauge invariant, and we found it by working with the gauge invariant electromagnetic field tensor. The interaction term has an explicit A_μ, which is not manifestly gauge invariant. Recall that a gauge transform takes the form of

$A_\mu(x) \to A'_\mu(x) = A_\mu(x) + \partial_\mu \chi(x)$, for some smooth scalar field χ. Using this we can show that the action is gauge invariant:

$$S_\text{int} \to \frac{1}{c} \int \left[-\frac{1}{c} J^\mu (A_\mu + \partial_\mu \chi) \right] \mathrm{d}^4 x \tag{12.3.10}$$

$$= S_\text{int} - \frac{1}{c^2} \int \left[J^\mu \partial_\mu \chi \right] \mathrm{d}^4 x \tag{12.3.11}$$

$$= S_\text{int} - \frac{1}{c^2} \int \left[\partial_\mu (J^\mu \chi) - (\partial_\mu J^\mu) \chi \right] \mathrm{d}^4 x. \tag{12.3.12}$$

Now, considering the first term we see that an application of the divergence theorem gives

$$\int_V \partial_\mu (J^\mu \chi) \mathrm{d}^4 x = \int_{\partial V} J^\mu \chi \, \mathrm{d}S_\mu \tag{12.3.13}$$

where S_μ is a normal vector to the boundary ∂V. Now, whenever we vary the action we always do so in such a way that variations on the boundary vanish, and hence this term doesn't contribute to the variation, and so doesn't effect the equations of motion.

The second term has $\partial_\mu J^\mu$, which vanishes by the continuity equation if J^μ is a conserved current. This hints at a link between the symmetry of gauge invariance and the conservation of current, which is a consequence of Noether's theorem.

The result is that while the action is not strictly gauge invariant the variation in the action is gauge invariant, and that's all we care about. More formally the action is gauge invariant up to a surface term if the current is conserved.

At this point we note that often we use terminology inherited from Lagrangian dynamics. In particular we might call A^μ a generalised coordinate, and the quantity

$$\frac{\partial \mathcal{L}}{\partial (\partial^\mu A^\nu)} \tag{12.3.14}$$

is the canonical momentum conjugate to A^μ.

12.3.3   Relativistic Interacting Point Particle

We have seen that we get the inhomogeneous Maxwell equation as the equation of motion of the field. Before we accept this form of the action we should show that it also recovers the correct particle dynamics, namely the Lorentz force law.

Suppose we have some particle of mass m and charge q. The four current splits into a temporal part and spatial part according to $J^\mu(x) = (\rho(x)c, \mathbf{J}(x))$. For a point particle the charge density is simply a Dirac delta distribution,

$$\rho(x) = q\delta^3(\mathbf{x} - \mathbf{r}(t)). \tag{12.3.15}$$

The superscript 3 is because this is a delta distribution in three dimensions. The charge density depends on the four-position, $x^\mu = (ct, \mathbf{x})$. This is the four-position of the point at which we are measuring the charge density, not the position of the particle, which is $r^\mu = (ct', \mathbf{r})$. The spatial part of the current density for a moving charge distribution is $\mathbf{J}(x) = \rho(x)\mathbf{v}(t)$. Therefore we can write the four-current as

$$J^\mu(x) = q\delta^3(\mathbf{x} - \mathbf{r}(t))(c, \mathbf{v}(t))^\mu. \tag{12.3.16}$$

This is not manifestly Lorentz covariant, but we know that J^μ is a contravariant four-vector, and so we should be able to massage this into a form making it manifestly Lorentz covariant. The first step is to get rid of the three-dimensional Dirac delta in favour for a four dimensional one. We can do this by using the fact that if we integrate a four-dimensional delta distribution over time we are left with a three-dimensional spatial delta-distribution. Let's introduce $r^0(t') = ct'$ as the time component of the particle's four-position. We therefore have

$$J^\mu(x) = q\delta^3(\boldsymbol{x} - \boldsymbol{r}(t))(c, \boldsymbol{v}(t))^\mu \tag{12.3.17}$$

$$= q \int \mathrm{d}(ct') \, \delta(ct - ct')\delta^3(\boldsymbol{x} - \boldsymbol{r}(t'))\frac{\mathrm{d}r^\mu(t')}{\mathrm{d}t'} \tag{12.3.18}$$

$$= qc \int \mathrm{d}t' \, \delta^4(x - r(t'))\frac{\mathrm{d}r^\mu}{\mathrm{d}t'}. \tag{12.3.19}$$

Here we have written the four-velocity of the particle as a derivative of its four-position. We can choose to use the proper time of the particle as the time parameter, and we therefore get

$$J^\mu(x) = qc \int \mathrm{d}\tau \frac{\mathrm{d}r^\mu}{\mathrm{d}\tau}\delta^4(x - r(\tau)). \tag{12.3.20}$$

It should be noted that a change of variables ("cancelling" the $\mathrm{d}\tau$ in the derivative with the measure of the integral) gives us

$$J^\mu(x) = qc \int \mathrm{d}r^\mu \, \delta^4(x - r). \tag{12.3.21}$$

Either of these two forms is manifestly Lorentz covariant, which is what we wanted.

