611: Electromagnetic Theory II

CONTENTS

- Special relativity; Lorentz covariance of Maxwell equations

- Scalar and vector potentials, and gauge invariance

- Relativistic motion of charged particles

- Action principle for electromagnetism; energy-momentum tensor

- Electromagnetic waves; waveguides

- Fields due to moving charges

- Radiation from accelerating charges

- Antennae

- Radiation reaction

- Magnetic monopoles, duality, Yang-Mills theory

Suggested textbooks:

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

- J.D. Jackson, *Classical Electrodynamics*

Contents

1 Electrodynamics and Special Relativity

1.1 Introduction

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

James Clerk Maxwell

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

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

$$\vec{r}\,' = \vec{r} - \vec{v}\,t\,. \tag{1.1}$$

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

$$t' = t \,. \tag{1.2}$$

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

$$\vec{r}\,' = \mathbf{M} \cdot \vec{r} - \vec{v}\,t \,, \qquad t' = t \,, \tag{1.3}$$

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

$$\vec{r} \leftrightarrow \begin{pmatrix} x \\ y \\ z \end{pmatrix} , \qquad \mathbf{M} \cdot \vec{r} \longleftrightarrow \mathbf{M} \begin{pmatrix} x \\ y \\ z \end{pmatrix} , \tag{1.4}$$

where $\mathbf{M}^T \mathbf{M} = 1$, and $\mathbf{M}^T$ denotes the transpose of the matrix $\mathbf{M}$.

Returning to our simplifying assumption that the two frames are parallel, i.e. that $M = \mathbb{1}$, it follows that if a particle having position vector $\vec{r}$ in S moves with velocity $\vec{u} = d\vec{r}/dt$, then its velocity $\vec{u}\,' = d\vec{r}\,'/dt$ as measured with respect to the frame S' is given by

$$\vec{u}\,' = \vec{u} - \vec{v} \,. \tag{1.5}$$

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

$$c' = c - v \,. \tag{1.6}$$

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

Of course, it should be emphasised that the discrepancies between experiment and the Galilean transformations are rather negligible if the relative speed v between the two inertial frames is of a typical "everyday" magnitude, such as the speed of a car or a plane. But if

v begins to become appreciable in comparison to the speed of light, then the discrepancy becomes appreciable too.

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

$$\vec{\nabla} \cdot \vec{E} \;\; = \;\; \frac{1}{\epsilon_0} \, \rho \,, \qquad \vec{\nabla} \times \vec{B} - \mu_0 \epsilon_0 \, \frac{\partial \vec{E}}{\partial t} = \mu_0 \, \vec{J} \,,$$
$$\vec{\nabla} \cdot \vec{B} \;\; = \;\; 0 \,, \qquad \vec{\nabla} \times \vec{E} + \frac{\partial \vec{B}}{\partial t} = 0 \,, \tag{1.7}$$

where $\vec{E}$ and $\vec{B}$ are the electric and magnetic fields, ρ and $\vec{J}$ are the charge density and current density, and ϵ_0 and μ_0 are the permittivity and permeability of free space.[1]

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

$$\vec{\nabla} \times (\vec{\nabla} \times \vec{E}) = -\frac{\partial}{\partial t} \vec{\nabla} \times \vec{B} = -\mu_0 \epsilon_0 \, \frac{\partial^2 \vec{E}}{\partial t^2} \,. \tag{1.8}$$

But using the vector identity $\vec{\nabla} \times (\vec{\nabla} \times \vec{E}) = \vec{\nabla}(\vec{\nabla} \cdot \vec{E}) - \nabla^2 \vec{E}$, it follows from $\vec{\nabla} \cdot \vec{E} = 0$ that the electric field satisfies the wave equation

$$\nabla^2 \vec{E} - \mu_0 \epsilon_0 \, \frac{\partial^2 \vec{E}}{\partial t^2} = 0 \,. \tag{1.9}$$

This admits plane-wave solutions of the form

$$\vec{E} = \vec{E}_0 \, e^{\mathrm{i}(\vec{k} \cdot \vec{r} - \omega t)} \,, \tag{1.10}$$

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

$$k^2 = \mu_0 \epsilon_0 \, \omega^2 \,. \tag{1.11}$$

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

$$c = \frac{\omega}{k} = \frac{1}{\sqrt{\mu_0 \epsilon_0}} \,. \tag{1.12}$$

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

---

[1]The equations here are written using the system of units known as SI, which could be said to stand for "Super Inconvenient." In these units, the number of unnecessary dimensionful "fundamental constants" is maximised. We shall pass speedily to more convenient units a little bit later.

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

$$\vec{B} = \frac{1}{\omega}\,\vec{k} \times \vec{E}\,. \tag{1.13}$$

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

In fact, as we shall see, the transformations that correctly describe the relation between observations in different inertial frames in uniform motion are the *Lorentz Transformations* of Special Relativity. Furthermore, even though the Maxwell equations were written down in the pre-relativity days of the nineteenth century, they are in fact perfectly invariant[2] under the Lorentz transformations. No further modification is required in order to incorporate Maxwell's theory of electromagnetism into special relativity.

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

Actually, before proceeding it is instructive to take a step back and look at what the Maxwell equations actually looked like in Maxwell's 1865 paper *A Dynamical Theory of the Electromagnetic Field*, published in the Philosophical Transactions of the Royal Society of London. It must be recalled that in 1865 three-dimensional vectors had not yet been invented, and so everything was written out explicitly in terms of the x, y and z components.[3] To make matters worse, Maxwell used a different letter of the alphabet for each component of each field. In terms of the now-familiar electric vector fields $\vec{E}$, $\vec{D}$, the magnetic fields $\vec{B}$, $\vec{H}$, the current density $\vec{J}$ and the charge density ρ, Maxwell's chosen names for the

---

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

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

components were

$$
\begin{aligned}
\vec{E} &= (P, Q, R), & \vec{D} &= (f, g, h), \\
\vec{B} &= (F, G, H), & \vec{H} &= (\alpha, \beta, \gamma), \\
\vec{J} &= (p, q, r), & \rho &= e.
\end{aligned}
\tag{1.14}
$$

Thus the Maxwell equations that we now write rather compactly as[4]

$$
\vec{\nabla} \cdot \vec{D} = 4\pi\rho, \qquad \vec{\nabla} \cdot \vec{B} = 0,
\tag{1.15}
$$

$$
\vec{\nabla} \times \vec{H} - \frac{\partial \vec{D}}{\partial t} = 4\pi\vec{J}, \qquad \vec{\nabla} \times \vec{E} + \frac{\partial \vec{B}}{\partial t} = 0,
\tag{1.16}
$$

took, in 1865, the highly inelegant forms

$$
\begin{aligned}
\frac{\partial f}{\partial x} + \frac{\partial g}{\partial y} + \frac{\partial h}{\partial z} &= 4\pi e, \\
\frac{\partial F}{\partial x} + \frac{\partial G}{\partial y} + \frac{\partial H}{\partial z} &= 0,
\end{aligned}
\tag{1.17}
$$

for the two equations in (1.15), and

$$
\begin{aligned}
\frac{\partial \gamma}{\partial y} - \frac{\partial \beta}{\partial z} - \frac{\partial f}{\partial t} &= 4\pi p, \\
\frac{\partial \alpha}{\partial z} - \frac{\partial \gamma}{\partial x} - \frac{\partial g}{\partial t} &= 4\pi q, \\
\frac{\partial \beta}{\partial x} - \frac{\partial \alpha}{\partial y} - \frac{\partial h}{\partial t} &= 4\pi r, \\
\frac{\partial R}{\partial y} - \frac{\partial Q}{\partial z} + \frac{\partial F}{\partial t} &= 0, \\
\frac{\partial P}{\partial z} - \frac{\partial R}{\partial x} + \frac{\partial G}{\partial t} &= 0, \\
\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} + \frac{\partial H}{\partial t} &= 0,
\end{aligned}
\tag{1.18}
$$

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

---

[4]Here, we are writing the equations in the so-called "Natural Units," which we shall be using throughout this course.

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

1.2 The Lorentz Transformation

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

- The laws of physics are the same for all inertial observers.

- The speed of light is the same for all inertial observers.

To derive the Lorentz transformation, let us suppose that we have two inertial frames S and S', whose origins coincide at time zero, that is to say, at $t = 0$ in the frame S, and at $t' = 0$ in the frame S'. If a flash of light is emitted at the origin at time zero, then it will spread out over a spherical wavefront given by

$$x^2 + y^2 + z^2 - c^2 t^2 = 0 \tag{1.19}$$

in the frame S, and by

$$x'^2 + y'^2 + z'^2 - c^2 t'^2 = 0 \tag{1.20}$$

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

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

$$y' = y, \qquad z' = z. \tag{1.21}$$

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

$$x' = Ax + Bt, \qquad t' = Cx + Dt, \tag{1.22}$$

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

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

$$x' = A(x - vt). \tag{1.23}$$

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

$$x = A(x' + vt'). \tag{1.24}$$

Note that it must be the *same* constant A in both these equations, since the two really just correspond to reversing the direction of the x axis, and the physics must be the same for the two cases.

Now we bring in the postulate that the speed of light is the same in the two frames, so if we have $x = ct$ then this must imply $x' = ct'$. Solving the resulting two equations

$$ct' = A(c - v)t, \qquad ct = A(c + v)t' \tag{1.25}$$

for A, we obtain

$$A = \frac{1}{\sqrt{1 - v^2/c^2}}. \tag{1.26}$$

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

$$t' = A\left(t - \frac{v}{c^2}x\right). \tag{1.27}$$

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

$$x' = \gamma(x - vt), \qquad y' = y, \qquad z' = z, \qquad t' = \gamma(t - \frac{v}{c^2}x), \tag{1.28}$$

where

$$\gamma = \frac{1}{\sqrt{1 - v^2/c^2}}, \tag{1.29}$$

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

At this point, for notational convenience, we shall introduce the simplification of working in a system of units in which the speed of light is set equal to 1. We can do this because the

speed of light is the same for all inertial observers, and so we may as well choose to measure length in terms of the time it takes for light *in vacuo* to traverse the distance. In fact, the metre is nowadays *defined* to be the distance travelled by light *in vacuo* in 1/299,792,458 of a second. By making the small change of taking the *light-second* as the basic unit of length, rather than the $1/299,792,458'th$ of a light-second, we end up with a system of units in which $c = 1$. Alternatively, we could measure time in "light metres," where the unit is the time taken for light to travel 1 metre. In these units, the Lorentz transformation (1.28) becomes

$$x' = \gamma(x - vt), \qquad y' = y, \qquad z' = z, \qquad t' = \gamma(t - vx), \tag{1.30}$$

where

$$\gamma = \frac{1}{\sqrt{1 - v^2}}. \tag{1.31}$$

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

$$\vec{r}\,' = \vec{r} + \frac{\gamma - 1}{v^2}\,(\vec{v} \cdot \vec{r})\,\vec{v} - \gamma \vec{v}\, t, \qquad t' = \gamma(t - \vec{v} \cdot \vec{r}), \tag{1.32}$$

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

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

$$x^2 + y^2 + z^2 - t^2 = x'^2 + y'^2 + z'^2 - t'^2. \tag{1.33}$$

11

1.3 An interlude on 3-vectors and suffix notation

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

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

$$\vec{A} = (A_1, A_2, A_3)\,. \tag{1.34}$$

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

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

$$\vec{A} \cdot \vec{B} = A_1 B_1 + A_2 B_2 + A_3 B_3 = \sum_{i=1}^{3} A_i B_i\,. \tag{1.35}$$

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

$$\vec{A} \cdot \vec{B} = A_i B_i\,. \tag{1.36}$$

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

The next concept to introduce is the *Kronecker delta* tensor δ_{ij}. This is defined by

$$\delta_{ij} = 1 \quad \text{if} \quad i = j\,, \qquad \delta_{ij} = 0 \quad \text{if} \quad i \neq j\,, \tag{1.37}$$

Thus

$$\delta_{11} = \delta_{22} = \delta_{33} = 1\,, \qquad \delta_{12} = \delta_{13} = \cdots = 0\,. \tag{1.38}$$

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

$$A_i = \delta_{ij} A_j\,, \tag{1.39}$$

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

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

$$\vec{A} \times \vec{B} = (A_2 B_3 - A_3 B_2, A_3 B_1 - A_1 B_3, A_1 B_2 - A_2 B_1) \,. \tag{1.40}$$

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

$$\epsilon_{ijk} = -\epsilon_{ikj} = -\epsilon_{jik} = -\epsilon_{kji} \,. \tag{1.41}$$

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

$$\epsilon_{123} = \epsilon_{231} = \epsilon_{312} = +1 \,, \qquad \epsilon_{132} = \epsilon_{321} = \epsilon_{213} = -1 \,, \tag{1.42}$$

with all other components vanishing.

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

$$(\vec{A} \times \vec{B})_i = \epsilon_{ijk} A_j B_k \,. \tag{1.43}$$

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

$$(\vec{A} \times \vec{B})_1 = \epsilon_{1jk} A_j B_k = \epsilon_{123} A_2 B_3 + \epsilon_{132} A_3 B_2 = A_2 B_3 - A_3 B_2 \,, \tag{1.44}$$

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

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

$$[\vec{A} \times (\vec{B} \times \vec{C})]_i = \epsilon_{ijk} A_j (\vec{B} \times \vec{C})_k = \epsilon_{ijk} \epsilon_{k\ell m} A_j B_\ell C_m \,. \tag{1.45}$$

For convenience, we may cycle the indices on the second ϵ tensor around and write this as

$$[\vec{A} \times (\vec{B} \times \vec{C})]_i = \epsilon_{ijk} \epsilon_{\ell m k} A_j B_\ell C_m \,. \tag{1.46}$$

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

$$\epsilon_{ijk} \epsilon_{\ell m k} = \delta_{i\ell} \delta_{jm} - \delta_{im} \delta_{j\ell} \,. \tag{1.47}$$

Using this in (1.46), we have

$$
\begin{aligned}
[\vec{A} \times (\vec{B} \times \vec{C})]_i &= (\delta_{i\ell}\delta_{jm} - \delta_{im}\delta_{j\ell})A_j B_\ell C_m \,, \\
&= \delta_{i\ell}\delta_{jm}A_j B_\ell C_m - \delta_{im}\delta_{j\ell}A_j B_\ell C_m \,, \\
&= B_i A_j C_j - C_i A_j B_j \,, \\
&= (\vec{A} \cdot \vec{C})\, B_i - (\vec{A} \cdot \vec{B})\, C_i \,.
\end{aligned}
\tag{1.48}
$$

In other words, we have proven that

$$
\vec{A} \times (\vec{B} \times \vec{C}) = (\vec{A} \cdot \vec{C})\, \vec{B} - (\vec{A} \cdot \vec{B})\, \vec{C} \,.
\tag{1.49}
$$

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

$$
\vec{\nabla} = \Big(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z} \Big) \,.
\tag{1.50}
$$

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

$$
\partial_i = \frac{\partial}{\partial x_i} \,.
\tag{1.51}
$$

Thus, the i'th component of $\vec{\nabla}$ is ∂_i.

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

$$
\mathrm{div}\,\vec{A} = \vec{\nabla} \cdot \vec{A} = \partial_i A_i \,, \qquad (\mathrm{curl}\,\vec{A})_i = (\vec{\nabla} \times \vec{A})_i = \epsilon_{ijk}\partial_j A_k \,.
\tag{1.52}
$$

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

$$
\nabla^2 = \partial_i \partial_i \,.
\tag{1.53}
$$

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

$$
r^2 = x_j x_j \,.
\tag{1.54}
$$

If we now act with ∂_i on both sides, we get

$$
2r\, \partial_i r = 2x_j\, \partial_i x_j = 2x_j\, \delta_{ij} = 2x_i \,.
\tag{1.55}
$$

Thus we have the very useful result that

$$
\partial_i r = \frac{x_i}{r} \,.
\tag{1.56}
$$

14

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

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \mathbf{M} \begin{pmatrix} x \\ y \\ z \end{pmatrix}, \tag{1.57}$$

where

$$\mathbf{M}^T \, \mathbf{M} = 1. \tag{1.58}$$

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

$$\mathbf{M} = \begin{pmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{1.59}$$

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

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

$$\mathbf{M} = \begin{pmatrix} M_{11} & M_{12} & M_{13} \\ M_{21} & M_{22} & M_{23} \\ M_{31} & M_{32} & M_{33} \end{pmatrix}. \tag{1.60}$$

The rotation (1.57) can then be expressed as

$$x_i' = M_{ij} \, x_j, \tag{1.61}$$

---

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

and the orthogonality condition (1.58) is

$$M_{ki}\, M_{kj} = \delta_{ij}\,. \tag{1.62}$$

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

$$x'_i\, x'_i = M_{ij}\, M_{ik}\, x_j\, x_k = \delta_{jk}\, x_j\, x_k = x_j\, x_j = x_i\, x_i\,. \tag{1.63}$$

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

$$A'_i = M_{ij}\, A_j\,, \qquad B'_i = M_{ij}\, B_j\,. \tag{1.64}$$

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

$$A'_i B'_i = M_{ij} A_j\, M_{ik} B_k\,. \tag{1.65}$$

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

$$A'_i B'_i = A_j B_k \delta_{jk} = A_j B_j\,. \tag{1.66}$$

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

It is useful to count up how many independent parameters are needed to specify the most general possible rotation matrix $\mathbf{M}$. Looking at (1.60), we can see that a general 3×3 matrix has 9 components. But our matrix $\mathbf{M}$ is required to be orthogonal, i.e. it must satisfy $\mathbf{M}^T \mathbf{M} - 1 = 0$. How many equations does this amount to? Naively, it is a 3×3 matrix equation, and so implies 9 conditions. But this is not correct, since the left-hand side of $\mathbf{M}^T \mathbf{M} - 1 = 0$ is in fact a *symmetric* matrix. (Take the transpose, and verify this.) A 3×3 symmetric matrix has $(3 \times 4)/2 = 6$ independent components, and so setting a symmetric 3×3 matrix to zero implies only 6 independent equations rather than 9. Thus the orthogonality condition imposes 6 constraints on the 9 components of a general 3×3 matrix, and so that leaves

$$9 - 6 = 3 \tag{1.67}$$

16

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

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

1.4 4-vectors and 4-tensors

The Lorentz transformations given in (1.32) are linear in the space and time coordinates. They can be written more succinctly if we first define the set of four spacetime coordinates denoted by x^μ, where μ is an index, or label, that ranges over the values 0, 1, 2 and 3. The case $\mu = 0$ corresponds to the time coordinate t, while $\mu = 1$, 2 and 3 corresponds to the space coordinates x, y and z respectively. Thus we have[6]

$$(x^0, x^1, x^2, x^3) = (t, x, y, z) \,. \tag{1.68}$$

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

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

$$\eta_{00} = -1 \,, \qquad \eta_{11} = \eta_{22} = \eta_{33} = 1 \,, \tag{1.69}$$

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

$$\eta_{\mu\nu} = \eta_{\nu\mu} \,. \tag{1.70}$$

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

$$x^2 + y^2 + z^2 - t^2 = \sum_{\mu=0}^{3} \sum_{\nu=0}^{3} \eta_{\mu\nu} \, x^\mu x^\nu \,. \tag{1.71}$$

---

[6]The choice to put the index label μ as a superscript, rather than a subscript, is purely conventional. But, unlike the situation with many arbitrary conventions, in this case the coordinate index is placed upstairs in *all* modern literature.

In the same way as we previously associated 2-index objects in 3-dimensional Euclidean space with 3×3 matrices, so here too we can associate $\eta_{\mu\nu}$ with a 4×4 matrix $\boldsymbol{\eta}$:

$$\boldsymbol{\eta} = \begin{pmatrix} \eta_{00} & \eta_{01} & \eta_{02} & \eta_{03} \\ \eta_{10} & \eta_{11} & \eta_{12} & \eta_{13} \\ \eta_{20} & \eta_{21} & \eta_{22} & \eta_{23} \\ \eta_{30} & \eta_{31} & \eta_{32} & \eta_{33} \end{pmatrix} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} . \tag{1.72}$$

Thus one can think of the rows of the matrix on the right as being labelled by the index μ and the columns being labelled by the index ν.

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

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

Since in (1.71) both μ and ν occur exactly twice, we can rewrite the expression, using the Einstein summation convention, as simply

$$x^2 + y^2 + z^2 - t^2 = \eta_{\mu\nu} \, x^\mu x^\nu \,. \tag{1.73}$$

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

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

$$x'^\mu = \Lambda^\mu{}_\nu \, x^\nu \,, \tag{1.74}$$

---

[7]As a side remark, it should be noted that in a valid vectorial or tensorial expression, a specific index can **NEVER** appear more than twice in a given term. If you have written down a term where a given index occurs 3, 4 or more times then there is no need to look further at it; it is **WRONG**. Thus, for example, it is totally meaningless to write $\eta_{\mu\mu} \, x^\mu x^\mu$. If you ever find such an expression in a calculation then you must stop, and go back to find the place where an error was made.

where $\Lambda^\mu{}_\nu$ are constants, and the Einstein summation convention is operative for the dummy index ν. By comparing (1.74) carefully with (1.32), we can see that the components $\Lambda^\mu{}_\nu$ are given by

$$
\begin{aligned}
\Lambda^0{}_0 &= \gamma\,, & \Lambda^0{}_i &= -\gamma v_i\,, \\
\Lambda^i{}_0 &= -\gamma\,v_i\,, & \Lambda^i{}_j &= \delta_{ij} + \frac{\gamma-1}{v^2} v_i v_j\,,
\end{aligned}
\tag{1.75}
$$

where δ_{ij} is the Kronecker delta symbol,

$$
\delta_{ij} = 1 \quad \text{if} \quad i = j\,, \qquad \delta_{ij} = 0 \quad \text{if} \quad i \neq j\,.
\tag{1.76}
$$

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

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

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

$$
\eta_{\mu\nu}\, x^\mu x^\nu = \eta_{\mu\nu}\, \Lambda^\mu{}_\rho\, \Lambda^\nu{}_\sigma\, x^\rho x^\sigma\,.
\tag{1.77}
$$

(Note that we have been careful to choose two different dummy indices for the two implicit summations over ρ and σ!) On the left-hand side, we can replace the dummy indices μ and ν by ρ and σ, and thus write

$$
\eta_{\rho\sigma}\, x^\rho x^\sigma = \eta_{\mu\nu}\, \Lambda^\mu{}_\rho\, \Lambda^\nu{}_\sigma\, x^\rho x^\sigma\,.
\tag{1.78}
$$

This can be grouped together as

$$
(\eta_{\rho\sigma} - \eta_{\mu\nu}\, \Lambda^\mu{}_\rho\, \Lambda^\nu{}_\sigma) x^\rho x^\sigma = 0\,,
\tag{1.79}
$$

and, since it is true for any x^μ, we must have that

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

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

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

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

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

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

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

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

independent parameters.

Notice that if η had been simply the 4×4 unit matrix, then (1.82) would have been a direct 4-dimensional analogue of the 3-dimensional orthogonality condition (1.58). In other words, were it not for the minus sign in the 00 component of η, the Lorentz transformations would just be spatial rotations in 4 dimensions, and they would be elements of the group $O(4)$. The counting of the number of independent such transformations would be identical to the one given above, and so the group $O(4)$ of orthogonal 4×4 matrices is characterised by 6 independent parameters.

Because of the minus sign in η, the group of 4×4 matrices satisfying (1.82) is called $O(1,3)$, with the numbers 1 and 3 indicating the number of *time* and *space* dimensions respectively. Thus the four-dimensional *Lorentz Group* is $O(1,3)$.