Now consider the total action of the fields and the particle, this can be written as

$$S_{\text{tot}} = \frac{1}{c} \int \left[-\frac{1}{4}F^{\mu\nu}F_{\mu\nu}\right] \mathrm{d}^4x - mc \int \sqrt{\frac{\mathrm{d}r}{\mathrm{d}\lambda} \cdot \frac{\mathrm{d}r}{\mathrm{d}\lambda}} \, \mathrm{d}\lambda + \frac{1}{c} \int \left[-\frac{1}{c}J^\mu A_\mu\right] \mathrm{d}^4x.$$

The first term is the action for a free electromagnetic field, the second term the action for a free particle, and the third accounts for the interactions between the field and the particle. It is this interacting term that we are interested in. In more detail we can write it as

$$S_{\text{int}} = \frac{1}{c} \int \mathrm{d}^4x \left[-\frac{1}{c}J^\mu A_\mu\right] \tag{12.3.22}$$

$$= -\frac{1}{c} \int \mathrm{d}^4x \, qc \int \mathrm{d}\tau \frac{\mathrm{d}r^\mu}{\mathrm{d}\tau}\delta^4(x - r(t))A_\mu(x) \tag{12.3.23}$$

$$= -\frac{q}{c} \int \mathrm{d}\tau \frac{\mathrm{d}r^\mu}{\mathrm{d}\tau}A_\mu(r(\tau)). \tag{12.3.24}$$

In the last step we swapped the integrals to perform the integral over d^4x first using the sifting property of the delta distribution. This leaves us with a term which depends on the four-potential at the position of the particle, this means that there is no interaction at a distance, which is what we would expect.

We can reparametrise the interaction action as

$$S_{\text{int}} = -\frac{q}{c} \int \mathrm{d}\lambda \frac{\mathrm{d}r^\mu(\lambda)}{\mathrm{d}\lambda}A_\mu(r(\lambda)). \tag{12.3.25}$$

The total action is then

$$S_{\text{tot}} = \frac{1}{c} \int \mathrm{d}^4 x \left[-\frac{1}{4} F^{\mu\nu} F_{\mu\nu} \right] - mc \int \mathrm{d}\lambda \sqrt{\frac{\mathrm{d}r}{\mathrm{d}\lambda} \cdot \frac{\mathrm{d}r}{\mathrm{d}\lambda}} - \frac{q}{c} \int \mathrm{d}\lambda \frac{\mathrm{d}r^\mu}{\mathrm{d}\lambda} A_\mu(r(\lambda)).$$

$$(12.3.26)$$

We can then find the equations of motion for the particle using the appropriate Euler-Lagrange equations, namely

$$\frac{\mathrm{d}}{\mathrm{d}\lambda} \frac{\mathrm{d}L}{\mathrm{d}\left(\frac{\mathrm{d}r^\mu}{\mathrm{d}\lambda}\right)} = \frac{\mathrm{d}L}{\mathrm{d}r^\mu} \tag{12.3.27}$$

where L is the integrand of the total action after performing all spatial integrals. There is no dependence on r^μ or $\mathrm{d}r^\mu/\mathrm{d}\lambda$ in the first term, so it doesn't contribute to the particle's equations of motion, which makes sense as that term is only to do with the electromagnetic field.

Performing the derivative of L we have

$$\frac{\mathrm{d}}{\mathrm{d}\lambda} \left[-mc \frac{1}{\sqrt{\frac{\mathrm{d}r}{\mathrm{d}\lambda} \cdot \frac{\mathrm{d}r}{\mathrm{d}\lambda}}} \frac{\mathrm{d}r^\mu}{\mathrm{d}\lambda} - \frac{q}{c} A^\mu(r(\lambda)) \right] = -\frac{q}{c} \frac{\mathrm{d}r_\nu}{\mathrm{d}\lambda} \partial^\mu A^\nu(r(\lambda)). \tag{12.3.28}$$

Taking $\lambda = \tau$ again we have that $(\mathrm{d}r/\mathrm{d}\tau)^2 = c^2$, and so computing the derivative with respect to λ we are left with

$$-m \frac{\mathrm{d}^2 r^\mu}{\mathrm{d}\tau^2} - \frac{q}{c} \partial^\nu A^\mu(r(\tau)) \frac{\mathrm{d}r_\nu}{\mathrm{d}\tau} = -\frac{q}{c} \partial^\mu A^\nu(r(\tau)) \frac{\mathrm{d}r_\nu}{\mathrm{d}\tau}. \tag{12.3.29}$$

Collecting the derivatives of the potential together we see that we get the electromagnetic tensor and are left with the final equation of motion

$$m \frac{\mathrm{d}^2 r^\mu}{\mathrm{d}\tau^2} = \frac{q}{c} F^{\mu\nu}(r(\tau)) \frac{\mathrm{d}r_\nu}{\mathrm{d}\tau}. \tag{12.3.30}$$

We can then identify the left hand side as the four-force and the whole equation is the Lorentz force law, in a covariant form.

We have seen two forms of the total action here. The second one, with the interaction in terms of λ, is good for deriving the particle's equation of motion. The first form, with the interaction in terms of the position, is better for the equations of motion of the field. These are unaffected by the particle's presence since A^μ and its derivative don't appear in the action for the free particle.

Thirteen

Remarks and Extensions

13.1 Recap

The total action for the electromagnetic field with a charged particle is

$$S = \frac{1}{c} \int \mathrm{d}^4 x \left[-\frac{1}{4} F^{\mu\nu} F_{\mu\nu} \right] - mc \int \mathrm{d}\lambda \sqrt{\frac{\mathrm{d}r}{\mathrm{d}\lambda} \cdot \frac{\mathrm{d}r}{\mathrm{d}\lambda}} + S_{\mathrm{int}} \tag{13.1.1}$$

where

$$S_{\mathrm{int}} = \frac{1}{c} \int \mathrm{d}^4 x \left[-\frac{1}{c} J^\mu(x) A_\mu(x) \right] \tag{13.1.2}$$

$$= -\frac{e}{c} \int \mathrm{d}\lambda \frac{\mathrm{d}r^\mu}{\mathrm{d}\lambda} A_\mu(r(\lambda)). \tag{13.1.3}$$

If we vary S with respect to A_μ then we get the Maxwell equations, which tell us how the field evolves. If we vary with respect to r^μ then we get the Lorentz force law, which tells us how the particle's motion evolves. In both cases we can instead use the appropriate Euler–Lagrange equations instead of varying the action explicitly. They are

$$\partial^\mu \frac{\partial \mathcal{L}}{\partial (\partial^\mu A^\nu)} = \frac{\partial \mathcal{L}}{\partial A^\nu}, \qquad \text{and} \qquad \frac{\mathrm{d}}{\mathrm{d}\lambda} \frac{\partial L}{\partial \left(\frac{\partial r^\mu}{\partial \lambda} \right)} = \frac{\partial L}{\partial r^\mu}. \tag{13.1.4}$$

After this we are free to set $\lambda = \tau$ and use simplifications like $(\mathrm{d}r/\mathrm{d}\lambda)^2 = c^2$. The two Lagrangian quantities, the Lagrangian, L, and Lagrangian density, $\mathcal{L}$, are related by

$$L = \int \mathcal{L} \, \mathrm{d}^3 x. \tag{13.1.5}$$

13.2 Radiation Reaction

The first form of the interaction, in terms of x, is good for deriving the Maxwell equations with a charged particle. The second form, with λ, is good for deriving the Lorentz force law. When we did this we got the Lorentz force law in the form

$$\frac{\mathrm{d}p^\mu}{\mathrm{d}\tau} = \frac{e}{c} F^{\mu\nu}(r(\tau)) \frac{\mathrm{d}r_\nu}{\mathrm{d}\tau}. \tag{13.2.1}$$

This can be done either by varying the action or applying the Euler–Lagrange equations.

The $F^{\mu\nu}$ here is the total field tensor, including the contribution to the field from the particle. This is different to how we usually think about the fields due to point particles, where we can ignore the Coulomb field of the particle in question. The field diverges at the position of the particle, causing problems, much like those we have to fix by renormalisation in QFT.

We cannot completely ignore the field from the particle though, if we do we miss something called the radiation reaction, which we will cover later in our discussion of radiation.

13.3 Alternative Terms

The form of the action is determined using two constraints, symmetries of the system and dimensional analysis. These are very powerful tools and allow us to specify actions which align well with the observed physics. However, they can't rule out all possible terms, but with some simple assumptions dimensional analysis can suggest the most important terms.

13.3.1 Why No Second Order Term?

> *Rigorous? This is physics. There is no rigour.*
>
> Donal O'Connel

We chose not to include a term involving $(F^{\mu\nu}F_{\mu\nu})^2$ in our action. This is justified by the end results, but can also be argued for by dimensional analysis. Suppose we did have such a term. It would be of the form

$$\frac{1}{c}\int \mathrm{d}^4x\, \kappa(F^{\mu\nu}F_{\mu\nu})^2 \tag{13.3.1}$$

where κ is a dimensionfull quantity such that the above has units of action. Comparing to the first order term we see that we have

$$\left[\frac{1}{c}\int \mathrm{d}^4x\, \kappa(F^{\mu\nu}F_{\mu\nu})^2\right] = \left[\frac{1}{c}\int \mathrm{d}^4x\left(-\frac{1}{4}F^{\mu\nu}F_{\mu\nu}\right)\right] \tag{13.3.2}$$

where square brackets denote "dimensions of". In particular

$$[\kappa F^{\mu\nu}F_{\mu\nu}] = 1. \tag{13.3.3}$$

Hence, we have

$$[\kappa] = \frac{1}{[F^{\mu\nu}F_{\mu\nu}]}. \tag{13.3.4}$$

Now, we know that

$$[S] = \left[\frac{1}{c}\int \mathrm{d}^4x\left(-\frac{1}{4}F^{\mu\nu}F_{\mu\nu}\right)\right] \tag{13.3.5}$$

so we have that

$$[F^{\mu\nu}F_{\mu\nu}] = \frac{[S][c]}{[\int \mathrm{d}^4x]} = \frac{M}{LT^2}, \tag{13.3.6}$$

where we've used the fact that S has dimensions of angular momentum, $[S] = ML^2T^{-1}$, $[c] = LT^{-1}$, and $[\int d^4x] = L^4$.

Combing this we have

$$[\kappa] = \frac{LT^2}{M} = \left(\frac{T}{L}\right)^2 \frac{L^3}{M}. \tag{13.3.7}$$

At this point we make the assumption that κ is not a new constant, rather it is built from other physical constants. In particular we assume that the dimensions of $(T/L)^2$ come from a factor of c^2, and the dimension of M comes from the mass of the particle, m. We are then just left with the dimensions of L^3. It turns out that there is a natural length scale in classical electrodynamics, called the **classical radius of an electron**. This is defined as the length, ℓ_c, such that the rest energy of the electron is equal to the electric potential from a pair of electrons:

$$\frac{e^2}{4\pi\ell_c} = mc^2 \implies \ell_c = \frac{e^2}{4\pi mc^2} \approx 2.82 \times 10^{-15} \text{ m}. \tag{13.3.8}$$

The idea is that, before the advent of quantum mechanics, one suggestion was that the mass of the electron was provided by the potential energy. Note that this quantity is on the order of length of one nuclear diameter.

We then posit that

$$\kappa = \frac{k\ell_c^3}{mc^2} \tag{13.3.9}$$

where k is some dimensionless number. Further, we assume that k is approximately on the order of 1, and so the resulting value of κ is

$$\kappa \sim \frac{(10^{-15} \text{ m})^3}{(10^{-31} \text{ kg})(10^8 \text{ m s}^{-1})^2} = 10^{-31} \text{ m s}^2 \text{ kg}^{-1}. \tag{13.3.10}$$

The exact value isn't important, the point is that, unless k is very large, κ is very *very* small. Hence, if there is an $(F^{\mu\nu}F_{\mu\nu})^2$ term in the action then it is suppressed by a factor of 10^{-31} compared to the $F^{\mu\nu}F_{\mu\nu}$ term.

This analysis relies on two assumptions, first, that κ depends only on other constants which are already relevant to classical electrodynamics, and second that k is of order 1. However, since we get reasonable results without need for an $(F^{\mu\nu}F_{\mu\nu})^2$ term this seems to be the case.