Obviously, the subset of Λ matrices of the form

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

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

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

$$\Lambda^\mu{}_\nu = \Lambda^\mu{}_\rho(R) \, \Lambda^\rho{}_\nu(B) \,, \qquad (1.85)$$

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

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

$$\Lambda^\mu{}_\nu = \widetilde{\Lambda}^\mu{}_\rho(B) \, \widetilde{\Lambda}^\rho{}_\nu(R) \,, \qquad (1.86)$$

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

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

The coordinates $x^\mu = (x^0, x^i)$ live in a four-dimensional *spacetime*, known as *Minkowski Spacetime*. This is the four-dimensional analogue of the three-dimensional *Euclidean Space*

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

$$ds^2 = \eta_{\mu\nu} \, dx^\mu dx^\nu \,. \tag{1.87}$$

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

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

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

$$
\begin{aligned}
ds'^2 &= \eta_{\mu\nu} \, dx'^\mu dx'^\nu = \eta_{\mu\nu} \, \Lambda^\mu{}_\rho \, \Lambda^\nu{}_\sigma \, dx^\rho dx^\sigma \,, \\
&= \eta_{\rho\sigma} \, dx^\rho dx^\sigma = ds^2 \,.
\end{aligned}
\tag{1.89}
$$

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

The Lorentz transformation rule of the coordinate differential dx^μ, i.e.

$$dx'^\mu = \Lambda^\mu{}_\nu \, dx^\nu \,, \tag{1.90}$$

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

$$U'^\mu = \Lambda^\mu{}_\nu \, U^\nu \,. \tag{1.91}$$

The Minkowski metric $\eta_{\mu\nu}$ may be thought of as a 4×4 matrix, whose rows are labelled by μ and columns labelled by ν, as in (1.72). Clearly, the inverse of this matrix takes the same form as the matrix itself. We denote the components of the inverse matrix by $\eta^{\mu\nu}$. This is called, not surprisingly, the *inverse Minkowksi metric*. Clearly it satisfies the relation

$$\eta_{\mu\nu}\,\eta^{\nu\rho} = \delta_\mu^\rho\,, \tag{1.92}$$

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

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

$$U_\mu = \eta_{\mu\nu}\,U^\nu\,. \tag{1.93}$$

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

$$\eta^{\mu\nu}\,U_\nu = \eta^{\mu\nu}\,\eta_{\nu\rho}\,U^\rho = \delta_\rho^\mu\,U^\rho = U^\mu\,. \tag{1.94}$$

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

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

$$\Lambda_\mu{}^\nu = \eta_{\mu\rho}\,\eta^{\nu\sigma}\,\Lambda^\rho{}_\sigma\,. \tag{1.95}$$

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

$$\Lambda^\mu{}_\nu\,\Lambda_\mu{}^\rho = \delta_\nu^\rho\,. \tag{1.96}$$

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

$$S_{\mu\nu} = T_{\mu\nu} \quad\Longleftrightarrow\quad S_\mu{}^\nu = T_\mu{}^\nu \quad\Longleftrightarrow\quad S^\mu{}_\nu = T^\mu{}_\nu \quad\Longleftrightarrow\quad S^{\mu\nu} = T^{\mu\nu}\,. \tag{1.97}$$

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

$$A^\mu\,B_\mu = A_\mu\,B^\mu\,, \tag{1.98}$$

23

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

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

$$x^\mu = \Lambda_\nu{}^\mu\, x'^\nu\,. \tag{1.99}$$

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

$$U'_\mu = \Lambda_\mu{}^\nu\, U_\nu\,. \tag{1.100}$$

Any set of 4 quantities U_μ which transform in this way under Lorentz transformations will be called a covariant 4-vector.

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

$$\frac{\partial}{\partial x'^\mu} = \frac{\partial x^\nu}{\partial x'^\mu}\, \frac{\partial}{\partial x^\nu}\,. \tag{1.101}$$

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

$$\frac{\partial x^\nu}{\partial x'^\mu} = \Lambda_\mu{}^\nu\,, \tag{1.102}$$

and hence (1.101) gives

$$\frac{\partial}{\partial x'^\mu} = \Lambda_\mu{}^\nu\, \frac{\partial}{\partial x^\nu}\,. \tag{1.103}$$

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

---

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

symbol to denote it. We therefore define

$$\partial_\mu \equiv \frac{\partial}{\partial x^\mu} \,. \tag{1.104}$$

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

$$\partial'_\mu = \Lambda_\mu{}^\nu \, \partial_\nu \,. \tag{1.105}$$

1.5   Lorentz tensors

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

$$T'_{\mu\nu} = \Lambda_\mu{}^\rho \, \Lambda_\nu{}^\sigma \, T_{\rho\sigma} \,. \tag{1.106}$$

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

$$T'^{\mu_1\cdots\mu_m}{}_{\nu_1\cdots\nu_n} = \Lambda^{\mu_1}{}_{\rho_1} \cdots \Lambda^{\mu_m}{}_{\rho_m} \, \Lambda_{\nu_1}{}^{\sigma_1} \cdots \Lambda_{\nu_n}{}^{\sigma_n} \, T^{\rho_1\cdots\rho_m}{}_{\sigma_1\cdots\sigma_n} \,. \tag{1.107}$$

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

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

$$
\begin{aligned}
T'^{\mu\nu} &= U'^\mu V'^\nu = \Lambda^\mu{}_\rho \, U^\rho \, \Lambda^\nu{}_\sigma \, V^\sigma \,, \\
&= \Lambda^\mu{}_\rho \, \Lambda^\nu{}_\sigma \, T^{\rho\sigma} \,.
\end{aligned} \tag{1.108}
$$

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

We make also define the operation of *Contraction*, which reduces a tensor to one with a smaller number of indices. A contraction is performed by setting an upstairs index on a

25

tensor equal to a downstairs index. The Einstein summation convention then automatically comes into play, and the result is that one has an object with one fewer upstairs indices and one fewer downstairs indices. Furthermore, a simple calculation shows that the new object is itself a tensor. Consider, for example, a tensor $T^\mu{}_\nu$. This, of course, transforms as

$$T'^\mu{}_\nu = \Lambda^\mu{}_\rho \, \Lambda_\nu{}^\sigma \, T^\rho{}_\sigma \tag{1.109}$$

under Lorentz transformations. If we form the contraction and define $\phi \equiv T^\mu{}_\mu$, then we see that under Lorentz transformations we shall have

$$\begin{aligned} \phi' &\equiv T'^\mu{}_\mu = \Lambda^\mu{}_\rho \, \Lambda_\mu{}^\sigma \, T^\rho{}_\sigma \,, \\ &= \delta^\sigma_\rho \, T^\rho{}_\sigma = \phi \,. \end{aligned} \tag{1.110}$$

Since $\phi' = \phi$, it follows, by definition, that ϕ is a scalar.

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

The Minkowski metric $\eta_{\mu\nu}$ is itself a tensor, but of a rather special type, known as an *invariant* tensor. This is because, unlike a generic 2-index tensor, the Minkowski metric is identical in all Lorentz frames. To see this, let us first write out how it would transform under Lorentz transformations, using the usual transformation rules in (1.107):

$$\eta'_{\mu\nu} = \Lambda_\mu{}^\rho \, \Lambda_\nu{}^\sigma \, \eta_{\rho\sigma} \,. \tag{1.111}$$

Our goal is to show that in fact $\eta'_{\mu\nu} = \eta_{\mu\nu}$, i.e. that it is actualy *invariant* under Lorentz transformations. Now although the right-hand side of (1.111) looks reminiscent of what one has in (1.80), which is the defining property of the $\Lambda^\mu{}_\nu$ Lorentz transformations, it is not the same. Specifically, in (1.111) the indices of the two Λ transformations are contracted with η on their second indices, rather than on the first indices as in (1.80). We can easily work out what the right-hand side of (1.111) is by going through the following steps. First, we rewrite (1.80) in matrix language as $\Lambda^T \eta \Lambda = \eta$. Then right-multiply by Λ^{-1} and left-multiply by η^{-1}; this gives $\eta^{-1} \Lambda^T \eta = \Lambda^{-1}$. Next left-multiply by Λ and right-multiply by η^{-1}, which gives $\Lambda \eta^{-1} \Lambda^T = \eta^{-1}$. (This is the analogue for the Lorentz transformations of the proof, for ordinary orthogonal matrices, that $M^T M = 1$ implies $M M^T = 1$.) Converting back to index notation gives $\Lambda^\mu{}_\rho \, \Lambda^\nu{}_\sigma \, \eta^{\rho\sigma} = \eta^{\mu\nu}$. After some index raising and lowering, this gives

$$\Lambda_\mu{}^\rho \, \Lambda_\nu{}^\sigma \, \eta_{\rho\sigma} = \eta_{\mu\nu} \,. \tag{1.112}$$

26

Applying this result to the right-hand side of (1.111) therefore gives the desired result,

$$\eta'_{\mu\nu} = \eta_{\mu\nu} \,. \tag{1.113}$$

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

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

$$\Box \equiv \partial^\mu \partial_\mu = \eta^{\mu\nu}\,\partial_\mu\partial_\nu \tag{1.114}$$

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

$$\Box = -\partial_0\partial_0 + \partial_i\partial_i = -\frac{\partial^2}{\partial t^2} + \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}\,. \tag{1.115}$$

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

$$\eta_{00} = -1\,, \qquad \eta_{ij} = \delta_{ij}\,, \qquad \eta_{0i} = \eta_{i0} = 0\,, \tag{1.116}$$

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

1.6   Proper time and 4-velocity

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

$$ds^2 = \eta_{\mu\nu}\,dx^\mu dx^\nu = -dt^2 + dx^2 + dy^2 + dz^2\,. \tag{1.117}$$

This is the Minkowskian generalisation of the spatial interval in Euclidean space. Note that ds^2 can be positive, negative or zero. These cases correspond to what are called spacelike, timelike or null separations, respectively.

On occasion, it is useful to define the negative of ds^2, and write

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

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

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

$$U^\mu \equiv \frac{dx^\mu}{d\tau} \tag{1.119}$$

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

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

$$u^i = \frac{dx^i}{dt}\,. \tag{1.120}$$

From (1.118), it follows that

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

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

$$\gamma \equiv \frac{1}{\sqrt{1 - u^2}}\,. \tag{1.122}$$

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

$$U^\mu = \frac{dt}{d\tau}\frac{dx^\mu}{dt} = \gamma\,\frac{dx^\mu}{dt}\,. \tag{1.123}$$

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

$$U^0 = \gamma\,, \qquad U^i = \gamma\,u^i\,. \tag{1.124}$$

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

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

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

$$U^\mu = (\gamma, \gamma\,u^i) \qquad \text{or} \qquad U^\mu = (\gamma, \gamma\,\vec{u})\,. \tag{1.126}$$

28

Having set up the 4-vector formalism, it is now completely straightforward write down how velocities transform under Lorentz transformations. We know that the 4-velocity U^μ will transform according to (1.91), and this is identical to the way that the coordinates x^μ transform:

$$U'^\mu = \Lambda^\mu{}_\nu U^\nu, \qquad x'^\mu = \Lambda^\mu{}_\nu x^\nu. \qquad (1.127)$$

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

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

$$
\begin{aligned}
U'^0 &= \gamma_v \left(U^0 - v U^1 \right), \\
U'^1 &= \gamma_v \left(U^1 - v U^0 \right), \\
U'^2 &= U^2, \\
U'^3 &= U^3,
\end{aligned}
\qquad (1.128)
$$

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

$$
\begin{aligned}
\gamma' &= \gamma \gamma_v \left(1 - v u_x \right), \\
\gamma' u'_x &= \gamma \gamma_v \left(u_x - v \right), \\
\gamma' u'_y &= \gamma u_y, \\
\gamma' u'_z &= \gamma u_z,
\end{aligned}
\qquad (1.129)
$$

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

$$u'_x = \frac{u_x - v}{1 - v u_x}, \qquad u'_y = \frac{u_y}{\gamma_v \left(1 - v u_x \right)}, \qquad u'_z = \frac{u_z}{\gamma_v \left(1 - v u_x \right)}. \qquad (1.130)$$

2 Electrodynamics and Maxwell's Equations

2.1 Natural units

We saw earlier that the supposition of the universal validity of Maxwell's equations in all inertial frames, which in particular would imply that the speed of light should be the same in all frames, is consistent with experiment. It is therefore reasonable to expect that Maxwell's

equations should be compatible with special relativity. However, written in their standard form (1.7), this compatibility is by no means apparent. Our next task will be to re-express the Maxwell equations, in terms of 4-tensors, in a way that makes their Lorentz covariance manifest.

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

$$\vec{E} \longrightarrow \frac{1}{\sqrt{4\pi\epsilon_0}}\,\vec{E}\,, \qquad \vec{B} \longrightarrow \sqrt{\frac{\mu_0}{4\pi}}\,\vec{B}\,,$$
$$\rho \longrightarrow \sqrt{4\pi\epsilon_0}\,\rho\,, \qquad \vec{J} \longrightarrow \sqrt{4\pi\epsilon_0}\,\vec{J}\,. \tag{2.1}$$

Bearing in mind that the speed of light is given by $c = 1/\sqrt{\mu_0\epsilon_0}$, we see that the Maxwell equations (1.7) become

$$\vec{\nabla} \cdot \vec{E} \;=\; 4\pi\,\rho\,, \qquad \vec{\nabla} \times \vec{B} - \frac{1}{c}\frac{\partial\vec{E}}{\partial t} = \frac{4\pi}{c}\,\vec{J}\,,$$
$$\vec{\nabla} \cdot \vec{B} \;=\; 0\,, \qquad \vec{\nabla} \times \vec{E} + \frac{1}{c}\frac{\partial\vec{B}}{\partial t} = 0\,, \tag{2.2}$$

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

$$\vec{\nabla} \cdot \vec{E} \;=\; 4\pi\,\rho\,, \qquad \vec{\nabla} \times \vec{B} - \frac{\partial\vec{E}}{\partial t} = 4\pi\,\vec{J}\,, \tag{2.3}$$

$$\vec{\nabla} \cdot \vec{B} \;=\; 0\,, \qquad \vec{\nabla} \times \vec{E} + \frac{\partial\vec{B}}{\partial t} = 0\,, \tag{2.4}$$

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

2.2   Gauge potentials and gauge invariance

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

$$\vec{B} = \vec{\nabla} \times \vec{A}\,, \tag{2.5}$$

where $\vec{A}$ is the magnetic 3-vector potential. Note that (2.5) *identically* solves $\vec{\nabla} \cdot \vec{B} = 0$, because of the vector identity that **div curl** $\equiv 0$. Substituting (2.5) into the second equation

in (2.4), we obtain

$$\vec{\nabla} \times \left(\vec{E} + \frac{\partial \vec{A}}{\partial t} \right) = 0 \,. \tag{2.6}$$

This can be solved, again *identically*, by writing

$$\vec{E} + \frac{\partial \vec{A}}{\partial t} = -\vec{\nabla}\phi \,, \tag{2.7}$$

where ϕ is the electric scalar potential. Thus we can *solve* the Bianchi identities (2.4) by writing $\vec{E}$ and $\vec{B}$ in terms of scalar and 3-vector potentials ϕ and $\vec{A}$:

$$\vec{E} = -\vec{\nabla}\phi - \frac{\partial \vec{A}}{\partial t} \,, \qquad \vec{B} = \vec{\nabla} \times \vec{A} \,. \tag{2.8}$$

Although we have now "disposed of" the two Maxwell equations in (2.4), it has been achieved at a price, in that there is a redundancy in the choice of gauge potentials ϕ and $\vec{A}$. First, we may note that that $\vec{B}$ in (2.8) is unchanged if we make the replacement

$$\vec{A} \longrightarrow \vec{A} + \vec{\nabla}\lambda \,, \tag{2.9}$$

where λ is an arbitrary function of position and time. The expression for $\vec{E}$ will also be invariant, if we simultaneously make the replacement

$$\phi \longrightarrow \phi - \frac{\partial \lambda}{\partial t} \,. \tag{2.10}$$

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

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

$$\vec{\nabla} \cdot \vec{A} + \frac{\partial \phi}{\partial t} = 0 \,. \tag{2.11}$$

Note that, contrary to the belief of many physicists, this gauge choice was introduced by the Danish physicist Ludvig Valentin Lorenz, and not the Dutch physicist Hendrik Antoon Lorentz who is responsible for the Lorentz transformation. Adding to the confusion is that
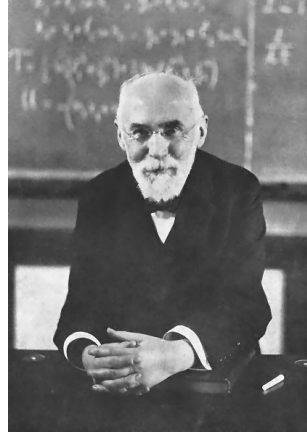
Ludvig Valentin Lorenz   Hendrik Antoon Lorentz

unlike many other gauge choices that one encounters, the Lorenz gauge condition is, as we shall see later, Lorentz invariant.

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

$$\nabla^2\phi - \frac{\partial^2\phi}{\partial t^2} = -4\pi\rho\,,$$
$$\nabla^2\vec{A} - \frac{\partial^2\vec{A}}{\partial t^2} = -4\pi\vec{J}.\qquad(2.12)$$

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

2.3 Maxwell's equations in 4-tensor notation

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

tensor. This is in fact exactly what is needed. Thus we introduce a tensor $F_{\mu\nu}$, satisfying

$$F_{\mu\nu} = -F_{\nu\mu} \,. \tag{2.13}$$

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

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

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

$$F_{23} = B_1 \,, \qquad F_{31} = B_2 \,, \qquad F_{12} = B_3 \,,$$
$$F_{32} = -B_1 \,, \qquad F_{13} = -B_2 \,, \qquad F_{21} = -B_3 \,. \tag{2.15}$$

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

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

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

$$J^0 = \rho \,, \qquad J^i = J^i \,. \tag{2.17}$$

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

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

$$\partial_\mu F^{\mu\nu} = -4\pi J^\nu \,, \tag{2.18}$$
$$\partial_\mu F_{\nu\rho} + \partial_\nu F_{\rho\mu} + \partial_\rho F_{\mu\nu} = 0 \,. \tag{2.19}$$

33

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

$$F'_{\mu\nu} = \Lambda_\mu{}^\rho \, \Lambda_\nu{}^\sigma \, F_{\rho\sigma} \,, \qquad J'^\mu = \Lambda^\mu{}_\nu \, J^\nu \,, \tag{2.20}$$

etc., where $\Lambda^\mu{}_\nu$ describes the Lorentz transformation from the frame S to the frame S'.

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

$$\partial_\mu F^{\mu 0} = \partial_0 F^{00} + \partial_i F^{i0} = \partial_i F^{i0} = -4\pi J^0 \,, \tag{2.21}$$

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

$$-\partial_i E_i = -4\pi\rho \,, \qquad \text{i.e.} \qquad \vec{\nabla} \cdot \vec{E} = 4\pi\rho \,. \tag{2.22}$$

For $\nu = j$, we shall have

$$\partial_\mu F^{\mu j} = \partial_0 F^{0j} + \partial_i F^{ij} = -4\pi J^j \,, \tag{2.23}$$

which gives

$$\partial_0 E_j + \epsilon_{ijk}\partial_i B_k = -4\pi J^j \,. \tag{2.24}$$

This is just[9]

$$-\frac{\partial \vec{E}}{\partial t} + \vec{\nabla} \times \vec{B} = 4\pi\vec{J} \,. \tag{2.25}$$

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

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

---

[9]Recall that the i'th component of $\vec{\nabla} \times \vec{V}$ is given by $(\vec{\nabla} \times \vec{V})_i = \epsilon_{ijk}\partial_j V_k$ for any 3-vector $\vec{V}$.

the $1 + 3$ decomposition $\mu = (0, i)$ etc.: Either one of the indices is a 0 with the other two Latin, or else all three are Latin. Consider first $(\mu, \nu, \rho) = (0, i, j)$:

$$\partial_0 F_{ij} + \partial_i F_{j0} + \partial_j F_{0i} = 0 \,, \tag{2.26}$$

which, from (2.14), means

$$\epsilon_{ijk} \frac{\partial B_k}{\partial t} + \partial_i E_j - \partial_j E_i = 0 \,. \tag{2.27}$$

Since this is antisymmetric in ij there is no loss of generality involved in contracting with $\epsilon_{ij\ell}$, which gives[10]

$$2 \frac{\partial B_\ell}{\partial t} + 2 \epsilon_{ij\ell} \, \partial_i E_j = 0 \,. \tag{2.28}$$

This is just the statement that

$$\vec{\nabla} \times \vec{E} + \frac{\partial \vec{B}}{\partial t} = 0 \,, \tag{2.29}$$

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

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

$$\partial_i F_{jk} + \partial_j F_{ki} + \partial_k F_{ij} = 0 \,. \tag{2.30}$$

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

$$3 \epsilon_{ijk} \, \partial_i F_{jk} = 0 \,. \tag{2.31}$$

From (2.14), this implies

$$3 \epsilon_{ijk} \epsilon_{jk\ell} \partial_i B_\ell = 0 \,, \qquad \text{and hence} \qquad 6 \partial_i B_i = 0 \,. \tag{2.32}$$

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

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

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

---

[10]Recall that $\epsilon_{ijm} \epsilon_{k\ell m} = \delta_{ik} \delta_{j\ell} - \delta_{i\ell} \delta_{jk}$, and hence $\epsilon_{ijm} \epsilon_{kjm} = 2 \delta_{ik}$.

partial derivatives commute and $F^{\mu\nu}$ is antisymmetric. Thus from the right-hand side we get

$$\partial_\mu J^\mu = 0\,. \tag{2.33}$$

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

$$\frac{\partial \rho}{\partial t} + \vec{\nabla} \cdot \vec{J} = 0\,. \tag{2.34}$$

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

$$\frac{\partial}{\partial t} \int_V \rho\, d^3 x = -\int_S \vec{J} \cdot d\vec{S}\,, \tag{2.35}$$

where $d^3 x = dxdydz$ is the spatial 3-volume element, and $d\vec{S} = (dydz, dzdx, dxdy)$ is the 2-area elemenet.

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

$$dQ \equiv \rho\, dxdydz \tag{2.36}$$

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

$$dv = dtdxdydz\,, \tag{2.37}$$

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

$$\mathcal{J} = \det\left(\frac{\partial x'^\mu}{\partial x^\nu}\right) = \det(\Lambda^\mu{}_\nu)\,. \tag{2.38}$$

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

$$\Lambda^T \eta\, \Lambda = \eta\,, \tag{2.39}$$

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

$$\det \Lambda = \pm 1\,. \tag{2.40}$$

Assuming that we restrict attention to Lorentz transformations without reflections, then they will be connected to the identity (we can take the boost velocity $\vec{v}$ to zero and/or

the rotation angle to zero and continuously approach the identity transformation), and so $\det \Lambda = 1$. Thus it follows from (2.38) that for Lorentz transformations without reflections, the 4-volume element $dv = dt\,dx\,dy\,dz$ is invariant.

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

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

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

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

$$d\Sigma_\mu = (dx\,dy\,dz, -dt\,dy\,dz, -dt\,dz\,dx, -dt\,dx\,dy)\,. \tag{2.41}$$

Charge conservation can then be described in 4-dimensional terms, by considering an arbitrary 4-volume V_4 in spacetime, which is bounded by a 3-dimensional surface S_3. Integrating $\partial_\mu J^\mu = 0$ over V_4 and then using the 4-dimensional analogue of the divergence theorem gives[11]

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

---

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

where $d^4x = dt\,dx\,dy\,dz$ is the infinitesimal 4-volume element.

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

$$0 = -\int_{\text{Cap 1}} J^0\,dx\,dy\,dz + +\int_{\text{Cap 2}} J^0\,dx\,dy\,dz + \int_{\text{Sphere at infinity}} J^i\,d\Sigma_i\,. \qquad (2.43)$$

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

$$\int_{\text{Cap 1}} J^0\,dx\,dy\,dz = \int_{\text{Cap 2}} J^0\,dx\,dy\,dz\,, \qquad \text{i.e.} \int_{t=t_1} \rho\,dx\,dy\,dz = \int_{t=t_2} \rho\,dx\,dy\,dz\,, \quad (2.44)$$

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

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

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

$$\Box A^\mu = -4\pi\,J^\mu\,, \qquad (2.45)$$

where

$$A^\mu = (\phi, \vec{A})\,, \qquad (2.46)$$

where the d'Alembertian, or wave operator, $\Box = \partial^\mu \partial_\mu = \partial_i \partial_i - \partial_0^2$ was introduced in (1.115). We saw that it is manifestly a Lorentz *scalar* operator, since it is built from the contraction

of indices on the two Lorentz-vector gradient operators. Since we have already established that J^μ is a 4-vector, it therefore follows that A^μ is a 4-vector. Note, *en passant*, that the Lorenz gauge condition (2.11) that we imposed earlier translates, in the four-dimensional language, into

$$\partial_\mu A^\mu = 0 \,, \tag{2.47}$$

which is nicely Lorentz invariant.

The final step is to note that our definition (2.14) is precisely consistent with (2.46) and (2.8), if we write

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu \,. \tag{2.48}$$

First, we note from (2.46) that because of the $\eta_{00} = -1$ needed when lowering the 0 index, we shall have

$$A_\mu = (-\phi, \vec{A}) \,. \tag{2.49}$$

Therefore we find

$$
\begin{aligned}
F_{0i} &= \partial_0 A_i - \partial_i A_0 = \frac{\partial A_i}{\partial t} + \partial_i \phi = -E_i \,, \\
F_{ij} &= \partial_i A_j - \partial_j A_i = \epsilon_{ijk} (\vec{\nabla} \times \vec{A})_k = \epsilon_{ijk} B_k \,.
\end{aligned}
\tag{2.50}
$$

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

Finally, it is worth remarking that in the 4-tensor description, the way in which the gauge invariance arises is very straightforward. First, it is manifest that the Bianchi identity (2.19) is solved identically by writing

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu \,, \tag{2.51}$$

for some 4-vector A_μ. This is because (2.19) is totally antisymmetric in $\mu\nu\rho$, and so, when (2.51) is substituted into it, one gets identically zero since partial derivatives commute. (Try making the substitution and verify this explicitly. The vanishing because of the commutativity of partial derivatives is essentially the same as the reason why **curl grad** $\equiv 0$ and **div curl** $\equiv 0$.) It is also clear from (2.51) that $F_{\mu\nu}$ will be unchanged if we make the replacement

$$A_\mu \longrightarrow A_\mu + \partial_\mu \lambda \,, \tag{2.52}$$

where λ is an arbitrary function of position and time. Again, the reason is that partial derivatives commute. Comparing (2.52) with (2.49), we see that (2.52) implies

$$\phi \longrightarrow \phi - \frac{\partial \lambda}{\partial t}, \qquad A_i \longrightarrow A_i + \partial_i \lambda, \tag{2.53}$$

and so we have reproduced the gauge transformations (2.9) and (2.10).

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

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

Although for many purposes the four-dimensional decsription of the Maxwell equations is the most convenient, it is sometimes useful to revert to the original description in terms of $\vec{E}$ and $\vec{B}$. For example, we may easily derive the Lorentz transformation properties of $\vec{E}$ and $\vec{B}$, making use of the four-dimensional formulation. In terms of $F_{\mu\nu}$, there is no work needed to write down its behaviour under Lorentz transformations. Raising the indices for convenience, we shall have

$$F'^{\mu\nu} = \Lambda^\mu{}_\rho \Lambda^\nu{}_\sigma F^{\rho\sigma} . \tag{2.54}$$

From this, and the fact (see (2.14) that $F^{0i} = E_i$, $F^{ij} = \epsilon_{ijk} B_k$, we can then immediately read of the Lorentz transformations for $\vec{E}$ and $\vec{B}$.

From the expressions (1.75) for the most general Lorentz boost transformation, we may first calculate $\vec{E}\,'$, calculated from

$$
\begin{aligned}
E'_i &= F'^{0i} = \Lambda^0{}_\rho \Lambda^i{}_\sigma F^{\rho\sigma} , \\
&= \Lambda^0{}_0 \Lambda^i{}_k F^{0k} + \Lambda^0{}_k \Lambda^i{}_0 F^{k0} + \Lambda^0{}_k \Lambda^i{}_\ell F^{k\ell} , \\
&= \gamma\Big(\delta_{ik} + \frac{\gamma-1}{v^2} v_i v_k\Big) E_k - \gamma^2 v_i v_k E_k - \gamma\, v_k \Big(\delta_{i\ell} + \frac{\gamma-1}{v^2} v_i v_\ell\Big)\epsilon_{k\ell m} B_m , \\
&= \gamma E_i + \gamma \epsilon_{ijk}\, v_j B_k - \frac{\gamma-1}{v^2} v_i v_k E_k . \tag{2.55}
\end{aligned}
$$

(Note that because $F^{\mu\nu}$ is antisymmetric, there is no F^{00} term on the right-hand side on the second line.) Thus, in terms of 3-vector notation, the Lorentz boost transformation of the electric field is given by

$$\vec{E}' = \gamma(\vec{E} + \vec{v} \times \vec{B}) - \frac{\gamma-1}{v^2} (\vec{v} \cdot \vec{E})\, \vec{v} . \tag{2.56}$$

An analogous calculation shows that the Lorentz boost transformation of the magnetic field is given by

$$\vec{B}' = \gamma(\vec{B} - \vec{v} \times \vec{E}) - \frac{\gamma - 1}{v^2}(\vec{v} \cdot \vec{B})\vec{v}. \tag{2.57}$$

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

$$\vec{B}' = \gamma\vec{B} - \frac{\gamma - 1}{v^2}(\vec{v} \cdot \vec{B})\vec{v}, \tag{2.58}$$

but also an electric field, given by

$$\vec{E}' = \gamma\vec{v} \times \vec{B}. \tag{2.59}$$

This, of course, is the principle of the dynamo.[12]

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

$$\begin{aligned}
E'_x = E_x, \qquad & E'_y = \gamma(E_y - vB_z), \qquad && E'_z = \gamma(E_z + vB_y), \\
B'_x = B_x, \qquad & B'_y = \gamma(B_y + vE_z), \qquad && B'_z = \gamma(B_z - vE_y).
\end{aligned} \tag{2.60}$$

2.5 The Lorentz force

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

$$U^\mu = \frac{dx^\mu(\tau)}{d\tau} = (\gamma, \gamma\,\vec{u}), \qquad \text{where} \qquad \gamma = \frac{1}{\sqrt{1 - u^2}}, \tag{2.61}$$

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

$$p^\mu = mU^\mu = (m\gamma, m\gamma\,\vec{u}). \tag{2.62}$$

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

$$p^\mu p_\mu = -m^2. \tag{2.63}$$

---

[12]In a practical dynamo the rotor is moving with a velocity $\vec{v}$ which is much less than the speed of light, i.e. $|\vec{v}| \ll 1$ in natural units. This means that the gamma factor $\gamma = (1 - v^2)^{-1/2}$ is approximately equal to unity in such cases.

We now define the relativistic 4-force f^μ acting on the particle to be

$$f^\mu = \frac{dp^\mu}{d\tau}\,, \tag{2.64}$$

where τ is the proper time. Clearly f^μ is indeed a 4-vector, since it is the 4-vector dp^μ divided by the scalar $d\tau$.

Using (2.62), we can write the 4-force as

$$f^\mu = \left(m\gamma^3 \vec{u} \cdot \frac{d\vec{u}}{d\tau}, m\gamma^3 \vec{u} \cdot \frac{d\vec{u}}{d\tau}\,\vec{u} + m\gamma\frac{d\vec{u}}{d\tau} \right). \tag{2.65}$$

It follows that if we move to the instantaneous rest frame of the particle, i.e. the frame in which $\vec{u} = 0$ at the particular moment we are considering, then f^μ reduces to

$$f^\mu\Big|_{\text{rest frame}} = (0, \vec{F})\,, \tag{2.66}$$

where

$$\vec{F} = m\frac{d\vec{u}}{dt} \tag{2.67}$$

is the Newtonian force measured in the rest frame of the particle.[13] Thus, we should interpret the 4-force physically as describing the Newtonian 3-force when measured in the instantaneous rest frame of the accelerating particle.

If we now suppose that the particle has electric charge e, and that it is moving under the influence of an electromagnetic field $F_{\mu\nu}$, then its motion is given by the *Lorentz force equation*

$$f^\mu = eF^{\mu\nu}\,U_\nu\,. \tag{2.68}$$

One can more or less justify this equation on the grounds of "what else could it be?", since we know that there must exist a relativistic equation (i.e. a Lorentz covariant equation) that describes the motion. In fact it is easy to see that (2.68) is correct. We calculate the spatial components:

$$\begin{aligned}
f^i &= eF^{i\nu}\,U_\nu = eF^{i0}\,U_0 + eF^{ij}\,U_j\,,\\
&= e(-E_i)(-\gamma) + e\epsilon_{ijk}\,B_k\gamma\,u_j\,, \tag{2.69}
\end{aligned}$$

and thus

$$\vec{f} = e\gamma\,(\vec{E} + \vec{u}\times\vec{B})\,. \tag{2.70}$$

---

[13]Note that we can replace the proper time τ by the coordinate time t in the instantaneous rest frame, since $d\tau = dt/\gamma$, and $\gamma = 1$ when $\vec{u} = 0$.

But $f^\mu = dp^\mu/d\tau$, and so $\vec{f} = d\vec{p}/d\tau = \gamma\, d\vec{p}/dt$ (recall from section 1.6 that $d\tau = dt/\gamma$) and so we have

$$\frac{d\vec{p}}{dt} = e\,(\vec{E} + \vec{u}\times\vec{B})\,, \tag{2.71}$$

where $d\vec{p}/dt$ is the rate of change of th relativistic 3-momentum $\vec{p} = m\,\gamma\,u\vec{u}$. This is indeed the standard expression for the motion of a charged particle under the Lorentz force.

2.6 Action principle for charged particles

In this section, we shall show how the equations of motion for a charged particle moving in an electromagnetic field can be derived from an action principle. To begin, we shall consider an uncharged particle of mass m, with no forces acting on it. It will, of course, move in a straight line. It turns out that its equation of motion can be derived from the Lorentz-invariant action

$$S = -m\int_{\tau_1}^{\tau_2} d\tau\,, \tag{2.72}$$

where τ is the proper time along the trajectory $x^\mu(\tau)$ of the particle, starting at proper time $\tau = \tau_1$ and ending at $\tau = \tau_2$. The action principle then states that if we consider all possible paths between the initial and final spacetime points on the path, then the actual path followed by the particle will be such that the action S is stationary. In other words, if we consider small variations of the path around the actual path, then to first order in the variations we shall have $\delta S = 0$.

To see how this works, we note that $d\tau^2 = dt^2 - dx^i dx^i = dt^2(1 - v_i v_i) = dt^2(1 - v^2)$, where $v_i = dx^i/dt$ is the 3-velocity of the particle. Thus $d\tau = dt/\gamma$ where $\gamma = (1 - v^2)^{-1/2}$ and so

$$S = -m\int_{t_1}^{t_2}\frac{dt}{\gamma} = -m\int_{t_1}^{t_2}(1 - v^2)^{1/2}\,dt = -m\int_{t_1}^{t_2}(1 - \dot{x}^i\dot{x}^i)^{1/2}\,dt\,. \tag{2.73}$$

In other words, the Lagrangian L, for which $S = \int_{t_1}^{t_2} Ldt$, is given by

$$L = -\frac{m}{\gamma} = -m(1 - \dot{x}^i\dot{x}^i)^{1/2}\,. \tag{2.74}$$

As a check, if we expand (2.74) for small velocities (i.e. small compared with the speed of light, so $|\dot{x}^i| << 1$), we shall have

$$L = -m + \tfrac{1}{2}mv^2 + \cdots\,. \tag{2.75}$$

Since the Lagrangian is given by $L = T - V$ we see that T is just the usual kinetic energy $\tfrac{1}{2}mv^2$ for a non-relativistic particle of mass m, while the potential energy is just m. Of course if we were not using units where the speed of light were unity, this energy would be

mc^2. Since it is just a constant, this rest-mass energy of the particle does not affect the equations of motion that will follow from the action principle.

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

$$\delta S = m \int_{t_1}^{t_2} (1 - \dot{x}^j \dot{x}^j)^{-1/2} \dot{x}^i \delta \dot{x}^i dt \,. \tag{2.76}$$

Integrating by parts then gives

$$\delta S = -m \int_{t_1}^{t_2} \frac{d}{dt}\Big((1 - \dot{x}^j \dot{x}^j)^{-1/2} \dot{x}^i\Big) \delta x^i dt + m\Big[(1 - \dot{x}^j \dot{x}^j)^{-1/2} \dot{x}^i \delta x^i\Big]_{t_1}^{t_2} \,. \tag{2.77}$$

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

$$\frac{d}{dt}\Big(m(1 - \dot{x}^j \dot{x}^j)^{-1/2} \dot{x}^i\Big) = 0 \,. \tag{2.78}$$

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

$$\frac{d(m\gamma \vec{v})}{dt} = 0 \,, \tag{2.79}$$

or, in other words,

$$\frac{d\vec{p}}{dt} = 0 \,, \tag{2.80}$$

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

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

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu \,. \tag{2.81}$$

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

$$S = \int_{\tau_1}^{\tau_2} (-m d\tau + e A_\mu dx^\mu) \,. \tag{2.82}$$

Note that it is again Lorentz invariant.

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

$$A_\mu dx^\mu = A_\mu \frac{dx^\mu}{dt} dt = (A_0 + A_i \dot{x}^i) dt = (-\phi + A_i \dot{x}^i) dt \,. \tag{2.83}$$

Thus we have $S = \int_{t_1}^{t_2} L\,dt$ with the Lagrangian L given by

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

where potentials ϕ and A_i depend on t and x. The first-order variation of the action under a variation δx^i in the path gives

$$
\begin{aligned}
\delta S &= \int_{t_1}^{t_2} \Big[m(1 - \dot{x}^j \dot{x}^j)^{-1/2} \dot{x}^i \delta \dot{x}^i - e\partial_i \phi\, \delta x^i + eA_i \delta \dot{x}^i + e\partial_j A_i \dot{x}^i \delta x^j \Big] dt \,, \\
&= \int_{t_1}^{t_2} \Big[-\frac{d}{dt}(m\gamma \dot{x}^i) - e\partial_i \phi - e\frac{dA_i}{dt} + e\partial_i A_j\, \dot{x}^j \Big] \delta x^i dt \,. \tag{2.85}
\end{aligned}
$$

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

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

Now, the total time derivative dA_i/dt has two contributions, and we may write it as

$$\frac{dA_i}{dt} = \frac{\partial A_i}{\partial t} + \partial_j A_i\, \frac{dx^j}{dt} = \frac{\partial A_i}{\partial t} + \partial_j A_i\, \dot{x}^j \,. \tag{2.87}$$

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

Putting all this together, we have

$$
\begin{aligned}
\frac{d(m\gamma \dot{x}^i)}{dt} &= e\Big(-\partial_i \phi - \frac{\partial A_i}{\partial t} \Big) + e(\partial_i A_j - \partial_j A_i)\, \dot{x}^j \,, \\
&= e(E_i + \epsilon_{ijk}\, \dot{x}^j B_k) \,. \tag{2.88}
\end{aligned}
$$

In other words, we have

$$\frac{d\vec{p}}{dt} = e(\vec{E} + \vec{v} \times \vec{B}) \,, \tag{2.89}$$

which is the Lorentz force equation (2.71).

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

$$
\begin{aligned}
\delta(d\tau) &= -\eta_{\mu\nu}\, \frac{dx^\nu}{d\tau}\, d(\delta x^\mu) \,, \\
&= -U_\mu\, d(\delta x^\mu) \,, \tag{2.90}
\end{aligned}
$$

45

where U_μ is the 4-velocity. Thus the variation of the action (2.82) gives

$$
\begin{aligned}
\delta S & = \int_{\tau_1}^{\tau_2} \Big(m U_\mu d\delta x^\mu + e A_\mu \, d\delta x^\mu + e \partial_\nu A_\mu \, \delta x^\nu dx^\mu \Big) , \\
& = \int_{\tau_1}^{\tau_2} \Big(d(m U_\mu \, \delta x^\mu) - m dU_\mu \, \delta x^\mu + d(e A_\mu \, \delta x^\mu) - e dA_\mu \, \delta x^\mu + e \partial_\mu A_\nu \, \delta x^\mu dx^\nu \Big) , \\
& = \int_{\tau_1}^{\tau_2} \Big(- m dU_\mu \, \delta x^\mu - e dA_\mu \, \delta x^\mu + e \partial_\mu A_\nu \, \delta x^\mu dx^\nu \Big) , \\
& = \int_{\tau_1}^{\tau_2} \Big(- m \frac{dU_\mu}{d\tau} - e \frac{dA_\mu}{d\tau} + e \partial_\mu A_\nu \frac{dx^\nu}{d\tau} \Big) \delta x^\mu d\tau ,
\end{aligned}
\tag{2.91}
$$

where we have dropped the boundary terms $\int_{\tau_1}^{\tau_2} d(m U_\mu \, \delta x^\mu + e A_\mu \, \delta x^\mu)$ in getting to the third line, since they integrate to give $[m U_\mu \, \delta x^\mu + e A_\mu \, \delta x^\mu]_{\tau_1}^{\tau_2}$ and therefore vanish, as we assume $\delta x^\mu = 0$ at the initial and final proper times τ_1 and τ_2. Now by the chain rule

$$
\frac{dA_\mu}{d\tau} = \partial_\nu A_\mu \frac{dx^\nu}{d\tau} = \partial_\nu A_\mu U^\nu ,
\tag{2.92}
$$

and so

$$
\begin{aligned}
\delta S & = \int_{\tau_1}^{\tau_2} \Big(- m \frac{dU_\mu}{d\tau} - e \partial_\nu A_\mu \, U^\nu + e \partial_\mu A_\nu \, U^\nu \Big) \delta x^\mu d\tau , \\
& = \int_{\tau_1}^{\tau_2} \Big(- m \frac{dU_\mu}{d\tau} + e F_{\mu\nu} \, U^\nu \Big) d\tau .
\end{aligned}
\tag{2.93}
$$

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

$$
m \frac{dU_\mu}{d\tau} = e F_{\mu\nu} \, U^\nu .
\tag{2.94}
$$

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

$$
\frac{dp^\mu}{d\tau} = e F^{\mu\nu} \, U_\nu ,
\tag{2.95}
$$

where $p^\mu = m U^\mu$ is the 4-momentum.

2.7 Gauge invariance of the action

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

can be seen from the fact that, as we demonstrated, the variational principle for the action (2.82) does in fact produce the correct gauge-invariant Lorentz force equation (2.71).

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

$$
\begin{aligned}
S' &= \int_{\tau_1}^{\tau_2} (-md\tau + eA_\mu dx^\mu + e\partial_\mu \lambda dx^\mu)\,, \\
&= S + e \int_{\tau_1}^{\tau_2} \partial_\mu \lambda dx^\mu = e \int_{\tau_1}^{\tau_2} d\lambda\,,
\end{aligned}
\tag{2.96}
$$

and so

$$
S' = S + e[\lambda(\tau_2) - \lambda(\tau_1)]\,.
\tag{2.97}
$$

The simplest situation to consider is where we restrict ourselves to gauge transformations that vanish at the endpoints, in which case the action will be gauge invariant, $S' = S$. Even if λ is non-vanishing at the endpoints, we see from (2.97) that S and S' merely differ by a constant that depends solely on the values of λ at τ_1 and τ_2. Clearly, the addition of this constant has no effect on the equations of motion that one derives from S'.

2.8 Canonical momentum, and Hamiltonian

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

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

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

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

or, in other words,

$$
\begin{aligned}
\pi_i &= m\gamma\, \dot{x}^i + eA_i\,, \\
&= p_i + eA_i\,,
\end{aligned}
$$

$$
\tag{2.100}
$$
$$
\tag{2.101}
$$

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

As usual, the Hamiltonian for the system is given by

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

and so we find

$$
\begin{aligned}
H &= m\gamma\dot{x}^i\dot{x}^i + \frac{m}{\gamma} + e\phi\,, \\
&= m\gamma v^2 + \frac{m}{\gamma} + e\phi\,.
\end{aligned}
\tag{2.103}
$$

47

Now, $m\gamma v^2 + m/\gamma = m\gamma(v^2 + (1-v^2)) = m\gamma$, so we have

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

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

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

with H expressed as a function of the coordinates x^i and the canonical momenta π_i.

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

$$\frac{\partial H}{\partial \pi_i} = \dot{x}^i, \qquad \frac{\partial H}{\partial x^i} = -\dot{\pi}_i. \tag{2.106}$$

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

$$\begin{aligned} H &= m\sqrt{1 + (\pi_i - eA_i)^2/m^2} + e\phi, \\ &= m + \frac{1}{2m}(\pi_i - eA_i)^2 + e\phi + \cdots. \end{aligned} \tag{2.107}$$

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

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

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

3 Particle Motion in Static Electromagnetic Fields

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

3.1 Description in terms of potentials

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

$$
\begin{aligned}
\vec{E} &= -\vec{\nabla}\phi - \frac{\partial \vec{A}}{\partial t} = -\vec{\nabla}\phi(\vec{r})\,, \\
\vec{B} &= \vec{\nabla} \times \vec{A}(\vec{r})\,.
\end{aligned}
\tag{3.1}
$$

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

$$
\lambda(\vec{r},t) = \lambda(\vec{r}) + k\,t\,,
\tag{3.2}
$$

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

$$
\phi \longrightarrow \phi - k\,, \qquad \vec{A} \longrightarrow \vec{A} + \vec{\nabla}\lambda(\vec{r})\,.
\tag{3.3}
$$

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

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

$$
H = m\gamma + e\phi\,,
\tag{3.4}
$$

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

$$
\begin{aligned}
\frac{dH}{dt} &= \frac{\partial H}{\partial t} + \frac{\partial H}{\partial x^i}\,\dot{x}^i + \frac{\partial H}{\partial \pi_i}\,\dot{\pi}_i\,, \\
&= 0 - \dot{\pi}_i\,\dot{x}^i + \dot{x}^i\,\dot{\pi}_i = 0\,.
\end{aligned}
\tag{3.5}
$$

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

$$
\mathcal{E} \equiv H = m\gamma + e\phi\,.
\tag{3.6}
$$

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

$$
\mathcal{E}_{\text{mech}} = m\gamma\,,
\tag{3.7}
$$

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

Recall that the Lorentz force equation can be written as

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

Multiplying by v^i we therefore have

$$m\gamma v^i \frac{dv^i}{dt} + m v^i v^i \frac{d\gamma}{dt} = e v^i E_i\,. \tag{3.9}$$

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

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

and so (3.9) gives

$$m\frac{d\gamma}{dt} = e v^i E_i\,. \tag{3.11}$$

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

$$\frac{d\mathcal{E}_{\text{mech}}}{dt} = e\vec{v} \cdot \vec{E}\,. \tag{3.12}$$

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

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

$$\begin{aligned}
\frac{d\mathcal{E}}{dt} &= \frac{d(m\gamma)}{dt} + e\frac{d\phi}{dt} \\
&= \frac{d\mathcal{E}_{\text{mech}}}{dt} + e\partial_i\phi\,\frac{dx^i}{dt}\,, \\
&= e\vec{v} \cdot \vec{E} - e\vec{v} \cdot \vec{E} = 0\,.
\end{aligned} \tag{3.13}$$

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

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

$$\phi = -\vec{E} \cdot \vec{r} = -E_i x^i\,. \tag{3.14}$$

Clearly this gives the correct electric field, since

$$-\partial_i \phi = \partial_i(E_j x^j) = E_j \partial_i x^j = E_j \delta_{ij} = E_i \,. \tag{3.15}$$

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

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

$$\vec{A} = \tfrac{1}{2}\vec{B} \times \vec{r} \,. \tag{3.16}$$

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

$$
\begin{aligned}
(\vec{\nabla} \times \vec{A})_i &= \epsilon_{ijk}\,\partial_j A_k = \epsilon_{ijk}\partial_j\big(\tfrac{1}{2}\epsilon_{k\ell m}B_\ell x^m\big)\,, \\
&= \tfrac{1}{2}\epsilon_{ijk}\epsilon_{\ell m k}\,B_\ell \partial_j x^m = \tfrac{1}{2}\epsilon_{ijk}\epsilon_{\ell jk}\,B_\ell\,, \\
&= \delta_{i\ell}\,B_\ell = B_i\,.
\end{aligned}
\tag{3.17}
$$

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

$$\vec{A} = (-\tfrac{1}{2}By, \tfrac{1}{2}Bx, 0)\,. \tag{3.18}$$

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

$$\vec{A}' = (-By, 0, 0)\,. \tag{3.19}$$

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

3.2.1 Motion in a static uniform electric field

From the Lorentz force equation, we shall have

$$\frac{d\vec{p}}{dt} = e\vec{E}\,, \tag{3.20}$$

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

$$\frac{dp_x}{dt} = eE\,, \qquad \frac{dp_y}{dt} = 0\,, \qquad \frac{dp_z}{dt} = 0\,. \tag{3.21}$$

Since p_y and p_z are therefore constants, we can without loss of generality rotate the coordinate system around the x axis so that $p_z = 0$. Thus we may integrate (3.21 to give

$$p_x = eEt\,, \qquad p_y = \bar{p}\,, \qquad p_z = 0\,, \tag{3.22}$$

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

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

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

Hence we may write

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

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

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

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

which can be integrated to give

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

(The constant of integration has been absorbed into a choice of origin for the x coordinate.) Note from (3.25) that the x-component of the 3-velocity asymptotically approaches 1 as t goes to infinity. Thus the particle is accelerated closer and closer to the speed of light, but never reaches it.