In principle higher order terms, including but not limited to an $(F^{\mu\nu}F_{\mu\nu})^2$ term, are present in a full quantum mechanical description, such as QED. This is in part because quantum mechanics brings in other dimensionfull quantities, such as $\hbar$, from which we can construct κ.

The Compton wavelength, λ_c, of the electron is the appropriate length scale in quantum mechanics. It is larger than the classical electron radius by a factor of α, the fine structure constant:

$$\ell_c = \alpha\lambda_c = \frac{e^2}{4\pi\hbar c}\lambda_c \approx \frac{\lambda_c}{137} \approx 2.43 \times 10^{-12} \text{ m}. \tag{13.3.11}$$

What this means is that quantum effects become important as distances decrease long before questions of the structure of the electron.

13.3.2　Theta Term

We left out one term with the correct dimensions to be included in the action. This is called the **theta term**:

$$S_\vartheta = \frac{1}{c}\vartheta \int \mathrm{d}^4 x\, F_{\mu\nu} F^{*\mu\nu}. \tag{13.3.12}$$

Here ϑ is a dimensionless constant. We justified leaving this out through claims of "parity conservation". We can actually show that this term doesn't have any classical effect because it can be reduced to a boundary term. Start by inserting the definition of $F^{*\mu\nu}$:

$$S_\vartheta = \frac{1}{c}\vartheta \int \mathrm{d}^4 x \left[-\frac{1}{2}\varepsilon^{\mu\nu\rho\sigma} F_{\mu\nu} F_{\rho\sigma} \right]. \tag{13.3.13}$$

Then insert the definition of $F_{\mu\nu}$ and $F_{\rho\sigma}$ in terms of the potential:

$$S_\vartheta = -\frac{\vartheta}{2c} \int \mathrm{d}^4 x\, \varepsilon^{\mu\nu\rho\sigma} (\partial_\mu A_\nu - \partial_\nu A_\mu)(\partial_\rho A_\sigma - \partial_\sigma A_\rho). \tag{13.3.14}$$

Now, notice that each term in brackets is antisymmetric in the two indices, and so is the Levi-Civita symbol. So, swapping the indices in the second term in each bracket, expanding the brackets, relabelling swapping the indices again, and then swapping the indices on the Levi-Civita symbol cancels the minus sign giving the first term in the brackets again. Each bracket therefore contributes a factor of 2. That is,

$$\varepsilon^{\mu\nu\rho\sigma}(\partial_\mu A_\nu - \partial_\nu A_\mu) = \varepsilon^{\mu\nu\rho\sigma}\partial_\mu A_\nu - \underbrace{\varepsilon^{\mu\nu\rho\sigma}\partial_\nu A_\mu}_{\mu\leftrightarrow\nu} \tag{13.3.15}$$

$$= \varepsilon^{\mu\nu\rho\sigma}\partial_\mu A_\nu - \varepsilon^{\nu\mu\rho\sigma}\partial_\mu A_\nu \tag{13.3.16}$$

$$= \varepsilon^{\mu\nu\rho\sigma}\partial_\mu A_\nu + \varepsilon^{\mu\nu\rho\sigma}\partial_\mu A_\nu \tag{13.3.17}$$

$$= 2\varepsilon^{\mu\nu\rho\sigma}\partial_\mu A_\nu. \tag{13.3.18}$$

Hence, doing this for both terms we get

$$S_\vartheta = -\frac{2\vartheta}{c} \int \mathrm{d}^4 x\, \varepsilon^{\mu\nu\rho\sigma} (\partial_\mu A_\nu)(\partial_\rho A_\sigma). \tag{13.3.19}$$

Now, pull the ∂_μ out the front of the term and apply it to all that remains. Using the product rule this will give an extra term, which we subtract away:

$$S_\vartheta = -\frac{2\vartheta}{c} \int \mathrm{d}^4 x [\varepsilon^{\mu\nu\rho\sigma}\partial_\mu(A_\nu \partial_\rho A_\sigma) - \varepsilon^{\mu\nu\rho\sigma} A_\nu \partial_\mu \partial_\rho A_\sigma]. \tag{13.3.20}$$

Notice that $\varepsilon^{\mu\nu\rho\sigma}$ is antisymmetric in μ and ρ, whereas $\partial_\mu \partial_\rho$ is symmetric in μ and ρ. Hence, their product, and thus the second term, vanishes, leaving

$$S_\vartheta = -\frac{2\vartheta}{c} \int \mathrm{d}^4 x\, \varepsilon^{\mu\nu\rho\sigma}\partial_\mu(A_\nu \partial_\rho A_\sigma). \tag{13.3.21}$$

Finally, an application of the divergence theorem results in an integral over the boundary of the volume of integration above, and hence this doesn't contribute to the variation which vanishes on the boundary.

The result of all of this is that S_ϑ is irrelevant classically. It turns out that S_ϑ also doesn't have an effect in QED. However, there is an analogous term in QCD which does have an effect, although a small one with $\vartheta < 10^{-14}$. This leads to the "strong CP problem", in which the CP violation due to this theta term happens to cancel with CP violation due to other terms in the Lagrangian resulting in a theory that preserves CP symmetry overall. One solution to this fine tuning problem is to introduce a new particle, called an **axion**, which break the symmetry. Axions are also considered to be a dark matter candidate.

13.4 Actions are Flexible

We like the Higgs here at Edinburgh

Donal O'Connel

We considered a vector force. This is demonstrated by the fact that $\partial^\mu F_{\mu\nu} = J_\nu/c$ has only a single free index. In the Lorenz gauge we have the even simpler

$$\partial^2 A_\mu = \frac{1}{c} J_\mu. \tag{13.4.1}$$

What happens if instead of a vector field we have a scalar field? The free case can be modelled on the free EM case, we replace $\partial^2 A_\mu = 0$ with $\partial^2 \varphi = 0$, where φ is our scalar field. Introducing this field creates a long range effect, which is not observed. A slight alteration is

$$\partial^2 \varphi + \frac{1}{\ell^2} \varphi = 0 \tag{13.4.2}$$

where ℓ is a constant with units of length. This is, for example, applicable to the Higgs.