We also have

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

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

$$eEt = \mathcal{E}_0\,\sinh u\,. \tag{3.28}$$

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

$$y = \frac{\bar{p}}{eE}\,\text{arcsinh}\Big(\frac{eEt}{\mathcal{E}_0}\Big)\,. \tag{3.29}$$

(Again, the constant of integration has been absorbed into the choice of origin for y.)

The solutions (3.26) and (3.29) for x and y as functions of t can be combined to give x as a function of y, leading to

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

(3.30)

This is a catenary.

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

$$x \approx \text{constant} + \frac{eE}{2m\bar{v}^2} y^2.$$

(3.31)

3.2.2 Motion in a static uniform magnetic field

From the Lorentz force equation we shall have

$$\frac{d\vec{p}}{dt} = e\vec{v} \times \vec{B}.$$

(3.32)

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

$$\frac{d\vec{v}}{dt} = \frac{e}{m\gamma} \vec{v} \times \vec{B} = \frac{e}{\mathcal{E}} \vec{v} \times \vec{B},$$

(3.33)

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

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

$$\omega \equiv \frac{eB}{\mathcal{E}} = \frac{eB}{m\gamma},$$

(3.34)

we then find

$$\frac{dv_x}{dt} = \omega \, v_y, \qquad \frac{dv_y}{dt} = -\omega \, v_x, \qquad \frac{dv_z}{dt} = 0.$$

(3.35)

From this, it follows that

$$\frac{d(v_x + \mathrm{i}\, v_y)}{dt} = -\mathrm{i}\, \omega \, (v_x + \mathrm{i}\, v_y),$$

(3.36)

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

$$v_x + \mathrm{i}\, v_y = v_0 \, e^{-\mathrm{i}\,(\omega t + \alpha)},$$

(3.37)

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

$$x = x_0 + r_0 \sin(\omega t + \alpha), \qquad y = y_0 + r_0 \cos(\omega t + \alpha), \qquad z = z_0 + \bar{v}_z t,$$

(3.38)

for constants r_0, x_0, y_0, z_0 and $\bar{v}_z$, with

$$r_0 = \frac{v_0}{\omega} = \frac{m\gamma v_0}{eB} = \frac{\bar{p}}{eB}\,, \tag{3.39}$$

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

3.2.3 Motion in uniform $\vec{E}$ and $\vec{B}$ fields

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

$$\frac{dU^\mu}{d\tau} = \frac{e}{m}\, F^\mu{}_\nu\, U^\nu\,, \tag{3.40}$$

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

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

$$\vec{B} = (0, 0, B)\,, \qquad \vec{E} = (0, E_y, E_z)\,, \tag{3.41}$$

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

$$m\frac{d\vec{v}}{dt} = e(\vec{E} + \vec{v} \times \vec{B})\,, \tag{3.42}$$

and so

$$m\ddot{x} = eB\dot{y}\,, \qquad m\ddot{y} = eE_y - eB\dot{x}\,, \qquad m\ddot{z} = eE_z\,. \tag{3.43}$$

We can immediately solve for z, finding

$$z = \frac{e}{2m}\, E_z\, t^2 + \bar{v}t\,, \tag{3.44}$$

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

$$\frac{d}{dt}(\dot{x} + \mathrm{i}\,\dot{y}) + \mathrm{i}\,\omega(\dot{x} + \mathrm{i}\,\dot{y}) = \frac{\mathrm{i}\,e}{m}\, E_y\,, \tag{3.45}$$

where $\omega = eB/m$. Thus we find

$$\dot{x} + \mathrm{i}\,\dot{y} = a e^{-\mathrm{i}\,\omega t} + \frac{e}{m\omega}\,E_y = a e^{-\mathrm{i}\,\omega t} + \frac{E_y}{B}\,. \tag{3.46}$$

Choosing the origin of time so that a is real, we have

$$\dot{x} = a\cos\omega t + \frac{E_y}{B}\,, \qquad \dot{y} = -a\sin\omega t\,. \tag{3.47}$$

Taking the time averages, we see that

$$\langle\dot{x}\rangle = \frac{E_y}{B}\,, \qquad \langle\dot{y}\rangle = 0\,. \tag{3.48}$$

The averaged velocity along the x direction is called the *drift velocity*. Notice that it is perpendicular to $\vec{E}$ and $\vec{B}$. It can be written in general as

$$\vec{v}_{\mathrm{drift}} = \frac{\vec{E}\times\vec{B}}{B^2}\,. \tag{3.49}$$

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

Integrating (3.47) once more, we find

$$x = \frac{a}{\omega}\,\sin\omega t + \frac{E_y}{B}\,t\,, \qquad y = \frac{a}{\omega}\,(\cos\omega t - 1)\,, \tag{3.50}$$

where the origins of x and y have been chosen so that $x = y = 0$ at $t = 0$. These equations describe the projection of the particle's motion onto the (x, y) plane. The curve is called a *trochoid*. If $|a| > E_y/B$ there will be loops in the motion, and in the special case $a = -E_y/B$ the curve becomes a *cycloid*, with cusps:

$$x = \frac{E_y}{\omega B}\,(\omega t - \sin\omega t)\,, \qquad y = \frac{E_y}{\omega B}\,(1 - \cos\omega t)\,. \tag{3.51}$$

4 Action Principle for Electrodynamics

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

4.1 Invariants of the electromagnetic field

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

4.1.1 The first invariant

The first quadratic invariant is very simple; we may write

$$I_1 \equiv F_{\mu\nu} \, F^{\mu\nu} \,. \tag{4.1}$$

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

$$
\begin{aligned}
I_1 &= F_{0i} \, F^{0i} + F_{i0} \, F^{i0} + F_{ij} \, F^{ij} \,, \\
&= 2F_{0i} \, F^{0i} + F_{ij} \, F^{ij} = -2E_i E_i + \epsilon_{ijk} \, B_k \, \epsilon_{ij\ell} \, B_\ell \,, \\
&= -2E_i E_i + 2B_i B_i \,,
\end{aligned}
\tag{4.2}
$$

and so

$$I_1 \equiv F_{\mu\nu} \, F^{\mu\nu} = 2(\vec{B}^2 - \vec{E}^2) \,. \tag{4.3}$$

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

4.1.2 The second invariant

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

$$I_2 \equiv \tfrac{1}{2} \epsilon^{\mu\nu\rho\sigma} \, F_{\mu\nu} F_{\rho\sigma} \,. \tag{4.4}$$

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

$$\epsilon^{\mu\nu\rho\sigma} = -\epsilon^{\nu\mu\rho\sigma} = -\epsilon^{\mu\nu\sigma\rho} = -\epsilon^{\sigma\nu\rho\mu} \,. \tag{4.5}$$

---

[14]Beware that in an odd dimension, such as 3, the process of "cycling" the indices on ϵ_{ijk} (for example, pushing one off the right-hand end and bringing it to the front) is an *even* permutation; $\epsilon_{kij} = \epsilon_{ijk}$. By contrast, in an even dimension, such as 4, the process of cycling is an *odd* permutation; $\epsilon^{\sigma\mu\nu\rho} = -\epsilon^{\mu\nu\rho\sigma}$. This is an elementary point, but easily overlooked if one is familiar only with three dimensions!

Since all the non-vanishing components of $\epsilon^{\mu\nu\rho\sigma}$ are related by the antisymmetry, we need only specify one non-vanishing component in order to define the tensor completely. We shall define

$$\epsilon^{0123} = -1\,, \qquad \text{or, equivalently} \qquad \epsilon_{0123} = +1\,. \tag{4.6}$$

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

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

$$\begin{aligned}
\epsilon'^{\mu\nu\rho\sigma} &\equiv& \Lambda^\mu{}_\alpha \Lambda^\nu{}_\beta \Lambda^\rho{}_\gamma \Lambda^\sigma{}_\delta \,\epsilon^{\alpha\beta\gamma\delta}\,, \\
&=& (\det \Lambda)\,\epsilon^{\mu\nu\rho\sigma}\,.
\end{aligned} \tag{4.7}$$

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

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

$$\begin{aligned}
I_2 &=& \tfrac{1}{2}\epsilon^{\mu\nu\rho\sigma}\,F_{\mu\nu}F_{\rho\sigma} = \tfrac{1}{2}\times 4\times \epsilon^{0ijk}F_{0i}\,F_{jk}\,, \\
&=& 2(-\epsilon_{ijk})(-E_i)\epsilon_{jk\ell}B_\ell\,, \\
&=& 4E_i B_i = 4\vec{E}\cdot\vec{B}\,.
\end{aligned} \tag{4.8}$$

Thus, to summarise, we have the two quadratic invariants

$$\begin{aligned}
I_1 &=& F_{\mu\nu}F^{\mu\nu} = 2(\vec{B}^2 - \vec{E}^2)\,, \\
I_2 &=& \tfrac{1}{2}\epsilon^{\mu\nu\rho\sigma}\,F_{\mu\nu}F_{\rho\sigma} = 4\vec{E}\cdot\vec{B}\,.
\end{aligned} \tag{4.9}$$

Since the two quantities I_1 and I_2 are (manifestly) Lorentz invariant, this means that, even though it is not directly evident in the three-dimensional language without quite a lot of work, the two quantities

$$\vec{B}^2 - \vec{E}^2\,, \qquad \text{and} \qquad \vec{E} \cdot \vec{B} \tag{4.10}$$

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

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

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

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

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

4.2 Action for electrodynamics

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

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

$$L = \int \mathcal{L} d^3 x\,. \tag{4.11}$$

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

$$S = \int_{t_1}^{t_2} L dt = \int_V \mathcal{L} d^4 x\,. \tag{4.12}$$

Here, we are defining the 4-volume V to be the entire infinite spatial 3-space in the sandwich between the initial time surface at $t = t_1$ and the final time surface at $t = t_2$. In order to

58

make the following discussion more precise, it will be helpful to think of the infinite 3-space as being the limit in which one sends the radius R of a 3-dimensional ball to infinity. While R is still finite, the corresponding 4-volume will be of the form of the interior of a "cylinder" that is coaxial with the time coordinate, with end caps at $t = t_1$ and $t = t_2$, an having radius R in the three spatial directions. We cannot visualise such a 4-dimensional cylinder, but it is really like the higher-dimensional analogue of a tin can, with the time coordinate coaxial with the can. A $t =$ constant slice through a familiar tin can would be a circle; in the present case, the $t =$ constant slice will be the surface of a 2-dimensional sphere of radius R. Eventually, when we send R to infinity this sphere becomes the "sphere at infinity" in 3-space.

Consider first the vacuum Maxwell equations without sources,

$$\partial_\mu F^{\mu\nu} = 0\,, \qquad \partial_\mu F_{\nu\rho} + \partial_\nu F_{\rho\mu} + \partial_\rho F_{\mu\nu} = 0\,. \tag{4.13}$$

We immediately solve the second equation (the Bianchi identity) by writing $F_{\mu\nu}$ in terms of a potential:

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu\,. \tag{4.14}$$

Since the Maxwell field equations are linear in the fields, it is natural to expect that the action should be quadratic. In fact, it turns out that the first invariant we considered above provides the appropriate Lagrangian density. We take

$$\mathcal{L} = -\frac{1}{16\pi}\,F_{\mu\nu}F^{\mu\nu}\,, \tag{4.15}$$

and so the action will be

$$S = -\frac{1}{16\pi}\int_V F_{\mu\nu}F^{\mu\nu}d^4x\,. \tag{4.16}$$

We can now derive the source-free Maxwell equations by requiring that this action be stationary with respect to variations of the gauge field A_μ. It must be emphasised that we treat A_μ as the fundamental field here.

The derivation goes as follows. We shall have

$$
\begin{aligned}
\delta S &= -\frac{1}{16\pi}\int_V (\delta F_{\mu\nu}F^{\mu\nu} + F_{\mu\nu}\delta F^{\mu\nu})d^4x = -\frac{1}{8\pi}\int_V \delta F_{\mu\nu}F^{\mu\nu}\,d^4x\,,\\
&= -\frac{1}{8\pi}\int_V F^{\mu\nu}\,(\partial_\mu \delta A_\nu - \partial_\nu \delta A_\mu)d^4x = -\frac{1}{4\pi}\int_V F^{\mu\nu}\,\partial_\mu \delta A_\nu d^4x\,,\\
&= -\frac{1}{4\pi}\int_V \partial_\mu(F^{\mu\nu}\,\delta A_\nu)d^4x + \frac{1}{4\pi}\int_V (\partial_\mu F^{\mu\nu})\,\delta A_\nu d^4x\,,\\
&= -\frac{1}{4\pi}\int_\Sigma F^{\mu\nu}\,\delta A_\nu d\Sigma_\mu + \frac{1}{4\pi}\int_V (\partial_\mu F^{\mu\nu})\,\delta A_\nu d^4x\,,\\
&= \frac{1}{4\pi}\int_V (\partial_\mu F^{\mu\nu})\,\delta A_\nu d^4x\,. \tag{4.17}
\end{aligned}
$$

Note that in the penultimate step, we have used the 4-dimensional analogue of the divergence theorem to turn the 4-volume integral of the divergence of a vector into a 3-volume integral over the bounding surface Σ. The four-dimensional divergence theorem says that for any 4-vector W^μ, we have

$$\int_V (\partial_\mu W^\mu)\, d^4x = \int_\Sigma W^\mu d\Sigma_\mu\,. \tag{4.18}$$

In our case, as discussed above, the 4-volume V consists of the "sandwich" of all of 3-space between the two surfaces $t = t_1$ and $t = t_2$. We think of this as the limit of a sandwich between $t = t_1$ and t_2 extending out to spatial radius R, where eventually R is sent to infinity. The 3-dimensional boundary Σ of this "tin can" consists of the endcaps at $t = t_1$ and $t = t_2$, plus the spatial sphere of radius R along the sides of the can. The contributions from the integration over Σ have been dropped in getting to the final line in eqn (4.17). This happens for two reasons: Firstly, by definition we shall require our variations δA_μ to vanish on the end-caps at $t = t_1$ and t_2. This is part of the specification of the variational problem. It is the analogue of what one always does in a variational problem in paricle mechanics: the variations in the path are required to vanish at the initial and final times. We also require the electromagnetic field to fall off at spatial infinity. This ensures that the portion of the surface integral in the penultimate line of (4.17) that is evaluated over the sides of the "tin can" goes to zero as we send the radius R to infinity.

Finally, we argue that if the action S is to be stationary for all possible infinitesimal variations δA_μ that vanish at $t = t_1$ and $t = t_2$, it must be that the cofactor of δA_ν in the final line of (4.17) must vanish:

$$\partial_\mu F^{\mu\nu} = 0\,. \tag{4.19}$$

Thus we have derived the source-free Maxwell field equation. Of course the Bianchi identity has already been taken care of by writing $F_{\mu\nu}$ in terms of the 4-vector potential A_μ.

The action (4.16), whose variation gave the Maxwell field equation, is written in what is called *second-order formalism*; that is, the action is expressed in terms of the 4-vector-potential A_μ as the fundamental field, with $F_{\mu\nu}$ just being a short-hand notation for $\partial_\mu A_\nu - \partial_\nu A_\mu$. It is sometimes convenient to use instead the *first-order formalism*, in which one treats A_μ and $F_{\mu\nu}$ as *independent* fields. In this formalism, the equation of motion coming from demanding that S be stationary under variations of $F_{\mu\nu}$ will *derive* the equation $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$. To do this, we need a different action as our starting point, namely

$$S_{\text{f.o.}} = \frac{1}{4\pi}\int \left(\tfrac{1}{4}F^{\mu\nu}F_{\mu\nu} - F^{\mu\nu}\,\partial_\mu A_\nu\right)d^4x\,. \tag{4.20}$$

First, consider the variation of $F^{\mu\nu}$, *now treated as an independent fundamental field*. This gives

$$
\begin{aligned}
\delta S_{\text{f.o.}} &= \frac{1}{4\pi} \int (\tfrac{1}{2} F_{\mu\nu} \delta F^{\mu\nu} - \delta F^{\mu\nu} \partial_\mu A_\nu) d^4 x \,, \\
&= \frac{1}{4\pi} \int [\tfrac{1}{2} F_{\mu\nu} \delta F^{\mu\nu} - \tfrac{1}{2} \delta F^{\mu\nu} (\partial_\mu A_\nu - \partial_\nu A_\mu)] d^4 x \,,
\end{aligned}
\tag{4.21}
$$

where, in getting to the second line, we have used the fact that $F^{\mu\nu}$ is antisymmetric. The reason for doing this is that when we vary $F^{\mu\nu}$ we can take $\delta F^{\mu\nu}$ to be arbitary, but it must still be *antisymmetric*. Thus it is helpful to force an explicit antisymmetrisation on the $\partial_\mu A_\nu$ that multiplies it, since the symmetric part automatically gives zero when contracted onto the antisymmetric $\delta F^{\mu\nu}$. Requiring $\delta S_{\text{f.o.}} = 0$ for arbitrary $\delta F^{\mu\nu}$ then implies the integrand must vanish. This gives, as promised, the equation of motion

$$
F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu \,.
\tag{4.22}
$$

Vraying $S_{\text{f.o.}}$ in (4.20) instead with respect to A_μ, we get

$$
\begin{aligned}
\delta S_{\text{f.o.}} &= -\frac{1}{4\pi} \int F^{\mu\nu} \partial_\mu \delta A_\nu \, d^4 x \,, \\
&= \frac{1}{4\pi} \int (\partial_\mu F^{\mu\nu}) \, \delta A_\nu \, d^4 x \,,
\end{aligned}
\tag{4.23}
$$

and hence reuiring that the variation of $S_{\text{f.o.}}$ with respect to A_μ vanish gives the Maxwell field equation

$$
\partial_\mu F^{\mu\nu} = 0
\tag{4.24}
$$

again. Note that in this calculation, we immediately dropped the boundary term coming from the integration by parts, for the usual reason that we only allow variations that vanish on the boundary.

In practice, we shall usually use the previous, second-order, formalism.

4.3 Inclusion of sources

In general, the Maxwell field equation reads

$$
\partial_\mu F^{\mu\nu} = -4\pi J^\nu \,.
\tag{4.25}
$$

So far, we have seen that by varying the second-order action (4.16) with respect to A_μ, we obtain

$$
\delta S = \frac{1}{4\pi} \int \partial_\mu F^{\mu\nu} \, \delta A_\nu \, d^4 x \,.
\tag{4.26}
$$

61

To derive the Maxwell field equation with a source current J^μ, we can simply add a term to the action, to give

$$S = \int \left(-\frac{1}{16\pi} F_{\mu\nu} F^{\mu\nu} + J^\mu A_\mu \right) d^4 x \,. \tag{4.27}$$

Treating J^μ as *independent* of A_μ, we therefore find

$$\delta S = \int \left(\frac{1}{4\pi} \partial_\mu F^{\mu\nu} + J^\nu \right) \delta A_\nu \, d^4 x \,, \tag{4.28}$$

and so requiring $\delta S = 0$ gives the Maxwell field equation (4.25) with the source on the right-hand side.

The form of the source current J^μ depends, of course, on the details of the situation one is considering. One might simply have a situation where J^μ is an externally-supplied source field. Alternatively, the source J^μ might itself be given dynamically in terms of some charged matter fields, or in terms of a set of moving point charges. Let us consider this possibility in more detail.

If there is a single point charge q at the location $\vec{r}_0$, then it will be described by the charge density

$$\rho = q \, \delta^3(\vec{r} - \vec{r}_0) \,, \tag{4.29}$$

where the three-dimensional delta-function $\delta^3(\vec{r})$, with $\vec{r} = (x, y, z)$, means

$$\delta^3(\vec{r}) = \delta(x)\delta(y)\delta(z) \,. \tag{4.30}$$

If the charge is moving, so that its location at time t is at $\vec{r} = \vec{r}_0(t)$, then of course we shall have

$$\rho = q \, \delta^3(\vec{r} - \vec{r}_0(t)) \,. \tag{4.31}$$

The 3-vector current will be given by multiplying this by the 3-velocity $\vec{v} = \frac{d\vec{r}_0(t)}{dt}$, giving

$$\vec{J} = q \, \delta^3(\vec{r} - \vec{r}_0(t)) \, \frac{d\vec{r}_0}{dt} \,, \tag{4.32}$$

and so the 4-current is

$$J^\mu = (\rho, \rho\vec{v}) \,, \qquad \text{where} \qquad \vec{v} = \frac{d\vec{r}_0}{dt} \,, \tag{4.33}$$

and ρ is given by (4.31). We can verify that this is the correct current vector, by checking that it properly satisfies the charge-conservation equation $\partial_\mu J^\mu = \frac{\partial\rho}{\partial t} + \partial_i J^i = 0$. Thus we have

$$
\begin{aligned}
\frac{\partial\rho}{\partial t} &= q\frac{\partial}{\partial t}\delta^3(\vec{r} - \vec{r}_0(t)) = q\frac{\partial}{\partial x_0^i}\delta^3(\vec{r} - \vec{r}_0(t))\frac{dx_0^i}{dt} \,, \\
&= -q\frac{\partial}{\partial x^i}\delta^3(\vec{r} - \vec{r}_0(t))\frac{dx_0^i}{dt} = -\partial_i\left(\rho\,\frac{dx_0^i}{dt}\right) \,, \\
&= -\partial_i(\rho v^i) = -\partial_i J_i \,.
\end{aligned}
\tag{4.34}
$$

Note that we used the chain rule for differentiation in the first line, and that in getting to the second line we used the result that $\frac{\partial f(x-y)}{\partial x} = f'(x-y) = -\frac{\partial f(x-y)}{\partial y}$ for any function f with argument $(x-y)$ (where f' denotes the derivative of f with respect to its argument). It is also useful to note that we can write (4.33) as

$$J^\mu = \rho \frac{dx_0^\mu}{dt}, \qquad (4.35)$$

where we simply define $x_0^\mu(t)$ with $\mu = 0$ to be t.

Note that the integral $\int J^\mu A_\mu$ for the point charge gives a contribution to the action that is precisely of the form we saw in equation (2.82):

$$\begin{aligned}
\int J^\mu A_\mu d^4x &=& \int q\delta^3(\vec{r} - \vec{r}_0)\frac{dx_0^\mu}{dt}A_\mu d^3x\, dt\,, \\
&=& \int_{\text{path}} q\frac{dx_0^\mu}{dt}A_\mu(x_0^\nu)\, dt = q\int_{\text{path}} A_\mu dx^\mu\,. \qquad (4.36)
\end{aligned}$$

Suppose now we have N charges q_a, following paths $\vec{r}_a(t)$. Then the total charge density will be given by

$$\rho = \sum_{a=1}^{N} q_a\, \delta^3(\vec{r} - \vec{r}_a(t))\,. \qquad (4.37)$$

Since we have alluded several times to the fact that $\partial_\mu J^\mu = 0$ is the equation of charge conservation, it is appropriate to examine this in a little more detail. The total charge Q at time t_1 is given by integrating the charge density over the spatial 3-volume:

$$Q(t_1) = \int_{t=t_1} \rho\, d^3x = \int_{t=t_1} J^0 d\Sigma_0\,, \qquad \text{where} \qquad d\Sigma_0 = dx dy dz\,. \qquad (4.38)$$

This can be written covariantly as

$$Q(t_1) = \int_{t=t_1} J^\mu d\Sigma_\mu\,, \qquad (4.39)$$

where we define also[15]

$$d\Sigma_1 = -dt dy dz\,, \qquad d\Sigma_2 = -dt dz dx\,, \qquad d\Sigma_3 = -dt dy dz\,. \qquad (4.40)$$

Because the integral in (4.38) is *defined* to be over the 3-surface at constant t, it follows that the extra terms, for $\mu = 1, 2, 3$, in (4.39) do not contribute.

If we now calculate the charge at a later time t_2, and then take the difference between the two charges, we will obtain

$$Q(t_2) - Q(t_1) = \int_\Sigma J^\mu d\Sigma_\mu\,, \qquad (4.41)$$

---

[15]As has been mentioned previously, a careful derivation of the 4-dimensional divergence theorem $\int_V \partial_\mu W^\mu\, d^4x = \int_\Sigma W_\mu\, d\Sigma_\mu$ shows that the components of the 3-area element $d\Sigma_\mu$ should be as stated in eqns (4.38) and (4.40).

where Σ is the boundary of the 4-cylinder V, and it consists of the two "end caps" formed by the surfaces $t = t_1$ and $t = t_2$, and by the sides at spatial infinity. We are assuming the charges are confined to a finite region, and so the current J^μ is zero on the sides of the cylinder, as the radius is sent to infinity.

By the 4-dimensional analogue of the divergence theorem we shall have

$$\int_\Sigma J^\mu d\Sigma_\mu = \int_V \partial_\mu J^\mu \, d^4x \,, \tag{4.42}$$

where V is the 4-volume bounded by Σ. Thus we have

$$Q(t_2) - Q(t_1) = \int_V \partial_\mu J^\mu d^4x = 0 \,, \tag{4.43}$$

since $\partial_\mu J^\mu = 0$. Thus we see that $\partial_\mu J^\mu = 0$ implies that the total charge in an isolated finite region is independent of time.

Note that the equation of charge conservation implies the gauge invariance of the action. We have

$$S = \int \Big(-\frac{1}{16\pi} F_{\mu\nu} F^{\mu\nu} + J^\mu A_\mu \Big) d^4x \,, \tag{4.44}$$

and so under a gauge transformation $A_\mu \to A_\mu + \partial_\mu \lambda$, we find

$$
\begin{aligned}
S \;\longrightarrow\; & \int \Big(-\frac{1}{16\pi} F_{\mu\nu} F^{\mu\nu} + J^\mu A_\mu \Big) d^4x + \int J^\mu \partial_\mu \lambda d^4x \,, \\
= \; & S + \int J^\mu \partial_\mu \lambda d^4x = S + \int \partial_\mu (\lambda\, J^\mu) \, d^4x - \int \lambda \partial_\mu J^\mu d^4x \,, \\
= \; & S + \int_\Sigma \lambda\, J^\mu \, d\Sigma_\mu \,.
\end{aligned}
\tag{4.45}
$$

As usual, Σ here is the 3-cylinder of infinite radius in the spatial directions, with endcaps at $t = t_1$ and $t = t_2$. The current J^μ will vanish on the sides of the cylinder, since they are at spatial infinity and we take J^μ to vanish there. If we restrict attention to gauge transformations that vanish at $t = t_1$ and $t = t_2$ then the surface integral will therefore give zero, and so S is unchanged. Even if λ is non-zero at $t = t_1$ and $t = t_2$ then the surface integral will just give a constant, independent of A_μ, and so the original and the gauge transformed actions will give the same equations of motion.

4.4 Energy density and energy flux

Here, we review the calculation of energy density and energy flux in the 3-dimensional language. After that, we shall give the more elegant 4-dimensional description.

Consider the two Maxwell equations

$$\vec\nabla \times \vec B - \frac{\partial \vec E}{\partial t} = 4\pi \vec J, \qquad \vec\nabla \times \vec E + \frac{\partial \vec B}{\partial t} = 0 \,. \tag{4.46}$$

From these, we can deduce

$$
\begin{aligned}
\vec{E} \cdot \frac{\partial \vec{E}}{\partial t} + \vec{B} \cdot \frac{\partial \vec{B}}{\partial t} &= \vec{E} \cdot (\vec{\nabla} \times \vec{B} - 4\pi \vec{J}) - \vec{B} \cdot (\vec{\nabla} \times \vec{E}), \\
&= \epsilon_{ijk}(E_i \partial_j B_k - B_i \partial_j E_k) - 4\pi \vec{J} \cdot \vec{E}, \\
&= -\epsilon_{ijk}(B_i \partial_j E_k + E_k \partial_j B_i) - 4\pi \vec{J} \cdot \vec{E}, \\
&= -\partial_j (\epsilon_{jki} E_k B_i) - 4\pi \vec{J} \cdot \vec{E}, \\
&= -\vec{\nabla} \cdot (\vec{E} \times \vec{B}) - 4\pi \vec{J} \cdot \vec{E}.
\end{aligned}
\tag{4.47}
$$

We then define the *Poynting vector*

$$
\vec{S} \equiv \frac{1}{4\pi} \vec{E} \times \vec{B},
\tag{4.48}
$$

and so

$$
\tfrac{1}{2} \frac{\partial}{\partial t}(\vec{E}^2 + \vec{B}^2) = -4\pi \vec{\nabla} \cdot \vec{S} - 4\pi \vec{J} \cdot \vec{E},
\tag{4.49}
$$

since $\vec{E} \cdot \partial \vec{E}/\partial t = \tfrac{1}{2}\partial/\partial t(\vec{E}^2)$, etc.

We now assume that the $\vec{E}$ and $\vec{B}$ fields are confined to some finite region of space. Integrating (4.49) over all space, we obtain

$$
\begin{aligned}
\int \vec{J} \cdot \vec{E} d^3 x + \frac{1}{8\pi} \frac{d}{dt} \int (\vec{E}^2 + \vec{B}^2) d^3 x &= -\int \vec{\nabla} \cdot \vec{S} d^3 x, \\
&= -\int_{\Sigma} \vec{S} \cdot d\vec{\Sigma}, \\
&= 0.
\end{aligned}
\tag{4.50}
$$

We get zero on the right-hand side because, having used the divergence theorem to convert it to an integral over Σ, the "sphere at infinity," the integral vanishes since $\vec{E}$ and $\vec{B}$, and hence $\vec{S}$, are assumed to vanish there.

If the current $\vec{J}$ is assumed to be due to the motion of a set of charges q_a with 3-velocities $\vec{v}_a$ and rest masses m_a, we shall have from (4.32) and (3.12) that

$$
\int \vec{J} \cdot \vec{E} d^3 x = \sum_a q_a \vec{v}_a \cdot \vec{E}(\vec{r}_a) = \frac{d\mathcal{E}_{\text{mech}}}{dt},
\tag{4.51}
$$

where

$$
\mathcal{E}_{\text{mech}} = \sum_a m_a \gamma_a
\tag{4.52}
$$

is the total mechanical energy for the set of particles, as defined in (3.7). Note that here

$$
\gamma_a \equiv (1 - v_a^2)^{-1/2}.
\tag{4.53}
$$

Thus we conclude that

$$\frac{d}{dt}\Big(\mathcal{E}_{\text{mech}} + \frac{1}{8\pi}\int(\vec{E}^2 + \vec{B}^2)d^3x\Big) = 0\,. \tag{4.54}$$

This is the equation of total energy conservation. It says that the sum of the total mechanical energy plus the energy contained in the electromagnetic fields is a constant. Thus we interpret

$$W \equiv \frac{1}{8\pi}(\vec{E}^2 + \vec{B}^2) \tag{4.55}$$

as the energy density of the electromagnetic field.

Returning now to equation (4.49), we can consider integrating it over just a finite volume V, bounded by a closed 2-surface Σ. We will have

$$\frac{d}{dt}\Big(\mathcal{E}_{\text{mech}} + \int_V W d^3x\Big) = -\int_\Sigma \vec{S}\cdot d\vec{\Sigma}\,. \tag{4.56}$$

We now know that the left-hand side should be interpreted as the rate of change of total energy in the volume V and so clearly, since the total energy must be conserved, we should interpret the right-hand side as the flux of energy passing through the boundary surface Σ. Thus we see that the Poynting vector

$$\vec{S} = \frac{1}{4\pi}\vec{E}\times\vec{B} \tag{4.57}$$

is to be interpreted as the energy flux across the boundary; i.e. the energy per unit area per unit time.

4.5 Energy-momentum tensor

The discussion above was presented within the 3-dimensional framework. In this section we shall give a 4-dimensional spacetime description, which involves the introduction of the energy-momentum tensor. We shall begin with a rather general introduction. In order to simplify this discussion, we shall first describe the construction of the energy-momentum tensor for a scalar field $\phi(x^\mu)$. When we then apply these ideas to electromagnetism, we shall need to make the rather simple generalisation to the case of a Lagrangian for the vector field $A_\mu(x^\nu)$.

Recall that if we write the Maxwell tensor $F_{\mu\nu}$ in terms of the 4-vector potential A_μ, namely $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$, then the Bianchi identity $\partial_\mu F_{\nu\rho} + \partial_\nu F_{\rho\mu} + \partial_\rho F_{\mu\nu} = 0$ is automatically solved, and so the remaining content of the source-free Maxwell equations is just the field equation $\partial^\mu F_{\mu\nu} = 0$, which implies

$$\Box A_\mu - \partial_\mu(\partial_\nu A^\nu) = 0\,, \tag{4.58}$$

66

where $\Box = \partial^\nu \partial_\nu$ is the d'Alembertian. If we choose to work in the Lorenz gauge, $\partial_\nu A^\nu = 0$, the field equation reduces to

$$\Box A_\mu = 0 \,. \tag{4.59}$$

In the analogous, but simpler, example of a scalar field theory, we could consider the field equation

$$\Box \phi = 0 \,. \tag{4.60}$$

A slightly more general possibility would be to add a "mass term" for the scalar field, and consider the equation of motion

$$\Box \phi - m^2 \phi = 0 \,, \tag{4.61}$$

where m is a constant, describing the mass of the field. (As we shall discuss in detail later in the course, electromagnetism is described by a *massless* field. At the level of a particle description, this corresponds to the fact that the photon is a massless particle.)

The equation of motion (4.61) for the scalar field can be derived from an action. Consider the Lagrangian density

$$\mathcal{L} = -\tfrac{1}{2}(\partial^\mu \phi)(\partial_\mu \phi) - \tfrac{1}{2} m^2 \, \phi^2 \,. \tag{4.62}$$

Varying the action $S = \int \mathcal{L} d^4 x$ with respect to ϕ, we obtain

$$\begin{aligned}
\delta S &= \int \left(- (\partial^\mu \phi) \, \partial_\mu \delta\phi - m^2 \, \phi \, \delta\phi \right) d^4 x \,, \\
&= \int \left(\partial_\mu \partial^\mu \phi - m^2 \, \phi \right) \delta\phi \, d^4 x \,,
\end{aligned} \tag{4.63}$$

where we have, as usual, dropped the boundary term when performing the integration by parts to obtain the second line. Requiring $\delta S = 0$ for all possible $\delta\phi$ consistent with the boundary conditions, we conclude that the quantity in the parentheses on the second line must vanish, and hence we arrive at the equation of motion (4.61).

We can now extend the discussion by considering an abstract Lagrangian density $\mathcal{L}$ describing a scalar field ϕ. We shall assume that $\mathcal{L}$ depends on ϕ, and on its first derivatives $\partial_\nu \phi$, but that it has no *explicit* dependence[16] on the spacetime coordinates x^μ:

$$\mathcal{L} = \mathcal{L}(\phi, \partial_\nu \phi) \,. \tag{4.64}$$

The action is then given by

$$S = \int \mathcal{L}(\phi, \partial_\nu \phi) \, d^4 x \,. \tag{4.65}$$

---

[16]This is the analogue of a Lagrangian in classical mechanics that depends on the coordinates q_i and velocities $\dot{q}^i$, but which does not have explicit time dependence. Energy is conserved in a system described by such a Lagrangian.

The Euler-Lagrange equations for the scalar field then follow from requiring that the action be stationary. Thus we have[17]

$$
\begin{aligned}
\delta S &= \int \Big[\frac{\partial \mathcal{L}}{\partial \phi} \, \delta\phi + \frac{\partial \mathcal{L}}{\partial(\partial_\nu \phi)} \, \partial_\nu \delta\phi \Big] d^4 x \,, \\
&= \int \Big[\frac{\partial \mathcal{L}}{\partial \phi} \, \delta\phi - \partial_\nu \Big(\frac{\partial \mathcal{L}}{\partial(\partial_\nu \phi)} \Big) \delta\phi \Big] d^4 x + \int_\Sigma \frac{\partial \mathcal{L}}{\partial(\partial_\nu \phi)} \delta\phi \, d\Sigma_\nu \,, \\
&= \int \Big[\frac{\partial \mathcal{L}}{\partial \phi} \, \delta\phi - \partial_\nu \Big(\frac{\partial \mathcal{L}}{\partial(\partial_\nu \phi)} \Big) \delta\phi \Big] d^4 x \,,
\end{aligned}
\tag{4.67}
$$

where, in getting to the last line, we have as usual dropped the surface term integrated over the boundary cylinder Σ, since we shall insist that $\delta\phi$ vanishes on the endcaps of Σ at $t = t_1$ and $t = t_2$, and that ϕ goes to zero sufficiently fast at spatial infinity. Thus the requirement that $\delta S = 0$ for all such $\delta\phi$ implies the Euler-Lagrange equations[18]

$$
\frac{\partial \mathcal{L}}{\partial \phi} - \partial_\nu \Big(\frac{\partial \mathcal{L}}{\partial(\partial_\nu \phi)} \Big) = 0 \,.
\tag{4.68}
$$

Now consider the expression $\partial_\rho \mathcal{L} = \partial\mathcal{L}/\partial x^\rho$. Since we are assuming $\mathcal{L}$ has no *explicit* dependence on the spacetime coordinates, it follows that $\partial_\rho \mathcal{L}$ is given by the chain rule,

$$
\partial_\rho \mathcal{L} = \frac{\partial \mathcal{L}}{\partial \phi} \partial_\rho \phi + \frac{\partial \mathcal{L}}{\partial(\partial_\nu \phi)} \partial_\rho \partial_\nu \phi \,.
\tag{4.69}
$$

Now, using the Euler-Lagrange equations (4.68), we can write this as

$$
\begin{aligned}
\partial_\rho \mathcal{L} &= \partial_\nu \Big(\frac{\partial \mathcal{L}}{\partial(\partial_\nu \phi)} \Big) \partial_\rho \phi + \frac{\partial \mathcal{L}}{\partial(\partial_\nu \phi)} \partial_\nu \partial_\rho \phi \,, \\
&= \partial_\nu \Big[\frac{\partial \mathcal{L}}{\partial(\partial_\nu \phi)} \, \partial_\rho \phi \Big] \,,
\end{aligned}
\tag{4.70}
$$

and thus we have

$$
\partial_\nu \Big[\frac{\partial \mathcal{L}}{\partial(\partial_\nu \phi)} \, \partial_\rho \phi - \delta_\rho^\nu \, \mathcal{L} \Big] = 0 \,.
\tag{4.71}
$$

We are therefore led to define the 2-index tensor

$$
T_\rho{}^\nu \equiv -\frac{\partial \mathcal{L}}{\partial(\partial_\nu \phi)} \, \partial_\rho \phi + \delta_\rho^\nu \, \mathcal{L} \,,
\tag{4.72}
$$

---

[17]Note that $\partial\mathcal{L}/\partial(\partial_\nu \phi)$ means taking the partial derivative of $\mathcal{L}$ viewed as a function of ϕ and $\partial_\mu \phi$, with respect to $\partial_\nu \phi$. For example, if $\mathcal{L} = -\frac{1}{2}(\partial_\mu \phi)(\partial^\mu \phi) + \frac{1}{2}m^2 \phi^2$, then

$$
\partial\mathcal{L}/\partial(\partial_\nu \phi) = -(\partial^\mu \phi) \frac{\partial(\partial_\mu \phi)}{\partial(\partial_\nu \phi)} = -(\partial^\mu \phi) \, \delta_\mu^\nu = -\partial^\nu \phi \,.
\tag{4.66}
$$

Of course, in this example $\partial\mathcal{L}/\partial\phi$ is just equal to $-m^2 \phi$.

[18]These are the analogue of the Euler-Lagrange equations $\partial L/\partial q_i - d/dt(\partial L/\partial \dot{q}_i) = 0$ in particle mechanics for a system of particles with coordinates q_i and velocities $\dot{q}_i$, derived from the Lagrangian $L = L(q_i, \dot{q}_i)$ by requiring stationarity of the action $S = \int L dt$.

which then satisfies

$$\partial_\nu T_\rho{}^\nu = 0 \,. \tag{4.73}$$

$T_\rho{}^\nu$ is called an *energy-momentum tensor*.

In the specific example of the Lagrangian density (4.62) for a free massive scalar field, we see that the energy-momentum tensor will be given by

$$T_\rho{}^\nu = \partial^\nu \phi \, \partial_\rho \phi - \tfrac{1}{2} \delta_\rho^\nu \, (\partial^\sigma \phi)(\partial_\sigma \phi) - \tfrac{1}{2} m^2 \, \phi^2 \, \delta_\rho^\nu \,. \tag{4.74}$$

Raising the lower index and relabelling indices, we therefore have

$$T^{\mu\nu} = \partial^\mu \phi \, \partial^\nu \phi - \tfrac{1}{2} \eta^{\mu\nu} \, (\partial^\sigma \phi) \, (\partial_\sigma \phi) - \tfrac{1}{2} m^2 \, \eta^{\mu\nu} \, \phi^2 \,. \tag{4.75}$$

It so happens in this example that $T_{\mu\nu}$ has turned out to be symmetric in the indices μ and ν, but for a more general Lagrangian density this may not necessarily happen. We shall discuss this further below.

We saw previously that the equation $\partial_\mu J^\mu = 0$ for the 4-vector current density J^μ implies that there is a conserved charge

$$Q = \int_{t=\text{const}} \rho \, d^3x = \int_{t=\text{const}} J^0 \, d^3x \,. \tag{4.76}$$

The conservation of this charge, i.e. $\frac{dQ}{dt} = 0$ follows from

$$
\begin{aligned}
\frac{dQ}{dt} &= \int_{t=\text{const}} \partial_0 \, J^0 \, d^3x \\
&= -\int_{t=\text{const}} \partial_i \, J^i \, d^3x \\
&= -\int_S J^i \, dS_i \,, \\
&= 0 \,.
\end{aligned}
\tag{4.77}
$$

(The first step follows from $0 = \partial_\mu J^\mu = \partial_0 J^0 + \partial_i J^i$; the second step from the 3-dimensional divergence theorem; and the final step from the assumption that J^i vanishes on S, the sphere at spatial infinity.)

By an identical argument, it follows that the equation $\partial_\nu T_\rho{}^\nu = 0$ implies that there is a conserved 4-vector:

$$P^\mu \equiv \int_{t=\text{const}} T^{\mu 0} d\Sigma_0 = \int_{t=\text{const}} T^{\mu\nu} d\Sigma_\nu \,. \tag{4.78}$$

(Of course $T^{\mu\nu} = \eta^{\mu\rho} \, T_\rho{}^\nu$.) Thus we may check

$$
\begin{aligned}
\frac{dP^\mu}{dt} &= \partial_0 \int_{t=\text{const}} T^{\mu 0} d^3x = \int_{t=\text{const}} \partial_0 T^{\mu 0} d^3x = -\int_{t=\text{const}} \partial_i T^{\mu i} d^3x \,, \\
&= -\int_S T^{\mu i} dS_i = 0 \,,
\end{aligned}
\tag{4.79}
$$

where in the last line we have used the divergence theorem to turn the integral into a 2-dimensional integral over the boundary sphere S at infinity. This vanishes since we shall assume the fields are zero at infinity.

Notice that $T^{00} = -T_0{}^0$ and from (4.72) we therefore have

$$T^{00} = \frac{\partial \mathcal{L}}{\partial \partial_0 \phi} \, \partial_0 \phi - \mathcal{L} \, . \tag{4.80}$$

Now for a Lagrangian $L = L(q^i, \dot{q}^i)$ we have the canonical momentum $\pi_i = \partial L/\partial \dot{q}^i$, and the Hamiltonian

$$H = \pi_i \dot{q}^i - L \, . \tag{4.81}$$

Since there is no explicit time dependence, H is conserved, and is equal to the total energy of the system. Comparing with (4.80), we can recognise that T^{00} is the energy density. From (4.78) we therefore have that

$$P^0 = \int T^{00} d^3 x \tag{4.82}$$

is the total energy. Since it is manifest from its construction that P^μ is a 4-vector, and since its 0 component is the energy, it follows that P^μ is the 4-momentum.

The essential point in the discussion above is that P^μ given in (4.78) should be conserved, which requires $\partial_\nu T_\rho{}^\nu = 0$. The quantity $T_\rho{}^\nu$ we constructed is not the unique tensor with this property. We can define a new one, according to

$$T_\rho{}^\nu \longrightarrow T_\rho{}^\nu + \partial_\sigma \psi_\rho{}^{\nu\sigma} \, , \tag{4.83}$$

where $\psi_\rho{}^{\nu\sigma}$ is an arbitrary tensor that is antisymmetric in its last two indices,

$$\psi_\rho{}^{\nu\sigma} = -\psi_\rho{}^{\sigma\nu} \, . \tag{4.84}$$

We shall take $\psi_\rho{}^{\nu\sigma}$ to vanish at spatial infinity.

The antisymmetry implies, since partial derivatives commute, that

$$\partial_\nu \partial_\sigma \psi_\rho{}^{\nu\sigma} = 0 \, , \tag{4.85}$$

and hence that the modified energy-momentum tensor defined by (4.83) is conserved too. Furthermore, the modification to $T_\rho{}^\nu$ does not alter P^μ, since, from (4.78), the extra term will be

$$
\begin{aligned}
P^\mu_{\text{extra}} &= \int_{t=\text{const}} \partial_\sigma \psi^{\mu\nu\sigma} d\Sigma_\nu = \int_{t=\text{const}} \partial_\sigma \psi^{\mu 0\sigma} d\Sigma_0 \, , \\
&= \int_{t=\text{const}} \partial_0 \psi^{\mu 00} d^3 x + \int_{t=\text{const}} \partial_i \psi^{\mu 0i} d^3 x = \int_{t=\text{const}} \partial_i \psi^{\mu 0i} d^3 x \, , \\
&= \int_S \psi^{\mu 0i} dS_i = 0 \, ,
\end{aligned}
\tag{4.86}
$$

where S is the sphere at spatial infinity. (The fact that the integral is over a $t = $ constant surface means that only the $d\Sigma_0$ term contributes on the first line. The antisymmetry of $\psi^{\mu\nu\sigma}$ in ν and σ implies that $\psi^{\mu 00}$ on the second line is zero.) The modification to P^μ therefore vanishes since we are requiring that $\psi_\rho{}^{\nu\sigma}$ vanishes at spatial infinity.

The energy-momentum tensor can be pinned down uniquely by requiring that the four-dimensional angular momentum $M^{\mu\nu}$, defined by

$$M^{\mu\nu} = \int (x^\mu dP^\nu - x^\nu dP^\mu) \tag{4.87}$$

be conserved, where $dP^\mu = T^{\mu\rho} d\Sigma_\rho$, i.e. it is the integrand of the 4-momentum integral (4.78). First, let us make a remark about angular momentum in four dimensions. In three dimensions, we define the angular momentum 3-vector as $\vec{L} = \vec{r} \times \vec{p}$. In other words,

$$L_i = \epsilon_{ijk} x^j p^k = \tfrac{1}{2}\epsilon_{ijk}(x^j p^k - x^k p^j) = \tfrac{1}{2}\epsilon_{ijk} M^{jk}, \tag{4.88}$$

where $M^{jk} \equiv x^j p^k - x^k p^j$. Thus taking $M^{\mu\nu} = x^\mu p^\nu - x^\nu p^\mu$ in four dimensions is a plausible-looking generalisation. It should be noted that in a general dimension, angular momentum is described by a 2-index antisymmetric tensor; in other words, angular momentum is associated with a rotation in a 2-dimensional plane. It is a very special feature of three dimensions that we can use the ϵ_{ijk} tensor to map the 2-index antisymmetric tensor M^{jk} into the vector $L_i = \tfrac{1}{2}\epsilon_{ijk} M^{jk}$. Put another way, a very special feature of three dimensions is that a rotation in the (x, y) plane can equivalently be described as a rotation about the orthogonal (i.e. z) axis. In higher dimensions, rotations do not occur around axes, but rather, in 2-planes. It is amusing, therefore, to try to imagine what the analogue of an axle is for a higher-dimensional car!

Getting back to our discussion of angular momentum and the energy-momentum tensor in four dimensions, we are defining

$$M^{\mu\nu} = \int (x^\mu dP^\nu - x^\nu dP^\mu) = \int (x^\mu T^{\nu\rho} - x^\nu T^{\mu\rho}) d\Sigma_\rho, . \tag{4.89}$$

By analogous arguments to those we used earlier, this will be conserved (i.e. $dM^{\mu\nu}/dt = 0$) if

$$\partial_\rho (x^\mu T^{\nu\rho} - x^\nu T^{\mu\rho}) = 0. \tag{4.90}$$

Distributing the derivative, we therefore have the requirement that

$$\delta_\rho^\mu T^{\nu\rho} + x^\mu \partial_\rho T^{\nu\rho} - \delta_\rho^\nu T^{\mu\rho} - x^\nu \partial_\rho T^{\mu\rho} = 0, \tag{4.91}$$

and hence, since $\partial_\rho T^{\mu\rho} = 0$, that $T^{\mu\nu}$ is symmetric,

$$T^{\mu\nu} = T^{\nu\mu}. \tag{4.92}$$

Using the freedom to add $\partial_\sigma \psi^{\mu\nu\sigma}$ to $T^{\mu\nu}$, as we discussed earlier, it is always possible to arrange for $T^{\mu\nu}$ to be symmetric. From now on, we shall assume that this is done.