In further analogy with electromagnetism we can add an interation due to the motion of some particle at position $r(\lambda)$:

$$\partial^2 \varphi + \frac{1}{\ell^2} \varphi = y \int d\lambda \, \delta^4(x - r(\lambda)). \tag{13.4.3}$$

Here the interaction term is modelled after the interaction term in EM, in particular, the motion of a particle of charge y (up to some factor of c) moving along the path parametrised by λ. We call this "charge" y after Yukawa.

In the static case, with the charge fixed at the origin we can set $\lambda = \tau = t$ and this reduces to

$$-\nabla^2 \varphi + \frac{1}{\ell^2} \varphi = y \delta^3(\boldsymbol{x}). \tag{13.4.4}$$

The solution to this is the **Yukawa potential**

$$\varphi \propto \frac{1}{4\pi|\boldsymbol{x}|} e^{-|\boldsymbol{x}|/\ell}. \tag{13.4.5}$$

This is the usual Coulomb potential suppressed exponentially, justifying our claim that the extra φ/ℓ^2 term leads to a short range effect.

The action for the free scalar field would be

$$S_{\text{free}} = \frac{1}{c} \int \mathrm{d}^4 x \left[\frac{1}{2}(\partial\varphi)^2 - \frac{1}{2\ell^2}\varphi^2 \right]. \tag{13.4.6}$$

Note that $(\partial\varphi)^2 = (\partial_\mu\varphi)(\partial^\mu\varphi)$. The interaction term with the particle is then

$$S_{\text{int}} = -y \int \mathrm{d}\lambda, \; \varphi(r(\lambda)). \tag{13.4.7}$$

This is analogous to the EM interaction term

$$S_{\text{int}}^{\text{EM}} = -\frac{e}{c} \int \mathrm{d}\lambda \, \frac{\mathrm{d}r^\mu}{\mathrm{d}\lambda} A_\mu(r(\lambda)). \tag{13.4.8}$$

13.4.0.1   Action Lego

> *Actions are like Lego, you just keep adding terms.*
>
> Donal O'Connel

Actions are very flexible. Every new field we come up with can be added into the action, and all possible interactions between fields and particles can be accounted for. As an example we'll create a rather complicated action as follows:

$$S_{\text{tot}} = S_{\text{free}}^{\text{EM}} + S_{\text{free}}^{\text{particle}} + S_{\text{int}}^{\text{EM}} + S_{\text{free}}^{\text{scalar}} + S_{\text{free}}^{\text{Higgs}}$$
$$+ S_{\text{int}}^{\text{scalar/EM}} + S_{\text{free}}^{\text{dark}} + S_{\text{free}}^{\text{dark particle}} + S_{\text{int}}^{\text{dark}}. \tag{13.4.9}$$

Fundamentally there are two types of terms here, free terms describing fields, and interaction terms describing the interaction of either particles and fields, or two different fields. The first term is the free electromagnetic field term:

$$S_{\text{free}}^{\text{EM}} = \frac{1}{c} \int \mathrm{d}^4 x \left[-\frac{1}{4} F^{\mu\nu} F_{\mu\nu} \right]. \tag{13.4.10}$$

The second term is the usual term for a free particle, although we can extend it to n particles interacting with the EM field, where each has mass m_i and position r_i:

$$S_{\text{particle}}^{\text{EM}} = \sum_{i=1}^{n} \int \mathrm{d}\tau \, \frac{1}{2} m_i \left(\frac{\mathrm{d}r_i}{\mathrm{d}\tau} \right)^2. \tag{13.4.11}$$

Next, is the interaction between each of these particles and the EM field, with each particle contributing a current J_i:

$$S_{\text{int}}^{\text{EM}} = \sum_{i=1}^{n} \frac{1}{c} \int \mathrm{d}^4 x \, J_i^\mu(x) A_\mu(x). \tag{13.4.12}$$

Notice that the potential, A_μ, is the combined potential due to all particles. Now we include our free scalar field:

$$S_{\text{free}}^{\text{scalar}} = \frac{1}{c} \int \mathrm{d}^4 x \left[\frac{1}{2}(\partial\varphi)^2 - \frac{1}{2\ell^2}\varphi^2 \right]. \tag{13.4.13}$$

We can then have each of our particles also carry a Yukawa charge, y_i, interacting with the scalar field and contributing a term

$$S_{\text{free}}^{\text{Higgs}} = -\sum_{i=1}^{n} y_i \int \mathrm{d}\tau \, \varphi(r_i(\tau)). \tag{13.4.14}$$

Now we have two fields, the EM field and the scalar field, and we can have these interact:

$$S_{\text{int}}^{\text{scalar/EM}} = \int \mathrm{d}^4 x \, \kappa \varphi F^{\mu\nu} F_{\mu\nu}, \tag{13.4.15}$$

where κ is a dimensionfull constant giving overall dimensions of action, so $[\kappa] = [\varphi]/[cS]$.

Now suppose we want to include some dark matter, and that it behaves a lot like electromagnetism. We can introduce a "dark EM field", $\mathcal{B}_{\mu\nu}$, defined analogously to the EM field:

$$\mathcal{B}_{\mu\nu} = \partial_\mu B_\nu - \partial_\nu B_\mu, \tag{13.4.16}$$

where B_μ is some sort of "dark potential", analogous to A_μ. This "dark EM field" will then contribute a free term

$$S_{\text{free}}^{\text{dark}} = \int \mathrm{d}^4 x \left[-\frac{1}{4} \mathcal{B}^{\mu\nu} \mathcal{B}_{\mu\nu} + \frac{1}{2L^2} B_\mu B^\mu \right]. \tag{13.4.17}$$

Here L is another constant with dimensions of length and the second term makes this a short ranged field. We can then include dark matter particles, which don't interact with the EM field or scalar field, but do interact with the "dark EM field". Suppose there are N of them, with mass M_i and position R_i, then they will contribute the following term:

$$S_{\text{free}}^{\text{dark particle}} = \sum_{i=1}^{N} \int \mathrm{d}\tau \, \frac{1}{2} M_i \left(\frac{\mathrm{d}R_i}{\mathrm{d}\tau} \right)^2. \tag{13.4.18}$$

As well as some interaction term, $S_{\text{int}}^{\text{dark}}$ similar to $S_{\text{int}}^{\text{EM}}$:

$$S_{\text{int}}^{\text{dark}} = \sum_{i=1}^{N} \frac{1}{c} \int \mathrm{d}^4 x \, K_i^\mu(x) B_\mu(x) \tag{13.4.19}$$

where K_i^μ is a "dark current" analogous to J_i^μ.

We can keep going like this, adding in fields, particles, and interactions, until we've covered every required term.

Part IV

Green's Functions

Fourteen

Green's Functions

14.1 Motivation

We've now derived the Maxwell equations:

$$\partial^\mu F_{\mu\nu}(x) = \frac{1}{c} J_\nu(x). \tag{14.1.1}$$

In the Lorenz gauge, $\partial^\nu A_\nu = 0$, we've seen that these become

$$\partial^2 A_\mu = \frac{e}{c} \int u_\nu(x) \delta^4(x - r(\tau)) \, d\tau. \tag{14.1.2}$$

Here u is the four velocity, and δ^4 is a four dimensional Dirac delta. This is a wave equation with the right hand side being a driving force. The wave equation can be solved in many ways, but one of the most common is to use Green's functions.

14.2 General Theory

R For an introduction to Green's functions and Dirac deltas see the notes from the Fourier analysis part of the Statistics and Fourier Analysis course. For more advanced use see the Methods of Mathematical Physics course notes. For their applications to electromagentism see the Electromagentism course notes.

Let $\mathfrak{L}$ be a linear differential operator. Consider the equation $\mathfrak{L}y(x) = f(x)$. Here we want to solve for some function, y, depending on x. We call $f(x)$ the source term. The **Green's function** for this equation is the solution to the related equation

$$\mathfrak{L}G(x) = \delta(x). \tag{14.2.1}$$

That is, it gives the response of the system to a sharp impulse.

The general solution, y, is then given by integrating:

$$y(x) = \int G(x - x') f(x') \, dx'. \tag{14.2.2}$$

This works by adding up the response to impulses at all possible positions, x'. Applying $\mathfrak{L}$ shows that y when defined like this is indeed a solution, under the

assumption that everything is sufficiently smooth for the differential operator to commute with the integral:

$$\mathfrak{L}y(x) = \int \mathfrak{L}G(x - x')f(x')\,\mathrm{d}x' = \int \delta(x - x')f(x')\,\mathrm{d}x' = f(x). \qquad (14.2.3)$$

From which we see that we recover the equation $\mathfrak{L}y(x) = f(x)$.

So, once we find the Green's function, often easier than solving the equation for a general source, we can find the general solution by integrating.

Since this is physics we also have to be concerned with making our solutions physically reasonable. In the context of Green's functions this means that we look for causal Green's functions, which vanish before the impulse, so when $x < 0$ above. We're also doing relativity, so we restrict all motion to be slower than the speed of light. In particular this means that we look for **retarded Green's functions**, which have the response at a distance x from the source occur at a time x/c after the source does whatever it is that the response is in response to.

14.3   Application to Wave Equation

The wave equation with a source in three dimensional space is

$$\frac{1}{c^2}\frac{\partial^2}{\partial t^2}\varphi(t, \boldsymbol{x}) - \nabla^2\varphi(t, \boldsymbol{x}) = \rho(t, \boldsymbol{x}). \qquad (14.3.1)$$

Here φ is the wave and ρ is the source. This is for a scalar field, but the equation works the same for a vector field, or for a higher rank tensor field.

The Green's function for this system, G, is then defined as the solution to

$$\frac{1}{c^2}\frac{\partial^2}{\partial t^2}G(t, \boldsymbol{x}) - \nabla^2\varphi(t, \boldsymbol{x}) = \frac{1}{c}\delta(t)\delta^3(\boldsymbol{x}). \qquad (14.3.2)$$

Here the factor of $1/c$ on the right hand side is just convention, and is a choice of normalisation such that the area under the delta distribution is $1/c$, instead of 1. Leaving it out would just result in a factor of c appearing elsewhere. It is also common to define the right hand side to be negative instead, again, this is just matter of convention.

We want to make this equation look Lorentz covariant. To start with we combine the time and space parts into the d'Alembertian:

$$\partial^2\varphi(x) = \frac{1}{c}\delta(t)\delta^3(\boldsymbol{x}). \qquad (14.3.3)$$

To proceed we need a basic property of delta distributions:

Lemma 14.3.4 Let $a > 0$, then $\delta(ax) = \delta(x)/a$ as distributions. That is, for a smooth test function, f, we have

$$\int_{x_0}^{x_1} f(x)\delta(ax)\,\mathrm{d}x = \frac{1}{a}\int_{x_0}^{x_1} f(x)\delta(x)\,\mathrm{d}x \qquad (14.3.5)$$

when 0 is in the range of integration.