We already saw that $P^\mu = \int T^{\mu 0} d^3 x$ is the 4-momentum, so T^{00} is the energy density, and T^{i0} is the 3-momentum density. Let us now look at the conservation equation $\partial_\nu T^{\mu\nu} = 0$ in more detail. Taking $\mu = 0$, we have $\partial_\nu T^{0\nu} = 0$, or

$$\frac{\partial}{\partial t} T^{00} + \partial_j T^{0j} = 0 \,. \tag{4.93}$$

integrating over a spatial 3-volume V with boundary S, we therefore find

$$\frac{d}{dt} \int_V T^{00} d^3 x = - \int_V \partial_j T^{0j} d^3 x = - \int_S T^{0j} dS_j \,. \tag{4.94}$$

The left-hand side is the rate of change of field energy in the volume V, and so we can deduce, from energy conservation, that T^{0j} is the energy flux 3-vector. But since we are now working with a *symmetric* energy-momentum tensor, we have that $T^{0j} = T^{j0}$, and we already identified T^{j0} as the 3-momentum density. Thus we have that

$$\text{energy flux} = \text{momentum density} \,. \tag{4.95}$$

From the $\mu = i$ components of $\partial_\nu T^{\mu\nu} = 0$, we have

$$\frac{\partial}{\partial t} T^{i0} + \partial_j T^{ij} = 0 \,, \tag{4.96}$$

and so, integrating over the 3-volume V, we get

$$\frac{d}{dt} \int_V T^{i0} d^3 x = - \int_V \partial_j T^{ij} d^3 x = - \int_S T^{ij} dS_j \,. \tag{4.97}$$

The left-hand side is the rate of change of 3-momentum, and so we deduce that T^{ij} is the 3-tensor of momentum flux density. It gives the i component of 3-momentum that flows, per unit time, through the 2-surface perpendicular to the x^j axis. T^{ij} is sometimes called the 3-dimensional stress tensor.

4.6 Energy-momentum tensor for the electromagnetic field

Recall that for a scalar field ϕ, the original construction of the energy-momentum tensor $T_\rho{}^\nu$ (which we later modified by adding $\partial_\sigma \psi_\rho{}^{\nu\sigma}$ where $\psi_\rho{}^{\nu\sigma} = -\psi_\rho{}^{\sigma\nu}$) was given by

$$T_\rho{}^\nu = -\frac{\partial \mathcal{L}}{\partial(\partial_\nu \phi)} \partial_\rho \phi + \delta_\rho^\nu \mathcal{L} \,. \tag{4.98}$$

If we have a set of N scalar fields ϕ_a, then it is easy to see that the analogous conserved tensor is

$$T_\rho{}^\nu = -\sum_{a=1}^N \frac{\partial \mathcal{L}}{\partial(\partial_\nu \phi_a)} \partial_\rho \phi_a + \delta_\rho^\nu \mathcal{L} \,. \tag{4.99}$$

A similar calculation shows that if we consider instead a vector field A_σ, with Lagrangian density $\mathcal{L}(A_\sigma, \partial_\nu A_\sigma)$, the construction will give a conserved energy-momentum tensor

$$T_\rho{}^\nu = -\frac{\partial \mathcal{L}}{\partial(\partial_\nu A_\sigma)} \partial_\rho A_\sigma + \delta_\rho^\nu \mathcal{L}. \tag{4.100}$$

Let us apply this to the Lagrangian density for pure electrodynamics (without sources),

$$\mathcal{L} = -\frac{1}{16\pi} F_{\mu\nu} F^{\mu\nu}. \tag{4.101}$$

We have

$$\delta \mathcal{L} = -\frac{1}{8\pi} F^{\mu\nu} \delta F_{\mu\nu} = -\frac{1}{4\pi} F^{\mu\nu} \partial_\mu \delta A_\nu, \tag{4.102}$$

and so

$$\frac{\partial \mathcal{L}}{\partial(\partial_\mu A_\nu)} = -\frac{1}{4\pi} F^{\mu\nu}. \tag{4.103}$$

Thus from (4.100) we find

$$T_\rho{}^\nu = \frac{1}{4\pi} F^{\nu\sigma} \partial_\rho A_\sigma - \frac{1}{16\pi} \delta_\rho^\nu F_{\sigma\lambda} F^{\sigma\lambda}, \tag{4.104}$$

and so

$$T^{\mu\nu} = \frac{1}{4\pi} F^{\nu\sigma} \partial^\mu A_\sigma - \frac{1}{16\pi} \eta^{\mu\nu} F_{\sigma\lambda} F^{\sigma\lambda}. \tag{4.105}$$

This expression is not symmetric in μ and ν. However, following our previous discussion, we can add a term $\partial_\sigma \psi^{\mu\nu\sigma}$ to it, where $\psi^{\mu\nu\sigma} = -\psi^{\mu\sigma\nu}$, without upsetting the conservation condition $\partial_\nu T^{\mu\nu} = 0$. Specifically, we shall choose $\psi^{\mu\nu\sigma} = -1/(4\pi) A^\mu F^{\nu\sigma}$, and so

$$\begin{aligned}
\partial_\sigma \psi^{\mu\nu\sigma} &= -\frac{1}{4\pi} \partial_\sigma(A^\mu F^{\nu\sigma}), \\
&= -\frac{1}{4\pi}(\partial_\sigma A^\mu) F^{\nu\sigma} - \frac{1}{4\pi} A^\mu \partial_\sigma F^{\nu\sigma} = -\frac{1}{4\pi}(\partial_\sigma A^\mu) F^{\nu\sigma}.
\end{aligned} \tag{4.106}$$

(the $\partial_\sigma F^{\nu\sigma}$ term drops as a consequence of the source-free field equation.) This leads to the new energy-momentum tensor

$$T^{\mu\nu} = \frac{1}{4\pi} F^{\nu\sigma}(\partial^\mu A_\sigma - \partial_\sigma A^\mu) - \frac{1}{16\pi} \eta^{\mu\nu} F_{\sigma\lambda} F^{\sigma\lambda}, \tag{4.107}$$

or, in other words,

$$T^{\mu\nu} = \frac{1}{4\pi}\left(F^\mu{}_\sigma F^{\nu\sigma} - \tfrac{1}{4} \eta^{\mu\nu} F_{\sigma\lambda} F^{\sigma\lambda}\right). \tag{4.108}$$

This is indeed manifestly symmetric in μ and ν. From now on, it will be understood when we speak of the energy-momentum tensor for electrodynamics that this is the one we mean.

It is a straightforward exercise to verify directly, using the source-free Maxwell field equation and the Bianchi identity, that indeed $T^{\mu\nu}$ given by (4.108) is conserved, $\partial_\nu T^{\mu\nu} = 0$. Note that it has another simple property, namely that it is trace-free, in the sense that

$$\eta_{\mu\nu} T^{\mu\nu} = 0. \tag{4.109}$$

73

This is easily seen from (4.108), as a consequence of the fact that $\eta^{\mu\nu}\eta_{\mu\nu} = 4$ in four dimensions. The trace-free property is related to a special feature of the Maxwell equations in four dimensions, known as *conformal invariance*.

Having obtained the energy-momentum tensor (4.108) for the electromagnetic field, it is instructive to look at its components from the three-dimensional point of view. First, recall that we showed earlier that

$$F_{\sigma\lambda}F^{\sigma\lambda} = 2(\vec{B}^2 - \vec{E}^2)\,. \tag{4.110}$$

Then, we find

$$
\begin{aligned}
T^{00} &= \frac{1}{4\pi}(F^0{}_\sigma F^{0\sigma} - \tfrac{1}{4}\eta^{00}F_{\sigma\lambda}F^{\sigma\lambda})\,, \\
&= \frac{1}{4\pi}(F^{0i}F^{0i} + \tfrac{1}{2}\vec{B}^2 - \tfrac{1}{2}\vec{E}^2)\,, \\
&= \frac{1}{4\pi}(\vec{E}^2 + \tfrac{1}{2}\vec{B}^2 - \tfrac{1}{2}\vec{E}^2)\,, \\
&= \frac{1}{8\pi}(\vec{E}^2 + \vec{B}^2)\,.
\end{aligned}
\tag{4.111}
$$

Thus T^{00} is equal to the energy density W that we introduced in (4.55).

Now consider T^{0i}. Since $\eta^{0i} = 0$, we have

$$
\begin{aligned}
T^{0i} &= \frac{1}{4\pi}F^0{}_\sigma F^{i\sigma} = \frac{1}{4\pi}F^0{}_j\,F^{ij}\,, \\
&= \frac{1}{4\pi}E_j\epsilon_{ijk}B_k = S_i\,,
\end{aligned}
\tag{4.112}
$$

where $\vec{S} = 1/(4\pi)\vec{E} \times \vec{B}$ is the Poynting vector introduced in (4.48). Thus T^{0i} is the energy flux. As we remarked earlier, since we now have $T^{0i} = T^{i0}$, it can be equivalently interpreted as the 3-momentum density vector.

Finally, we consider the components T^{ij}. We have

$$
\begin{aligned}
T^{ij} &= \frac{1}{4\pi}\Big(F^i{}_\sigma F^{j\sigma} - \tfrac{1}{4}\eta^{ij}2(\vec{B}^2 - \vec{E}^2)\Big)\,, \\
&= \frac{1}{4\pi}\Big(F^i{}_0 F^{j0} + F^i{}_k F^{jk} - \tfrac{1}{2}\delta_{ij}(\vec{B}^2 - \vec{E}^2)\Big)\,, \\
&= \frac{1}{4\pi}\Big(-E_i E_j + \epsilon_{ik\ell}\epsilon_{jkm}B_\ell B_m - \tfrac{1}{2}\delta_{ij}(\vec{B}^2 - \vec{E}^2)\Big)\,, \\
&= \frac{1}{4\pi}\Big(-E_i E_j + \delta_{ij}\vec{B}^2 - B_i B_j - \tfrac{1}{2}\delta_{ij}(\vec{B}^2 - \vec{E}^2)\Big)\,, \\
&= \frac{1}{4\pi}\Big(-E_i E_j - B_i B_j + \tfrac{1}{2}\delta_{ij}(\vec{E}^2 + \vec{B}^2)\Big)\,.
\end{aligned}
\tag{4.113}
$$

To summarise, we have

$$T^{\mu\nu} = \begin{pmatrix} T^{00} & T^{0j} \\ T^{i0} & \sigma_{ij} \end{pmatrix} = \begin{pmatrix} W & S_j \\ S_i & \sigma_{ij} \end{pmatrix}\,, \tag{4.114}$$

74

where W and $\vec{S}$ are the energy density and Poynting flux,

$$W = \frac{1}{8\pi} \left(\vec{E}^2 + \vec{B}^2 \right), \qquad \vec{S} = \frac{1}{4\pi} \vec{E} \times \vec{B}, \tag{4.115}$$

and

$$\sigma_{ij} = -\frac{1}{4\pi} \left(E_i E_j + B_i B_j \right) + W \delta_{ij}. \tag{4.116}$$

Canonical Forms for $T_{\mu\nu}$:

- Unless $\vec{E}$ and $\vec{B}$ are perpendicular and equal in magnitude, we can always choose a Lorentz frame where $\vec{E}$ and $\vec{B}$ are parallel *at a point*. (In the case that $\vec{E}$ and $\vec{B}$ are perpendicular (but unequal in magnitude), one or other of $\vec{E}$ or $\vec{B}$ will be zero, at the point, in the new Lorentz frame.)

  Let the direction of $\vec{E}$ and $\vec{B}$ then be along z:

$$\vec{E} = (0, 0, E), \qquad \vec{B} = (0, 0, B). \tag{4.117}$$

  Then we have $\vec{S} = 1/(4\pi) \vec{E} \times \vec{B} = 0$ and

$$\sigma_{11} = \sigma_{22} = W, \qquad \sigma_{33} = -W, \qquad \sigma_{ij} = 0 \quad \text{otherwise}, \tag{4.118}$$

  and so $T^{\mu\nu}$ is diagonal, given by

$$T^{\mu\nu} = \begin{pmatrix} W & 0 & 0 & 0 \\ 0 & W & 0 & 0 \\ 0 & 0 & W & 0 \\ 0 & 0 & 0 & -W \end{pmatrix}, \tag{4.119}$$

  with $W = 1/(8\pi)(E^2 + B^2)$.

- If $\vec{E}$ and $\vec{B}$ are perpendicular and $|\vec{E}| = |\vec{B}|$ at a point, then at that point we can choose axes so that

$$\vec{E} = (E, 0, 0), \qquad \vec{B} = (0, B, 0) = (0, E, 0). \tag{4.120}$$

  (We shall see later on that this $\vec{E} \cdot \vec{B} = 0$ and $|\vec{E}| = |\vec{B}|$ case arises for electromagnetic plane waves.) Then we have

$$\begin{aligned} W &= \frac{1}{4\pi} E^2, \qquad \vec{S} = (0, 0, W), \\ \sigma_{11} &= \sigma_{22} = 0, \qquad \sigma_{33} = W, \qquad \sigma_{ij} = 0 \quad \text{otherwise}, \end{aligned} \tag{4.121}$$

and therefore $T^{\mu\nu}$ is given by

$$T^{\mu\nu} = \begin{pmatrix} W & 0 & 0 & W \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ W & 0 & 0 & W \end{pmatrix}. \tag{4.122}$$

4.7   Inclusion of massive charged particles

We now consider the energy-momentum tensor for a particle with rest mass m. We proceed by analogy with the construction of the 4-current density J^μ for charged non-interacting particles. Thus we define first a mass density, ε, for a point mass m located at $\vec{r} = \vec{r}_0(t)$. This will simply be given by a 3-dimensional delta function, with strength m, located at the instantaneous position of the mass point:

$$\varepsilon = m\delta^3(\vec{r} - \vec{r}_0(t)). \tag{4.123}$$

The energy density for the particle will then be its mass density times the corresponding γ factor, i.e. $T^{00} = \varepsilon\,\gamma$, where $\gamma = (1 - v^2)^{-1/2}$ and $\vec{v} = \frac{d\vec{r}_0(t)}{dt}$ is the velocity of the particle. Since the coordinate time t and the proper time τ in the frame of the particle are related, as usual, by $dt = \gamma d\tau$, we then have

$$T^{00} = \varepsilon\,\frac{dt}{d\tau}. \tag{4.124}$$

The 3-momentum density will be v^i times the energy density, and so

$$T^{0i} = \varepsilon\gamma\frac{dx^i}{dt} = \varepsilon\frac{dt}{d\tau}\frac{dx^i}{dt}. \tag{4.125}$$

We can therefore write

$$T^{0\nu} = \varepsilon\frac{dt}{d\tau}\frac{dx^\nu}{dt} = \varepsilon\frac{dx^0}{d\tau}\frac{dx^\nu}{dt}. \tag{4.126}$$

On general grounds of Lorentz covariance, it must therefore be that

$$\begin{aligned} T^{\mu\nu} &= \varepsilon\frac{dx^\mu}{d\tau}\frac{dx^\nu}{dt}, \\ &= \varepsilon\frac{dx^\mu}{d\tau}\frac{dx^\nu}{d\tau}\frac{d\tau}{dt}, \\ &= \frac{\varepsilon}{\gamma}\frac{dx^\mu}{d\tau}\frac{dx^\nu}{d\tau} \end{aligned} \tag{4.127}$$

By writing it as we have done in the second line here, it becomes manifest that $T^{\mu\nu}$ for the particle is symmetric in μ and ν.

Consider now a system consisting of a particle with mass m and charge q, moving in an electromagnetic field. Clearly, since the particle interacts with the field, we should not expect either the energy-momentum tensor (4.108) for the electromagnetic field or the energy-momentum tensor (4.127) for the particle to be conserved separately. This is because energy, and momentum, is being exchanged between the particle and the field. We *can* expect, however, that the total energy-momentum tensor for the system, i.e. the sum of (4.108) and (4.127), to be conserved.

In order to distinguish clearly between the various energy-momentum tensors, let us define

$$T_{\text{tot.}}^{\mu\nu} = T_{\text{e.m.}}^{\mu\nu} + T_{\text{part.}}^{\mu\nu}, \tag{4.128}$$

where $T_{\text{e.m.}}^{\mu\nu}$ and $T_{\text{part.}}^{\mu\nu}$ are the energy-momentum tensors for the electromagnetic field and the particle respectively:

$$
\begin{aligned}
T_{\text{e.m.}}^{\mu\nu} &= \frac{1}{4\pi}\left(F^{\mu}{}_{\sigma}\,F^{\nu\sigma} - \tfrac{1}{4}\eta^{\mu\nu}\,F_{\sigma\lambda}F^{\sigma\lambda}\right), \\
T_{\text{part.}}^{\mu\nu} &= \varepsilon\frac{dx^{\mu}}{d\tau}\frac{dx^{\nu}}{dt},
\end{aligned}
\tag{4.129}
$$

where $\varepsilon = m\delta^3(\vec{r} - \vec{r}_0(t))$.

Consider $T_{\text{e.m.}}^{\mu\nu}$ first. Taking the divergence, we find

$$
\begin{aligned}
\partial_{\nu}T_{\text{e.m.}}^{\mu\nu} &= \frac{1}{4\pi}\left(\partial_{\nu}F^{\mu}{}_{\sigma}\,F^{\nu\sigma} + F^{\mu}{}_{\sigma}\,\partial_{\nu}F^{\nu\sigma} - \tfrac{1}{2}F^{\sigma\lambda}\partial^{\mu}F_{\sigma\lambda}\right), \\
&= \frac{1}{4\pi}\left(\partial_{\nu}F^{\mu}{}_{\sigma}\,F^{\nu\sigma} + F^{\mu}{}_{\sigma}\,\partial_{\nu}F^{\nu\sigma} + \tfrac{1}{2}F^{\sigma\lambda}\partial_{\sigma}F_{\lambda}{}^{\mu} + \tfrac{1}{2}F^{\sigma\lambda}\partial_{\lambda}F^{\mu}{}_{\sigma}\right), \\
&= \frac{1}{4\pi}\left(\partial_{\nu}F^{\mu}{}_{\sigma}\,F^{\nu\sigma} - \tfrac{1}{2}F^{\sigma\lambda}\partial_{\sigma}F^{\mu}{}_{\lambda} - \tfrac{1}{2}F^{\lambda\sigma}\partial_{\lambda}F^{\mu}{}_{\sigma} + F^{\mu}{}_{\sigma}\,\partial_{\nu}F^{\nu\sigma}\right), \\
&= \frac{1}{4\pi}F^{\mu}{}_{\sigma}\,\partial_{\nu}F^{\nu\sigma}, \\
&= -F^{\mu}{}_{\nu}\,J^{\nu}.
\end{aligned}
\tag{4.130}
$$

In getting to the second line we used the Bianchi identity on the last term in the top line. The third line is obtained by swapping indices on a field strength in the terms with the $\frac{1}{2}$ factors, and this reveals that all except one term cancel, leading to the result on the fourth line. Finally, the result on the fifth lines follows after using the Maxwell equation of motion $\partial_{\nu}F^{\nu\sigma} = -4\pi J^{\sigma}$. As expected, the energy-momentum tensor for the electromagnetic field by itself is not conserved when there are sources.

Now we want to show that this non-conservation is balanced by an equal and opposite non-conservation for the energy-momentum tensor of the particle, which is given in (4.129). We have

$$\partial_{\nu}T_{\text{part.}}^{\mu\nu} = \partial_{\nu}\left(\varepsilon\frac{dx^{\nu}}{dt}\right)\frac{dx^{\mu}}{d\tau} + \varepsilon\frac{dx^{\nu}}{dt}\partial_{\nu}\left(\frac{dx^{\mu}}{d\tau}\right). \tag{4.131}$$

The first term is zero. This can be seen from the fact that the calculation is identical to the one which we used a while back in section 4.3 to show that the 4-current $J^\mu = \rho\frac{dx^\mu}{dt}$ for a charged particle is conserved. Thus we have

$$
\begin{aligned}
\partial_\nu T^{\mu\nu}_{\text{part.}} &= \varepsilon\frac{dx^\nu}{dt}\partial_\nu\Big(\frac{dx^\mu}{d\tau}\Big) = \varepsilon\frac{dx^\nu}{dt}\partial_\nu U^\mu\,, \\
&= \varepsilon\frac{dU^\mu}{dt}\,.
\end{aligned}
\tag{4.132}
$$

By the Lorentz force equation $m\frac{dU^\mu}{d\tau} = qF^\mu{}_\nu U^\nu$, we have

$$
\varepsilon\frac{dU^\mu}{d\tau} = \rho F^\mu{}_\nu U^\nu = \rho F^\mu{}_\nu\frac{dx^\nu}{d\tau}\,,
\tag{4.133}
$$

and so

$$
\varepsilon\frac{dU^\mu}{dt} = \rho F^\mu{}_\nu\frac{dx^\nu}{dt} = F^\mu{}_\nu J^\nu\,,
\tag{4.134}
$$

since $J^\mu = \rho\frac{dx^\mu}{dt}$. Thus we conclude that

$$
\partial_\nu T^{\mu\nu}_{\text{part.}} = F^\mu{}_\nu J^\nu\,,
\tag{4.135}
$$

and so, combining this with (4.130), we conclude that the total energy-momentum tensor for the particle plus electromagnetic field, defined in (4.128) is conserved,

$$
\partial_\nu T^{\mu\nu}_{\text{tot.}} = 0\,.
\tag{4.136}
$$

5   Coulomb's Law

5.1   Potential of a point charge

Consider first a static point charge, for which the Maxwell equations therefore reduce to

$$
\vec{\nabla}\times\vec{E} = 0\,, \qquad \vec{\nabla}\cdot\vec{E} = 4\pi\rho\,.
\tag{5.1}
$$

The first equation implies, of course, that we can write

$$
\vec{E} = -\vec{\nabla}\phi\,,
\tag{5.2}
$$

and then the second equation implies that ϕ satisfies the Poisson equation

$$
\nabla^2\phi = -4\pi\rho\,.
\tag{5.3}
$$

If the point charge is located at the origin, and the charge is e, then the charge density ρ is given by

$$
\rho = e\,\delta^3(\vec{r})\,.
\tag{5.4}
$$

Away from the origin, (5.3) implies that ϕ should satisfy the Laplace equation,

$$\nabla^2\phi = 0\,, \qquad |\vec{r}| > 0\,. \tag{5.5}$$

Since the charge density (5.4) is spherically symmetric, we can assume that ϕ will be spherically symmetric too, $\phi(\vec{r}) = \phi(r)$, where $r = |\vec{r}|$. From $r^2 = x^j x^j$ we deduce, by acting with ∂_i, that

$$\partial_i r = \frac{x^i}{r}\,. \tag{5.6}$$

From this it follows by the chain rule that

$$\partial_i \phi = \phi' \partial_i r = \phi' \frac{x^i}{r}\,, \tag{5.7}$$

where $\phi' \equiv d\phi/dr$, and hence

$$\begin{aligned}
\nabla^2\phi &= \partial_i\partial_i\phi = \partial_i\left(\phi'\frac{x^i}{r}\right) = \phi''\frac{x^i}{r}\frac{x^i}{r} + \phi'\frac{\partial_i x^i}{r} + \phi' x^i \partial_i \frac{1}{r}\,, \\
&= \phi'' + \frac{2}{r}\phi'\,. \tag{5.8}
\end{aligned}$$

Thus the Laplace equation (5.5) can be written as

$$(r^2\phi')' = 0\,, \qquad r > 0\,, \tag{5.9}$$

which integrates to give

$$\phi = \frac{q}{r}\,, \tag{5.10}$$

where q is a constant, and we have dropped an additive constant of integration by using the gauge freedom to choose $\phi(\infty) = 0$.