Proof. Consider the left hand side of the integral equation above:

$$I = \int_{x_0}^{x_1} f(x)\delta(ax)\,dx. \tag{14.3.6}$$

Complete a change of variables, defining $u = ax$. Then we have $a\,dx = du$, and so

$$I = \frac{1}{a}\int_{x_0}^{x_1} f(u/a)\delta(u)\,du = \frac{1}{a}f(0/a) = \frac{1}{a}f(0). \tag{14.3.7}$$

The right hand side of the integral equation is

$$\frac{1}{a}\int_{x_0}^{x_1} f(x)\delta(x)\,dx = \frac{1}{a}f(0) = I, \tag{14.3.8}$$

proving the lemma. $\square$

Using this we have $\delta(t)/c = \delta(ct) = \delta(x^0)$ as distributions. Hence, we have

$$\partial^2\varphi(x) = \frac{1}{c}\delta(t)\delta^3(\boldsymbol{x}) = \delta(x^0)\delta^3(\boldsymbol{x}) = \delta^4(x). \tag{14.3.9}$$

We will mainly be interested in the retarded Green's function, which for the wave equation is

$$G_{\text{R}}(t, \boldsymbol{x}) = \frac{1}{4\pi c|\boldsymbol{x}|}\delta(t - |\boldsymbol{x}|/c). \tag{14.3.10}$$

Note that the factor of $1/c$ out the front is due to the choice of including a factor of $1/c$ in front of the Dirac deltas in the definition of the Green's function.

The solution for φ is then

$$\varphi(x) = \int d^4x\, \frac{1}{4\pi c|\boldsymbol{x} - \boldsymbol{x}'|}\delta^4(t - t' - |\boldsymbol{x} - \boldsymbol{x}'|/c)\rho(t', \boldsymbol{x}'). \tag{14.3.11}$$

We can rewrite this in terms of t' as

$$\varphi(x) = \int d^3x'\, dt'\, \frac{1}{4\pi|\boldsymbol{x} - \boldsymbol{x}'|}\delta^4(t - t' - |\boldsymbol{x} - \boldsymbol{x}'|/c)\rho(t', \boldsymbol{x}'). \tag{14.3.12}$$

Here the factor of c from $x'^0 = ct'$ cancels with the factor of $1/c$ already present. We can then perform the integration over t using the sifting property of the delta distribution and we get

$$\varphi(x) = \int d^3x'\, \frac{1}{4\pi|\boldsymbol{x} - \boldsymbol{x}'|}\rho(t - |\boldsymbol{x} - \boldsymbol{x}'|/c, \boldsymbol{x}'). \tag{14.3.13}$$

This is as far as we can go without knowing what the source term is.

Note the one over distance dependence familiar from the static case, $\nabla^2 1/(4\pi|\boldsymbol{x}|) = -\delta^3(\boldsymbol{x})$. We can also see that at a time t and a distance $|\boldsymbol{x} - \boldsymbol{x}'|$ from the source the observer will see the effects of the source at time $t - |\boldsymbol{x} - \boldsymbol{x}'|/c$, since we choose the retarded Green's function.

14.3.1 Dirac Delta at a Function

Consider $\delta(f(x))$. It can be shown that

$$\delta(f(x)) = \sum_i \frac{\delta(x - a_i)}{|f'(a_i)|} \tag{14.3.14}$$

as distributions where a_i are the zeros of f, $f'(a_i) \neq 0$, and f is smooth. We also assume f has a finite number of zeros. This is what we shall do in this section.

To start with we suppose that f is a monotone increasing function. This means it has at most one zero. In the case with no zeros everything vanishes and its not interesting, so suppose f has one zero, call it a. Further, we suppose that a is in the region of interest, that is $a \in (x_0, x_1)$ when we are interested in

$$I = \int_{x_0}^{x_1} \delta(f(x))g(x)\,\mathrm{d}x \tag{14.3.15}$$

for some smooth test function, g. We want to show that

$$I = \frac{1}{|f'(a)|} \int_{x_0}^{x_1} \delta(x - a)g(x) = \frac{1}{|f'(a)|}g(a). \tag{14.3.16}$$

First, notice that for a monotone increasing function $f'(x) \geq 0$ for all x, so we can ignore the absolute values as we have $|f'(a)| = f'(a)$. Now, let $u = f(x)$, then we have $u = 0$ at $x = a$. We also have $u(x_0) = f(x_0)$ and $u(x_1) = f(x_1)$, and importantly $f(x_0) < f(x_1)$. Further, $\mathrm{d}u = f'(x)\,\mathrm{d}x$ and so a change of variables gives

$$I = \int_{f(x_0)}^{f(x_1)} \frac{1}{f'(x)}\delta(u)g(x(u))\,\mathrm{d}u. \tag{14.3.17}$$

Performing the integration using the sifting property we get the integrand evaluated at the point at which u vanishes, which is at $x = a$, and so $x(u) = a$, hence

$$I = \frac{1}{f'(a)}g(a). \tag{14.3.18}$$

This proves the statement for the special case where f is a monotone increasing function.

Now, suppose that f is a monotone decreasing function. Then let

$$I = \int_{x_0}^{x_1} \delta(f(x))g(x)\,\mathrm{d}x \tag{14.3.19}$$

again. The same change of variables gives

$$I = \int_{f(x_0)}^{f(x_1)} \frac{1}{f'(x)}\delta(u)g(x(u))\,\mathrm{d}u. \tag{14.3.20}$$

However, now $f(x_0) > f(x_1)$, and so we need to swap the limits picking up a negative sign:

$$I = -\int_{f(x_1)}^{f(x_0)} \frac{1}{f'(x)}\delta(u)g(x(u))\,\mathrm{d}u. \tag{14.3.21}$$

Since f is monotone decreasing we have $f'(x) \leq 0$ for all x and so $|f'(x)| = -f'(x)$ giving

$$I = \int_{f(x_1)}^{f(x_0)} \frac{1}{|f'|(x)} \delta(u) g(x(u)) \, \mathrm{d}u. \tag{14.3.22}$$

Completing the integration,

$$I = \frac{1}{|f'|(a)} g(a). \tag{14.3.23}$$

This proves the statement for the special case where f is a monotone decreasing function.