To determine the constant q, we integrate the Poisson equation (5.3) over the interior V_R of a sphere of radius R centred on the origin, and use the divergence theorem:

$$\begin{aligned}
\int_{V_R} \nabla^2\phi\, d^3x &= -4\pi e \int_{V_R} \delta^3(\vec{r})d^3x = -4\pi e\,, \\
&= \int_{S_R} \vec{\nabla}\phi \cdot d\vec{S} = \int_{S_R} \partial_i\left(\frac{q}{r}\right)dS_i\,, \\
&= -q\int_{S_R} \frac{x^i dS_i}{r^3} = -q\int_{S_R} \frac{n^i dS_i}{R^2}\,, \tag{5.11}
\end{aligned}$$

where S_R is the surface of the sphere of radius R that bounds the volume V_R, and $n^i \equiv x^i/r$ is the outward-pointing unit vector. Clearly we have

$$n^i dS_i = R^2 d\Omega\,, \tag{5.12}$$

79

where $d\Omega$ is the area element on the unit-radius sphere, and so

$$-q \int_{S_R} \frac{n^i dS_i}{r^2} = -q \int d\Omega = -4\pi q \,, \tag{5.13}$$

and so we conclude that q is equal to e, the charge on the point charge at $r = 0$.

Note that if the point charge e were located at $\vec{r}'$, rather than at the origin, then by trivially translating the coordinate system we will have the potential

$$\phi(\vec{r}) = \frac{e}{|\vec{r} - \vec{r}'|} \,, \tag{5.14}$$

and this will satisfy

$$\nabla^2 \phi = -4\pi e \delta^3(\vec{r} - \vec{r}') \,. \tag{5.15}$$

5.2 Electrostatic energy

In general, the energy density of an electromagnetic field is given by $W = 1/(8\pi)(\vec{E}^2 + \vec{B}^2)$. A purely electrostatic system therefore has a field energy U given by

$$\begin{aligned}
U &= \int W d^3x = \frac{1}{8\pi} \int \vec{E}^2 d^3x \,, \\
&= -\frac{1}{8\pi} \int \vec{E} \cdot \vec{\nabla}\phi \, d^3x \,, \\
&= -\frac{1}{8\pi} \int \vec{\nabla} \cdot (\vec{E}\,\phi) d^3x + \frac{1}{8\pi} \int (\vec{\nabla} \cdot \vec{E})\phi \, d^3x \,, \\
&= -\frac{1}{8\pi} \int_S \phi\, \vec{E} \cdot d\vec{S} + \tfrac{1}{2} \int \rho\, \phi \, d^3x \,, \\
&= \tfrac{1}{2} \int \rho\, \phi \, d^3x \,. \tag{5.16}
\end{aligned}$$

Note that the surface integral over the sphere at infinity gives zero because the electric field is assumed to die away to zero there. Thus we conclude that the electrostatic field energy is given by

$$U = \tfrac{1}{2} \int \rho\, \phi \, d^3x \,. \tag{5.17}$$

We can apply this formula to a system of N charges q_a, located at points $\vec{r}_a$, for which we shall have

$$\rho = \sum_{a=1}^{N} q_a \delta^3(\vec{r} - \vec{r}_a) \,. \tag{5.18}$$

However, a naive application of (5.17) would give nonsense, since we find

$$U = \tfrac{1}{2} \sum_{a=1}^{N} q_a \int \delta^3(\vec{r} - \vec{r}_a)\, \phi(\vec{r}) d^3x = \tfrac{1}{2} \sum_{a=1}^{N} q_a \phi(\vec{r}_a) \,, \tag{5.19}$$

80

where $\phi(\vec{r})$ is given by (5.14),

$$\phi(\vec{r}) = \sum_{b=1}^{N} \frac{q_b}{|\vec{r} - \vec{r}_b|}\,, \tag{5.20}$$

This means that (5.19) will give infinity since $\phi(\vec{r})$, not unreasonably, diverges at the location of each point charge.

This is the classic "self-energy" problem, which one encounters even for a single point charge. There is no totally satisfactory way around this in classical electromagnetism, and so one has to adopt a "fudge." The fudge consists of observing that the true self-energy of a charge, whatever that might mean, is a constant. Naively, it appears to be an infinite constant, but that is clearly the result of making the idealised assumption that the charge is literally located at a single point. In any case, one can argue that the constant self-energy will not be observable, as far as energy-conservation considerations are concerned, and so one might as well just drop it for now. Thus the way to make sense of the ostensibly divergent energy (5.19) for the system of point charges is to replace $\phi(\vec{r}_a)$, which means the potential at $\vec{r} = \vec{r}_a$ due to *all* the charges, by ϕ_a, which is defined to be the potential at $\vec{r} = \vec{r}_a$ due to all the charges *except the charge q_a that is itself located at $\vec{r} = \vec{r}_a$*. Thus we have

$$\phi_a \equiv \sum_{b \neq a} \frac{q_b}{|\vec{r}_a - \vec{r}_b|}\,, \tag{5.21}$$

and so (5.19) is now interpreted to mean that the total energy of the system of charges is

$$U = \tfrac{1}{2} \sum_a \sum_{b \neq a} \frac{q_a q_b}{|\vec{r}_a - \vec{r}_b|}\,. \tag{5.22}$$

5.3   Field of a uniformly moving charge

Suppose a charge e is moving with uniform velocity $\vec{v}$ in the Lorentz frame S. We may transform to a frame S', moving with velocity $\vec{v}$ relative to S, in which the charge is at rest. For convenience, we shall choose the origin of axes so that the charge is located at the origin of the frame S'.

It follows that in the frame S', the field due to the charge can be described purely by the electric scalar potential ϕ':

$$\text{In the frame } S': \qquad \phi' = \frac{e}{r'}\,, \qquad \vec{A}' = 0\,. \tag{5.23}$$

(Note that the primes here all signify that the quantities are those of the primed frame S'.)

We know that $A^\mu = (\phi, \vec{A})$ is a 4-vector, and so the components A^μ transform under Lorentz boosts in exactly the same way as the components of x^μ. Thus we shall have

$$\phi' = \gamma\,(\phi - \vec{v}\cdot\vec{A})\,, \qquad \vec{A}' = \vec{A} + \frac{\gamma-1}{v^2}\,(\vec{v}\cdot\vec{A})\vec{v} - \gamma\vec{v}\,\phi\,, \tag{5.24}$$

where $\gamma = (1-v^2)^{-1/2}$. Clearly the inverse Lorentz transformation is obtained by switching the roles of the primed and unprimed fields and sending $\vec{v} \to -\vec{v}$, and so we shall have

$$\phi = \gamma\,(\phi' + \vec{v}\cdot\vec{A}')\,, \qquad \vec{A} = \vec{A}' + \frac{\gamma-1}{v^2}\,(\vec{v}\cdot\vec{A}')\vec{v} + \gamma\vec{v}\,\phi'\,. \tag{5.25}$$

From (5.23), we therefore find that the potentials in the frame S, in which the particle is moving with velocity $\vec{v}$, are given by

$$\phi = \gamma\phi' = \frac{e\gamma}{r'}\,, \qquad \vec{A} = \gamma\vec{v}\,\phi' = \frac{e\gamma\vec{v}}{r'}\,. \tag{5.26}$$

Note that we still have r' appearing in the denominator, which we would now like to express in terms of the unprimed coordinates.

Suppose, for example, that we orient the axes so that $\vec{v}$ lies along the x direction. Then we shall have

$$x' = \gamma(x - vt)\,, \qquad y' = y\,, \qquad z' = z\,, \tag{5.27}$$

and so

$$
\begin{aligned}
r'^2 &= x'^2 + y'^2 + z'^2 = \gamma^2(x - vt)^2 + y^2 + z^2\,, & (5.28) \\
&= \gamma^2\left[(x - vt)^2 + (1 - v^2)(y^2 + z^2)\right]\,. & (5.29)
\end{aligned}
$$

It follows therefore from (5.26) that the scalar and 3-vector potentials in the frame S are given by

$$\phi = \frac{e}{R_*}\,, \qquad \vec{A} = \frac{e\vec{v}}{R_*}\,, \tag{5.30}$$

where we have defined

$$R_*^2 \equiv (x - vt)^2 + (1 - v^2)(y^2 + z^2)\,. \tag{5.31}$$

The electric and magnetic fields can now be calculated in the standard way from ϕ and $\vec{A}$, as in (2.8). Alternatively, and equivalently, we can first calculate $\vec{E}'$ and $\vec{B}'$ in the primed frame, and then Lorentz transform these back to the unprimed frame. In the frame S', we shall of course have

$$\vec{E}' = \frac{e\vec{r}'}{r'^3}\,, \qquad \vec{B}' = 0\,. \tag{5.32}$$

The transformation to the unprimed frame is then given by inverting the standard results (2.56) and (2.57) that express $\vec{E}'$ and $\vec{B}'$ in terms of $\vec{E}$ and $\vec{B}$. Again, this is simply achieved by interchanging the primed and unprimed fields, and sending $\vec{v}$ to $-\vec{v}$. This gives

$$
\begin{aligned}
\vec{E} &= \gamma(\vec{E}' - \vec{v} \times \vec{B}') - \frac{\gamma - 1}{v^2}(\vec{v} \cdot \vec{E}')\, \vec{v}\,, \\
\vec{B} &= \gamma(\vec{B}' + \vec{v} \times \vec{E}') - \frac{\gamma - 1}{v^2}(\vec{v} \cdot \vec{B}')\, \vec{v}\,,
\end{aligned}
\tag{5.33}
$$

and so from (5.32), we find that $\vec{E}$ and $\vec{B}$ in the frame S are given by

$$
\begin{aligned}
\vec{E} &= \frac{e\gamma \vec{r}'}{r'^3} - \frac{\gamma - 1}{v^2}\frac{e\vec{v} \cdot \vec{r}'}{r'^3}\,\vec{v}\,, \\
\vec{B} &= \gamma \vec{v} \times \vec{E}' = \frac{e\gamma \vec{v} \times \vec{r}'}{r'^3}\,.
\end{aligned}
\tag{5.34}
$$

Let us again assume that we orient the axes so that $\vec{v}$ lies along the x direction. Then from the above we find that

$$
E_x = \frac{ex'}{r'^3}\,, \qquad E_y = \frac{e\gamma y'}{r'^3}\,, \qquad E_z = \frac{e\gamma z'}{r'^3}\,,
\tag{5.35}
$$

and so

$$
E_x = \frac{e\gamma(x - vt)}{r'^3}\,, \qquad E_y = \frac{e\gamma y}{r'^3}\,, \qquad E_z = \frac{e\gamma z}{r'^3}\,.
\tag{5.36}
$$

Since the charge is located at the (uniformly moving) point $(vt, 0, 0)$ in the frame S, it follows that the vector from the charge to the point $\vec{r} = (x, y, z)$ is

$$
\vec{R} = (x - vt, y, z)\,.
\tag{5.37}
$$

From (5.36), we then find that the electric field is given by

$$
\vec{E} = \frac{e\gamma \vec{R}}{r'^3} = \frac{e(1 - v^2)\vec{R}}{R_*^3}\,,
\tag{5.38}
$$

where R_* was defined in (5.31).

If we now define θ to be the angle between the vector $\vec{R}$ and the x axis, then the coordinates (x, y, z) of the observation point P will be such that

$$
y^2 + z^2 = R^2 \sin^2\theta\,, \qquad \text{where} \qquad R^2 = |\vec{R}|^2 = (x - vt)^2 + y^2 + z^2\,.
\tag{5.39}
$$

This implies, from (5.31), that

$$
R_*^2 = R^2 - v^2(y^2 + z^2) = R^2(1 - v^2 \sin^2\theta)\,,
\tag{5.40}
$$

and so the electric field due to the moving charge is

$$
\vec{E} = \frac{e\vec{R}}{R^3}\frac{1 - v^2}{(1 - v^2 \sin^2\theta)^{3/2}}\,.
\tag{5.41}
$$

For an observation point P located on the x axis, the electric field will be $E_\parallel$ (parallel to the x axis), and given by setting $\theta = 0$ in (5.41). On the other hand, we can define the electric field $E_\perp$ in the (y, z) plane (corresponding to $\theta = \pi/2$). From (5.41) we therefore have

$$E_\parallel = \frac{e(1 - v^2)}{R^2}, \qquad E_\perp = \frac{e(1 - v^2)^{-1/2}}{R^2}. \tag{5.42}$$

Note that $E_\parallel$ has the smallest magnitude, and $E_\perp$ has the largest magnitude, that $\vec{E}$ attains as a function of θ.

When the velocity is very small, the magnitude of the electric field is (as one would expect) more or less independent of θ. However, as v approaches 1 (the speed of light), we find that $E_\parallel$ decreases to zero, while $E_\perp$ diverges. Thus for v near to the speed of light the electric field is very sharply peaked around $\theta = \frac{\pi}{2}$. If we set

$$\theta = \frac{\pi}{2} - \psi, \tag{5.43}$$

then

$$|\vec{E}| = \frac{e(1 - v^2)}{R^2(1 - v^2 \cos^2 \psi)^{3/2}} \approx \frac{e(1 - v^2)}{(1 - v^2 + \frac{1}{2}\psi^2)^{3/2}} \tag{5.44}$$

if $v \approx 1$. Thus the angular width of the peak is of the order of

$$\psi \sim \sqrt{1 - v^2}. \tag{5.45}$$

We saw previously that the magnetic field in the frame S is given by $\vec{B} = \gamma \vec{v} \times \vec{E}'$. From (5.34) we have $\vec{v} \times \vec{E} = \gamma \vec{v} \times \vec{E}'$, and so therefore

$$\vec{B} = \vec{v} \times \vec{E} = \frac{e(1 - v^2)\vec{v} \times \vec{R}}{R_*^3}. \tag{5.46}$$

Note that if $|\vec{v}| << 1$ we get the usual non-relativistic expressions

$$\vec{E} \approx \frac{e\vec{R}}{R^3}, \qquad \vec{B} \approx \frac{e\vec{v} \times \vec{R}}{R^3}. \tag{5.47}$$

5.4 Motion of a charge in a Coulomb potential

We shall consider a particle of mass m and charge e moving in the field of a static charge Q. The classic "Newtonian" result is very familiar, with the orbit of the particle being a conic section; an ellipse, a parabola or a hyperbola, depending on the charges and the orbital parameters. In this section we shall consider the fully relativistic problem, when the velocity of the particle is not necessarily small compared with the speed of light. Note that we shall be assuming here that the charge Q is fixed, at the origin. In practice, if Q

represented a nucleus and e an electron, the two objects would move around their common centre of gravity. Allowing for this would be a straightforward, if rather tedious, minor distraction from the main points to be studied here, and so we shall make the idealisation that the centre of attraction or repulsion, namely the charge Q, is fixed.

The Lagrangian for the system is given by (2.84), with $\phi = \frac{Q}{r}$ and $\vec{A} = 0$:

$$L = -m(1 - \dot{x}^i \dot{x}^i)^{1/2} - \frac{eQ}{r}, \qquad (5.48)$$

where $\dot{x}^i = \frac{dx^i}{dt}$, and $r^2 = x^i x^i$. The charges occur in the combination eQ throughout the calculation, and so for convenience we shall define

$$\kappa \equiv eQ. \qquad (5.49)$$

Note that if $\kappa > 0$ the force between the charges will be repulsive, while if $\kappa < 0$ it will be attractive.

It is convenient to introduce spherical polar coordinates in the standard way,

$$x = r \sin\theta \cos\varphi, \qquad y = r \sin\theta \sin\varphi, \qquad z = r \cos\theta, \qquad (5.50)$$

and then the Lagrangian becomes

$$L = -m(1 - \dot{r}^2 - r^2\dot{\theta}^2 - r^2 \sin^2\theta \dot{\varphi}^2)^{1/2} - \frac{\kappa}{r}. \qquad (5.51)$$

The Lagrangian is of the form $L = L(q_i, \dot{q}_i)$ for coordinates q_i and velocities $\dot{q}_i$. The Euler-Lagrange equations are

$$\frac{\partial L}{\partial q_i} - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_i}\right) = 0. \qquad (5.52)$$

Note that if the Lagrangian is independent of a particular coordinate, say q_j for some particular value j, then there is an associated conserved quantity $\partial L/\partial \dot{q}_i$:

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_j} = 0. \qquad (5.53)$$

(A coordinate that enters the Lagrangian only through its velocity is sometimes known as an *ignorable coordinate*.)

The Euler-Lagrange equation for θ gives

$$r^2 \sin\theta \cos\theta \dot{\varphi}^2 (1 - \dot{r}^2 - r^2\dot{\theta}^2 - r^2 \sin^2\theta \dot{\varphi}^2)^{-1/2} - \frac{d}{dt}\left(r^2\dot{\theta}(1 - \dot{r}^2 - r^2\dot{\theta}^2 - r^2 \sin^2\theta \dot{\varphi}^2)^{-1/2}\right) = 0. \qquad (5.54)$$

It can be seen that a solution to this equation is to take $\theta = \frac{\pi}{2}$, and $\dot{\theta} = 0$. In other words, if the particle starts out moving in the $\theta = \frac{\pi}{2}$ plane (i.e. the (x, y) plane at $z = 0$), it will

remain in this plane. This is just the familiar result that the motion of a particle moving in a central force lies in a plane. We may therefore assume now, without loss of generality, that $\theta = \frac{\pi}{2}$ for all time. We are left with just r and φ as polar coordinates in the (x, y) plane. The Lagrangian for the reduced system, where we consistently can set $\theta = \frac{\pi}{2}$, is then simply

$$L = -m(1 - \dot{r}^2 - r^2\dot{\varphi}^2)^{1/2} - \frac{\kappa}{r} \,. \tag{5.55}$$

We note that $\frac{\partial L}{\partial \varphi} = 0$, and so there is a conserved quantity

$$\frac{\partial L}{\partial \dot{\varphi}} = mr^2\dot{\varphi}(1 - \dot{r}^2 - r^2\dot{\varphi}^2)^{-1/2} = \ell \,, \tag{5.56}$$

where ℓ is a constant. Since $(1 - \dot{r}^2 - r^2\dot{\varphi}^2)^{-1/2} = (1 - \dot{x}_i \dot{x}_i)^{-1/2} = \gamma$, we simply have

$$m\gamma r^2\dot{\varphi} = \ell \,. \tag{5.57}$$

In fact ℓ is the angular momentum of the particle, measured with respect to the origin. Note that we can also write (5.56 as

$$mr^2\frac{d\varphi}{d\tau} = \ell \,, \tag{5.58}$$

since coordinate time t and proper time τ are related by $d\tau = dt/\gamma$.

Since the Lagrangian does not depend explicitly on t, the total energy $\mathcal{E}$ is also conserved. Thus we have that

$$\mathcal{E} = H = \sqrt{\vec{p}^2 + m^2} + \frac{\kappa}{r} \tag{5.59}$$

is a constant. Here,

$$\begin{aligned} \vec{p}^2 &= m^2\gamma^2\vec{v}^2 = m^2\gamma^2\dot{r}^2 + m^2\gamma^2 r^2\dot{\varphi}^2 \,, \\ &= m^2\Big(\frac{dr}{d\tau}\Big)^2 + m^2 r^2\Big(\frac{d\varphi}{d\tau}\Big)^2 \,, \end{aligned} \tag{5.60}$$

since, as usual, coordinate time and proper time are related by $d\tau = dt/\gamma$.

We therefore have

$$\Big(\mathcal{E} - \frac{\kappa}{r}\Big)^2 = \vec{p}^2 + m^2 = m^2\Big(\frac{dr}{d\tau}\Big)^2 + m^2 r^2\Big(\frac{d\varphi}{d\tau}\Big)^2 + m^2 \,. \tag{5.61}$$

We now perform the standard change of variables in orbit calculations, and let

$$r = \frac{1}{u} \,. \tag{5.62}$$

This implies

$$\frac{dr}{d\tau} = -\frac{1}{u^2}\frac{du}{d\tau} = -\frac{1}{u^2}\frac{du}{d\varphi}\frac{d\varphi}{d\tau} = -\frac{\ell}{m}u' \,, \tag{5.63}$$

86

where we have used (5.58) and also we have defined

$$u' \equiv \frac{du}{d\varphi} \,. \tag{5.64}$$

It now follows that (5.61) becomes

$$(\mathcal{E} - \kappa u)^2 = \ell^2 u'^2 + \ell^2 u^2 + m^2 \,. \tag{5.65}$$

This ordinary differential equation can be solved in order to find u as a function of φ, and hence r as a function of φ. The solution determines the shape of the orbit of the particle around the fixed charge Q.

Rewriting (5.65) as[19]

$$\ell^2 u'^2 = \left(u \sqrt{\kappa^2 - \ell^2} - \frac{\kappa \mathcal{E}}{\sqrt{\kappa^2 - \ell^2}} \right)^2 - m^2 - \frac{\mathcal{E}^2 \ell^2}{\kappa^2 - \ell^2} \,, \tag{5.66}$$

we see that it is convenient to make a change of variable from u to w, defined by

$$u \sqrt{\kappa^2 - \ell^2} - \frac{\kappa \mathcal{E}}{\sqrt{\kappa^2 - \ell^2}} = \pm \sqrt{m^2 + \frac{\mathcal{E}^2 \ell^2}{\kappa^2 - \ell^2}} \, \cosh w \,, \tag{5.67}$$

where the $+$ sign is chosen if $\kappa < 0$ (attractive potential, since $\kappa = eQ$), and the $-$ sign if $\kappa > 0$ (repulsive potential). We can then integrate (5.66), to obtain

$$\frac{\ell}{\sqrt{\kappa^2 - \ell^2}} w = \varphi \,. \tag{5.68}$$

Here, we have made a convenient choice, without loss of generality, for the additive constant of integration, by absorbing it into a choice of origin for the azimuthal angle φ. Thus we have

$$\sqrt{\kappa^2 - \ell^2} \, u = \pm \sqrt{m^2 + \frac{\mathcal{E}^2 \ell^2}{\kappa^2 - \ell^2}} \, \cosh \left[\left(\frac{\kappa^2}{\ell^2} - 1 \right)^{1/2} \varphi \right] + \frac{\kappa \mathcal{E}}{\sqrt{\kappa^2 - \ell^2}} \,. \tag{5.69}$$

In other words, the orbit is given, in terms of $r = r(\varphi) = 1/u(\varphi)$, by

$$\frac{\kappa^2 - \ell^2}{r} = \pm \sqrt{\mathcal{E}^2 \ell^2 + m^2(\kappa^2 - \ell^2)} \, \cosh \left[\left(\frac{\kappa^2}{\ell^2} - 1 \right)^{1/2} \varphi \right] + \kappa \mathcal{E} \,. \tag{5.70}$$

The solution (5.70) is presented for the case where $|\ell| < |\kappa|$. If instead $|\ell| > |\kappa|$ we should instead complete the square as given in footnote 19, leading us to make a cosine substitution rather than a hyperbolic cosine, and the solution becomes

$$\frac{\ell^2 - \kappa^2}{r} = \sqrt{\mathcal{E}^2 \ell^2 - m^2(\ell^2 - \kappa^2)} \, \cos \left[\left(1 - \frac{\kappa^2}{\ell^2} \right)^{1/2} \varphi \right] - \kappa \mathcal{E} \,. \tag{5.71}$$

---

[19]This "completing of the square" is appropriate for the case where $|\ell| < |\kappa|$. If instead $|\ell| > |\kappa|$, we would write

$$\ell^2 u'^2 = -\left(u \sqrt{\ell^2 - \kappa^2} + \frac{\kappa \mathcal{E}}{\sqrt{\ell^2 - \kappa^2}} \right)^2 - m^2 + \frac{\mathcal{E}^2 \ell^2}{\ell^2 - \kappa^2} \,.$$

Finally, if $|\ell| = |\kappa|$, we see that equation (5.65) leads to the solution

$$\frac{2\kappa\mathcal{E}}{r} = \mathcal{E}^2 - m^2 - \mathcal{E}^2\,\varphi^2\,. \tag{5.72}$$

The situation described above for relativistic orbits should be contrasted with what happens in the non-relativistic case. In this limit, the Lagrangian (after restricting to motion in the (x, y) plane again) is simply given by

$$L = \tfrac{1}{2}m(\dot{r}^2 + r^2\dot{\varphi}^2) - \frac{\kappa}{r}\,. \tag{5.73}$$

Note that this can be obtained from the relativisitic Lagrangian (5.55) we studied above, by taking $\dot{r}$ and $r\dot{\varphi}$ to be small compared to 1 (the speed of light), and then expanding the square root to quadratic order in velocities. As discussed previously, one can ignore the leading-order term $-m$ in the expansion, since this is just a constant (the rest-mass energy of the particle) and so it does not enter in the Euler-Lagrange equations. The analysis of the Euler-Lagrange equations for the non-relativistic Lagrangian (5.73) is a standard one. There energy E and angular momentum ℓ are conserved, and given by

$$E = \tfrac{1}{2}m(\dot{r}^2 + r^2\dot{\varphi}^2) + \frac{\kappa}{r}\,, \qquad \ell = mr^2\dot{\varphi}\,. \tag{5.74}$$

Substituting the latter into the former give the standard radial equation, whose solution implies orbits given by

$$\frac{1}{r} = -\frac{m\kappa}{\ell^2}\left(\sqrt{1 + \frac{2E\ell^2}{m\kappa^2}}\,\cos\varphi + 1\right). \tag{5.75}$$

If $\kappa \equiv eQ < 0$ (attractive potential), solutions exist for $E \geq -\frac{m\kappa^2}{2\ell^2}$. Closed orbits occur when $-\frac{m\kappa^2}{2\ell^2} \leq E < 0$; these are elliptical generically, and circular if the lower bound is saturated. The orbits are parabolic if $E = 0$ and hyperbolic if $E > 0$. If $\kappa > 0$ (repulsive potential), solutions exist if $E > 0$ and they are all hyperbolic.