Now, let f be any smooth function. Let a_i be its zeros and by hypothesis we have $f(a_i) \neq 0$. For each zero we split our integral over (x_0, x_1) into an integral over only a small neighbourhood of a_i, say $(a_i - \varepsilon_i, a_i + \varepsilon_i)$ where we choose ε_i such that these intervals are non-intersecting. Further, by choosing sufficiently small intervals f will be monotone increasing or monotone decreasing on each interval, allowing us to apply the special cases of the theorem above. Outside of these intervals since $f(x) \neq 0$ we have $\delta(f(x)) = 0$ and so there is no contribution to the integral. We then have

$$I = \int_{x_0}^{x_1} \delta(f(x)) g(x) \, \mathrm{d}x \tag{14.3.24}$$

$$= \sum_i \int_{a_i - \varepsilon_i}^{a_i + \varepsilon_i} \delta(f(x)) g(x) \, \mathrm{d}x \tag{14.3.25}$$

$$= \sum_i \frac{1}{|f'(a_i)|} g(a_i) \tag{14.3.26}$$

$$= \sum_i \int_{a_i - \varepsilon_i}^{a_i + \varepsilon_i} \frac{\delta(x - a_i)}{|f'(a_i)|} g(x) \, \mathrm{d}x \tag{14.3.27}$$

$$= \int_{x_0}^{x_1} \sum_i \frac{\delta(x - a_i)}{|f'(a_i)|} g(x) \, \mathrm{d}x. \tag{14.3.28}$$

Therefore, as distributions, we have

$$\delta(f(x)) = \sum_i \frac{\delta(x - a_i)}{|f'(a_i)|}. \tag{14.3.29}$$

Note that if we have $f(x) = ax$ for $a \neq 0$ then the only zero is at $x = 0$ and we have $f'(x) = a$ so we have

$$\delta(ax) = \frac{1}{|a|} \delta(x) \tag{14.3.30}$$

as distributions. This further simplifies to $\delta(ax) = \delta(x)/a$ for $a > 0$, which is Lemma 14.3.4.

14.3.2 Application to Green's Function

Motivated by this property of Dirac deltas we look for some function, f, such that

$$G_{\mathrm{R}}(x) = \delta(f(t)). \tag{14.3.31}$$

We start by considering $g(t) = c^2 t^2 - |x|^2$, where we treat $\boldsymbol{x}$ as fixed:

$$\delta(g(t)) = \delta(c^2 t^2 - |\boldsymbol{x}|^2). \tag{14.3.32}$$

This function has two zeros, $t_\pm = \pm |\boldsymbol{x}|/c$. The derivative is $g'(t) = 2c^2 t$, so $|g'(t_\pm)| = 2c|\boldsymbol{x}|$. Hence,

$$\delta(c^2 t^2 - |\boldsymbol{x}|^2) = \frac{1}{2c|\boldsymbol{x}|}[\delta(t - |x|/c) + \delta(t + |x|/c)]. \tag{14.3.33}$$

Comparing this to the retarded Green's function we see that the first term matches up to a factor of 2π. We can also introduce the **advanced Green's function**, which differs by the sign in the delta:

$$G_{\mathrm{A}}(x) = \frac{1}{4\pi c|\boldsymbol{x}|}\delta(t + |\boldsymbol{x}|/c). \tag{14.3.34}$$

Then we have

$$\delta(c^2 t^2 - |\boldsymbol{x}|^2) = 2\pi(G_{\mathrm{R}}(x) + G_{\mathrm{A}}). \tag{14.3.35}$$

We don't want the advanced Green's function to appear in our solution. Notice that the support of the retarded Green's function, that is the set upon which the function doesn't vanish, is positive, whereas the support of the advanced Green's function is negative. We can then multiply by a Heaviside step function,

$$\theta(x) = \begin{cases} 1 & x > 0, \\ 0 & x < 0. \end{cases} \tag{14.3.36}$$

This gives

$$G_{\mathrm{R}}(x) = \frac{1}{2\pi}\delta(c^2 t^2 - \boldsymbol{x}^2)\theta(x^0) = \frac{1}{2\pi}\delta(x^2)\theta(x^0). \tag{14.3.37}$$

This is (sufficiently) relativisticly invariant. The delta term is clearly invariant, since it depends only on the length of a four-vector. The theta term is not quite invariant, since if the direction of time changes then $\theta(x^0)$ will change. However, it is invariant under time dilation, since $\gamma > 0$, so the sign of time doesn't change.

The Lorentz group, $\mathrm{O}(1,3)$, is the group of all Lorentz transformations. The $\theta(x^0)$ term is invariant under the subgroup $\mathrm{SO}^+(1,3)$, called the **proper ortho-chronous Lorentz group**. Here the S stands for special, and as with the usual special orthogonal group it means that the transformation has a determinant of 1. The + denotes that the orientation of time is preserved. It is fine that $\theta(x^0)$ is only invariant under $\mathrm{SO}^+(1,3)$ since this is the component of the Lorentz group which is connected to the identity and so is the group generated by the Lie algebra $\mathfrak{so}(1,3)$.[1]

Solutions to the wave equation aren't unique. If φ is a solution to $\partial^2 \varphi = \rho$ then so is $\varphi + \varphi_{\mathrm{H}}$ where φ_{H} is a solution to the homogeneous equation, $\partial^2 \varphi_{\mathrm{H}} = 0$. This corresponds to the fact that the Green's function is not unique unless we impose boundary conditions. We choose boundary conditions corresponding to outgoing radiation, such as a particle causing electromagnetic waves. This gives the retarded Green's function. We could also consider boundary conditions corresponding to incoming radiation, such as firing a laser at a particle, which corresponds to the

[1]See the notes for the symmetries of quantum mechanics or symmetries of particles and fields courses for details.

advanced Green's function, which we get with $\theta(-x^0)$. If we had two particles interacting electromagnetically then we would need a mixture of these two boundary conditions as radiation leaves one particle and arrives at the other. This leads to the Feynman Green's function, which has a $\theta(x^0)$ term and a $\theta(-x^0)$ term. This is common in QFT[2].

[2]See the quantum theory notes

Index