The key difference in the relativistic case is that the orbits do not have a 2π periodicity in φ, even when $|\ell| > |\kappa|$, as in (5.71), for which the radius r is a trigonometric function of φ. The reason for this is that the argument of the trigonometric function is

$$\left(1 - \frac{\kappa^2}{\ell^2}\right)^{1/2}\varphi\,, \tag{5.76}$$

and so φ has to increase through an angle $\Delta\varphi$ given by

$$\Delta\varphi = 2\pi\left(1 - \frac{\kappa^2}{\ell^2}\right)^{-1/2} \tag{5.77}$$

before the cosine completes one cycle. If we assume that $|\kappa/\ell|$ is small compared with 1, then the shape of the orbit is still approximately like an ellipse, except that the "perihelion"

of the ellipse advances by an angle

$$\delta\varphi = 2\pi \left[\left(1 - \frac{\kappa^2}{\ell^2}\right)^{-1/2} - 1\right] \approx \frac{\pi\kappa^2}{\ell^2} \tag{5.78}$$

per orbit. Generically, the orbits are not closed, although they will be in the special case that $\left(1 - \frac{\kappa^2}{\ell^2}\right)^{-1/2}$ is rational.

The fact that the major axis of the ellipse remains fixed in the non-relativistic case is a reflection of the fact that there is an additional "hidden" symmetry in the non-relativistic system, which is broken by the relativistic corrections. Specifically, there is a conserved quantity called the *Runge-Lenz vector* in the non-relativistic theory of a particle of mass m moving in a central $1/r^2$ force $\vec{F} = \kappa\vec{r}/r^3$. It is given by

$$\vec{W} = \vec{p} \times \vec{L} + \frac{m\kappa\vec{r}}{r}, \tag{5.79}$$

where $\vec{L} = \vec{r} \times \vec{p}$ is the angular momentum of the particle about the force centre. It is straightforward to verify that the equations of motion following from the non-relativistic Lagrangian $L = \frac{1}{2}m\dot{x}_i^2 + \kappa/r$ imply that $d\vec{W}/dt = 0$. In the case of an elliptical orbit (i.e. $\kappa < 0$ and $-\frac{m\kappa^2}{2\ell^2} < E < 0$), the Runge-Lenz vector points along the major axis of the ellipse.

Going back to the full relativistic discussion, if $|\ell| \leq |\kappa|$ and if $\kappa < 0$ (which means $eQ < 0$ and hence an attractive force between the charges), we see from the solution (5.70) that the particle spirals inwards and eventually reaches $r = 0$ within a finite time. This can never happen in the non-relativisitic case; the orbit of the particle can never reach the origin at $r = 0$, unless the angular momentum ℓ is exactly zero. The reason for this is that in the non-relativistic case the centrifugal potential term ℓ^2/r^2 always tends to throw the particle away from the origin if r tries to get too small. By contrast, in the relativisitic case the effect of the centrifugal term is reduced at small r, and it cannot prevent the collapse of the orbit to $r = 0$. This can be seen by looking at the conserved quantity $\mathcal{E}$ in the fully relativisitic analysis, which, from our discussion above, can be written as

$$\mathcal{E} = \left(m^2 + m^2\left(\frac{dr}{d\tau}\right)^2 + \frac{\ell^2}{r^2}\right)^{1/2} + \frac{\kappa}{r}. \tag{5.80}$$

First, consider the non-relativistic limit, for which the rest-mass term dominates inside the square root:

$$\mathcal{E} \approx m + \frac{1}{2}m\left(\frac{dr}{dt}\right)^2 + \frac{\ell^2}{2mr^2} + \frac{\kappa}{r}. \tag{5.81}$$

In other words, written in terms of the non-relativistic energy E of eqn (5.74) we have

$$E = \frac{1}{2}m\dot{r}^2 + \frac{\ell^2}{2mr^2} + \frac{\kappa}{r}. \tag{5.82}$$

Here, we see that even if $\kappa < 0$ (an attractive force), the repulsive centrifugal term always wins over the attractive charge term κ/r at small enough r.

On the other hand, if we keep the full relativistic expression (5.80), then at small enough r the competition between the centrifugal term and the charge term becomes evenly matched, in the sense that each has the same $1/r$ power-law behaviour

$$\mathcal{E} \approx \frac{|\ell|}{r} + \frac{\kappa}{r} \,, \tag{5.83}$$

and clearly if $\kappa < -|\ell|$ the attraction between the charges wins the contest.

6 Electromagnetic Waves

6.1 Wave equation

As discussed at the beginning of the course (see section 1.1), Maxwell's equations admit wave-like solutions. These solutions can esist in free space, in a region where there are no source currents, for which the equations take the form

$$\begin{aligned}
\vec{\nabla} \cdot \vec{E} &= 0 \,, & \vec{\nabla} \times \vec{B} - \frac{\partial \vec{E}}{\partial t} &= 0 \,, \\
\vec{\nabla} \cdot \vec{B} &= 0 \,, & \vec{\nabla} \times \vec{E} + \frac{\partial \vec{B}}{\partial t} &= 0 \,.
\end{aligned} \tag{6.1}$$

As discussed in section 1.1, taking the curl of the $\vec{\nabla} \times \vec{E}$ equation, and using the $\vec{\nabla} \times \vec{B}$ equation, one finds

$$\nabla^2 \vec{E} - \frac{\partial^2 \vec{E}}{\partial t^2} = 0 \,, \tag{6.2}$$

and similarly,

$$\nabla^2 \vec{B} - \frac{\partial^2 \vec{B}}{\partial t^2} = 0 \,. \tag{6.3}$$

Thus each component of $\vec{E}$ and each component of $\vec{B}$ satisfies d'Alembert's equation

$$\nabla^2 f - \frac{\partial^2 f}{\partial t^2} = 0 \,. \tag{6.4}$$

This can, of course, be written as

$$\Box f \equiv \partial^\mu \partial_\mu f = 0 \,, \tag{6.5}$$

which shows that d'Alembert's operator is Lorentz invariant.

The wave equation (6.4) admits plane-wave solutions, where f depends on t and on a single linear combination of the x, y and z coordinates. By choosing the orientation of the

90

axes appropriately, we can make this linear combination become simply x. Thus we may seek solutions of (6.4) of the form $f = f(t, x)$. The function f will then satisfy

$$\frac{\partial^2 f}{\partial x^2} - \frac{\partial^2 f}{\partial t^2} = 0 \,, \tag{6.6}$$

which can be written in the factorised form

$$\Big(\frac{\partial}{\partial x} - \frac{\partial}{\partial t}\Big)\Big(\frac{\partial}{\partial x} + \frac{\partial}{\partial t}\Big) f(t, x) = 0 \,. \tag{6.7}$$

Now introduce "light-cone coordinates"

$$u = x - t \,, \qquad v = x + t \,. \tag{6.8}$$

We see that

$$\frac{\partial}{\partial x} = \frac{\partial}{\partial u} + \frac{\partial}{\partial v} \,, \qquad \frac{\partial}{\partial t} = -\frac{\partial}{\partial u} + \frac{\partial}{\partial v} \,, \tag{6.9}$$

and so (6.7) becomes

$$\frac{\partial^2 f}{\partial u \partial v} = 0 \,. \tag{6.10}$$

The general solution to this is

$$f = f_+(u) + f_-(v) = f_+(x - t) + f_-(x + t) \,, \tag{6.11}$$

where f_+ and f_- are arbitrary functions.

The functions $f_\pm$ determine the *profile* of a wave-like disturbance that propagates at the speed of light (i.e. at speed 1). In the case of a wave described by $f_+(x - t)$, the disturbance propagates at the speed of light in the *positive x* direction. This can be seen from the fact that if we sit at a given point on the profile (i.e. at a fixed value of the argument of the function f_+), then as t increases the x value must increase too. This means that the disturbance moves, with speed 1, along the positive x direction. Likewise, a wave described by $f_-(x + t)$ moves in the negative x direction as time increases.

More generally, we can consider a plane-wave disturbance moving along the direction of a unit 3-vector $\vec{n}$:

$$f(t, \vec{r}) = f_+(\vec{n} \cdot \vec{r} - t) + f_-(\vec{n} \cdot \vec{r} + t) \,. \tag{6.12}$$

The f_+ wave moves in the direction of $\vec{n}$ as t increases, while the f_- wave moves in the direction of $-\vec{n}$. The previous case of propagation along the x axis, corresponds to taking $\vec{n} = (1, 0, 0)$.

Let us now return to the discussion of electromagnetic waves. Following the discussion above, there will exist plane-wave solutions of (6.2), propagating along the $\vec{n}$ direction, of the form

$$\vec{E}(\vec{r}, t) = \vec{E}(\vec{n} \cdot \vec{r} - t) \,. \tag{6.13}$$

That is, $\vec{E}$ is a function of the single argument $(\vec{n} \cdot \vec{r} - t)$.

From the Maxwell equation $\partial \vec{B}/\partial t = -\vec{\nabla} \times \vec{E}$, we shall therefore have

$$
\begin{aligned}
\frac{\partial B_i}{\partial t} &= -\epsilon_{ijk} \partial_j E_k(n_\ell x_\ell - t) \,, \\
&= -\epsilon_{ijk} n_j E'(n_\ell x_\ell - t) \,,
\end{aligned}
\tag{6.14}
$$

where E'_k denotes the derivative of E_k with respect to its argument. We also have that $\partial E_k(n_\ell x_\ell - t)/\partial t = -E'_k(n_\ell x_\ell - t)$, and so we can write (6.14) as

$$
\frac{\partial B_i}{\partial t} = \epsilon_{ijk} \, n_j \frac{\partial}{\partial t} E_k(n_\ell x_\ell - t) \,.
\tag{6.15}
$$

This can be integrated with respect to t, dropping the constant of integration since an additional static $\vec{B}$ field term is of no interest to us when discussing electromagnetic waves. Thus we have

$$
B_i = \epsilon_{ijk} n_j E_k \,, \qquad \text{i.e.} \qquad \vec{B} = \vec{n} \times \vec{E} \,.
\tag{6.16}
$$

The source-free Maxwell equation $\vec{\nabla} \cdot \vec{E} = 0$ implies

$$
\partial_i E_i(n_j x_j - t) = n_i E'_i(n_j x_j - t) = -\frac{\partial}{\partial t} \vec{n} \cdot \vec{E} = 0 \,.
\tag{6.17}
$$

Again, we can drop the constant of integration since we are not interested in including static electric fields in this discussion about electromagnetic waves, and conclude that for the plane wave

$$
\vec{n} \cdot \vec{E} = 0 \,.
\tag{6.18}
$$

Since $\vec{B} = \vec{n} \times \vec{E}$, it immediately follows that $\vec{n} \cdot \vec{B} = 0$ and $\vec{E} \cdot \vec{B} = 0$ also. Thus we see that for a plane electromagnetic wave propagating along the $\vec{n}$ direction, the $\vec{E}$ and $\vec{B}$ vectors are orthogonal to $\vec{n}$ and also orthogonal to each other:

$$
\vec{n} \cdot \vec{E} = 0 \,, \qquad \vec{n} \cdot \vec{B} = 0 \,, \qquad \vec{E} \cdot \vec{B} = 0 \,.
\tag{6.19}
$$

It also follows from $\vec{B} = \vec{n} \times \vec{E}$ that

$$
|\vec{E}| = |\vec{B}| \,, \qquad \text{i.e.} \qquad E = B \,.
\tag{6.20}
$$

Thus we find that the energy density W is given by

$$
W = \frac{1}{8\pi}(E^2 + B^2) = \frac{1}{4\pi} E^2 \,.
\tag{6.21}
$$

The Poynting flux $\vec{S} = (\vec{E} \times \vec{B})/(4\pi)$ is given by

$$
\begin{aligned}
S_i &= \frac{1}{4\pi} \epsilon_{ijk} E_j \epsilon_{k\ell m} n_\ell E_m = \frac{1}{4\pi} n_i E_j E_j - \frac{1}{4\pi} E_i n_j E_j \,, \\
&= \frac{1}{4\pi} n_i E_j E_j \,,
\end{aligned}
\tag{6.22}
$$

and so we have

$$W = \frac{1}{4\pi} E^2 , \qquad \vec{S} = \frac{1}{4\pi} \vec{n} E^2 = \vec{n} W . \tag{6.23}$$

Note that the argument $\vec{n} \cdot \vec{r} - t$ can be written as

$$\vec{n} \cdot \vec{r} - t = n_\mu x^\mu , \tag{6.24}$$

where $n_\mu = (-1, \vec{n})$ and hence

$$n^\mu = (1, \vec{n}) . \tag{6.25}$$

Since $\vec{n}$ is a unit vector, $\vec{n} \cdot \vec{n} = 1$, we have

$$n^\mu n_\mu = \eta_{\mu\nu} \, n^\mu n^\nu = 0 . \tag{6.26}$$

n^μ is called a *Null Vector*. This is a non-vanishing vector whose norm $n^\mu n_\mu$ vanishes. Such vectors can arise because of the minus sign in the η_{00} component of the 4-metric. By contrast, in a metric of positive-definite signature, such as the 3-dimensional Euclidean metric δ_{ij}, a vector whose norm vanishes is itself necessarily zero.

We can now evaluate the various components of the energy-momentum tensor, which are given by (4.114) and the equations that follow it. Thus we have

$$
\begin{aligned}
T^{00} &= W = \frac{1}{4\pi} E^2 = \frac{1}{4\pi} B^2 , \\
T^{0i} &= T^{i0} = S_i = n_i W , \\
T^{ij} &= \frac{1}{4\pi} (-E_i E_j - B_i B_j + \tfrac{1}{2}(E^2 + B^2)\delta_{ij}) , \\
&= \frac{1}{4\pi} (-E_i E_j - \epsilon_{ik\ell}\epsilon_{jmn} n_k n_m E_\ell E_n + E^2 \delta_{ij}) , \\
&= \frac{1}{4\pi} (-E_i E_j - \delta_{ij} E^2 - n_i n_k E_k E_j - n_j n_\ell E_\ell E_i + \delta_{ij} n_k n_\ell E_k E_\ell \\
&\qquad + n_i n_j E_\ell E_\ell + n_k n_k E_i E_j + E^2 \delta_{ij}) , \\
&= \frac{1}{4\pi} n_i n_j E^2 = n_i n_j W . \tag{6.27}
\end{aligned}
$$

Note that in deriving this last result, we have used the identity

$$\epsilon_{ik\ell}\epsilon_{jmn} = \delta_{ij}\delta_{km}\delta_{\ell n} + \delta_{im}\delta_{kn}\delta_{\ell j} + \delta_{in}\delta_{kj}\delta_{\ell m} - \delta_{im}\delta_{kj}\delta_{\ell n} - \delta_{ij}\delta_{kn}\delta_{\ell m} - \delta_{in}\delta_{km}\delta_{\ell j} . \tag{6.28}$$

The expressions for T^{00}, T^{0i} and T^{ij} can be combined into the single Lorentz-covariant expression

$$T^{\mu\nu} = n^\mu n^\nu W . \tag{6.29}$$

From this, we can compute the conserved 4-momentum

$$
\begin{aligned}
P^\mu &= \int_{t=\text{const.}} T^{\mu\nu} d\Sigma_\nu = \int T^{\mu 0} d^3 x , \\
&= \int n^\mu W d^3 x = n^\mu \int W d^3 x , \tag{6.30}
\end{aligned}
$$

93

and hence we have

$$P^\mu = n^\mu \, \mathcal{E} \, , \tag{6.31}$$

where

$$\mathcal{E} = \int W d^3x \, , \tag{6.32}$$

the total energy of the electromagnetic field. Note that P^μ is also a null vector,

$$P^\mu P_\mu = \mathcal{E}^2 \, n^\mu n_\mu = 0 \, . \tag{6.33}$$

6.2 Monochromatic plane waves

In the discussion above, we considered plane electromagnetic waves with an arbitrary profile. A special case is to consider the situation when the plane wave has a definite frequency ω, so that its time dependence is of the form $\cos \omega t$. Thus we can write

$$\vec{E} = \vec{E}_0 \, e^{\mathrm{i}(\vec{k} \cdot \vec{r} - \omega t)} \, , \qquad \vec{B} = \vec{B}_0 \, e^{\mathrm{i}(\vec{k} \cdot \vec{r} - \omega t)} \, , \tag{6.34}$$

where $\vec{E}_0$ and $\vec{B}_0$ are (possibly complex) constants. The physical $\vec{E}$ and $\vec{B}$ fields are obtained by taking the real parts of $\vec{E}$ and $\vec{B}$. (Since the Maxwell equations are linear, we can always choose to work in such a complex notation, with the understanding that we take the real parts to get the physical quantities.) It is customary to use the same symbols $\vec{E}$ and $\vec{B}$ both for the complex fields and for the physical fields obtained by taking their real parts. There is usually no risk of confusion, since it should be clear from the context which fields are intended.

As we shall discuss in some detail later, more general electromagnetic wave solutions, including the plane wave with arbitrary profile discussed previously, can be built up as linear combinations of the monochromatic plane-wave solutions. The most general wave solutions correspond to superpositions of monochromatic plane waves summed over all wavevectors and frequencies.

Of course, for the fields in (6.34) to solve the Maxwell equations, there must be relations among the constants $\vec{k}$, ω, $\vec{E}_0$ and $\vec{B}_0$. Specifically, since $\vec{E}$ and $\vec{B}$ must satisfy the wave equations (6.2) and (6.3), we must have

$$\vec{k}^2 = \omega^2 \, , \tag{6.35}$$

and since $\vec{\nabla} \cdot \vec{E} = 0$ and $\vec{\nabla} \cdot \vec{B} = 0$, we must have

$$\vec{k} \cdot \vec{E}_0 = 0 \, , \qquad \vec{k} \cdot \vec{B}_0 = 0 \, . \tag{6.36}$$

Finally, following the discussion in the more general case above, it follows from $\vec{\nabla} \times \vec{E} + \partial \vec{B}/\partial t = 0$ and $\vec{\nabla} \times \vec{B} - \partial \vec{E}/\partial t = 0$ that

$$\vec{B} = \frac{\vec{k} \times \vec{E}}{\omega} \,. \tag{6.37}$$

It is natural, therefore, to introduce the 4-vector

$$k^\mu = (\omega, \vec{k}) = \omega \, n^\mu \,, \tag{6.38}$$

where $n^\mu = (1, \vec{n})$ and $\vec{n} = \vec{k}/|\vec{k}| = \vec{k}/\omega$. Equation (6.35) then becomes simply the statement that k^μ is a null vector,

$$k^\mu k_\mu = 0 \,. \tag{6.39}$$

Note that the argument of the exponentials in (6.34) can now be written as

$$\vec{k} \cdot \vec{r} - \omega t = k_\mu x^\mu \,, \tag{6.40}$$

which we shall commonly write as $k \cdot x$. Thus we may rewrite (6.34) more briefly as

$$\vec{E} = \vec{E}_0 \, e^{\mathrm{i} k \cdot x} \,, \qquad \vec{B} = \vec{B}_0 \, e^{\mathrm{i} k \cdot x} \,. \tag{6.41}$$

As usual, we have a plane transverse wave, propagating in the direction of the unit 3-vector $\vec{n} = \vec{k}/\omega$. The term "transverse" here signifies that $\vec{E}$ and $\vec{B}$ are perpendicular to the direction in which the wave is propagating. In fact, we have

$$\vec{n} \cdot \vec{E} = \vec{n} \cdot \vec{B} = 0 \,, \qquad \vec{B} = \vec{n} \times \vec{E} \,, \tag{6.42}$$

and so we have also that $\vec{E}$ and $\vec{B}$ are perpendicular to each other, and that $|\vec{E}| = \vec{B}|$.

Consider the case where $\vec{E}_0$ is taken to be real, which means that $\vec{B}_0$ is real too. Then the physical fields (obtained by taking the real parts of the fields given in (6.34)), are given by

$$\vec{E} = \vec{E}_0 \, \cos(\vec{k} \cdot \vec{r} - \omega t) \,, \qquad \vec{B} = \vec{B}_0 \, \cos(\vec{k} \cdot \vec{r} - \omega t) \,. \tag{6.43}$$

The energy density is then given by

$$W = \frac{1}{8\pi} (E^2 + B^2) = \frac{1}{4\pi} E_0^2 \, \cos^2(\vec{k} \cdot \vec{r} - \omega t) \,. \tag{6.44}$$

If we define the time average of W by

$$\langle W \rangle \equiv \frac{1}{T} \int_0^T W dt \,, \tag{6.45}$$

where $T = 2\pi/\omega$ is the period of the oscillation, then we shall have

$$\langle W \rangle = \frac{1}{8\pi} E_0^2 = \frac{1}{8\pi} B_0^2 \,. \tag{6.46}$$

Note that in terms of the *complex* expressions (6.34), we can write this as

$$\langle W \rangle = \frac{1}{8\pi} \vec{E} \cdot \vec{E}^* = \frac{1}{8\pi} \vec{B} \cdot \vec{B}^* \,, \tag{6.47}$$

where the $*$ denotes complex conjugation, since the time and position dependence of $\vec{E}$ or $\vec{B}$ is cancelled when multiplied by the complex conjugate field.[20]

In general, when $\vec{E}_0$ and $\vec{B}_0$ are not real, we shall also have the same expressions (6.47) for the time-averaged energy density.

In a similar manner, we can evaluate the time average of the Poynting flux vector $\vec{S} = (\vec{E} \times \vec{B})/(4\pi)$. If we first consider the case where $\vec{E}_0$ is real, we shall have

$$\vec{S} = \frac{1}{4\pi} \vec{E} \times \vec{B} = \frac{1}{4\pi} \vec{E}_0 \times \vec{B}_0 \cos^2(\vec{n} \cdot \vec{r} - \omega t) = \frac{1}{4\pi} \vec{n} \, E_0^2 \cos^2(\vec{n} \cdot \vec{r} - \omega t) \,, \tag{6.48}$$

and so

$$\langle \vec{S} \rangle = \frac{1}{8\pi} \vec{E}_0 \times \vec{B}_0 = \frac{1}{8\pi} \vec{n} \, E_0^2 \,. \tag{6.49}$$

In general, even if $\vec{E}_0$ and $\vec{B}_0$ are not real, we can write $\langle \vec{S} \rangle$ in terms of the complex $\vec{E}$ and $\vec{B}$ fields as

$$\langle \vec{S} \rangle = \frac{1}{8\pi} \vec{E} \times \vec{B}^* = \frac{1}{8\pi} \vec{n} \, \vec{E} \cdot \vec{E}^* \,. \tag{6.50}$$

Note that we have

$$\langle \vec{S} \rangle = \vec{n} \, \langle W \rangle \,. \tag{6.51}$$

6.3 Motion of a point charge in a linearly-polarised E.M. wave

Consider a plane wave propagating in the z direction, with

$$\vec{E} = (E_0 \, \cos \omega(z - t), 0, 0) \,, \qquad \vec{B} = (0, E_0 \, \cos \omega(z - t), 0) \,. \tag{6.52}$$

(Recall that $|\vec{k}| = \omega$.) Suppose now that there is a particle of mass m and charge e in this field. By the Lorentz force equation we shall have

$$\frac{d\vec{p}}{dt} = e\vec{E} + e\vec{v} \times \vec{B} \,. \tag{6.53}$$

---

[20]This "trick," of expressing the time-averaged energy density in terms of the dot product of the complex field with its complex conjugate, is rather specific to this situation, where the quantity being time-averaged is quadratic in the electric and magnetic fields.

For simplicity, we shall make the assumption that the motion of the particle can be treated non-relativistically, and so

$$\vec{p} = m\vec{v} = m\frac{d\vec{r}}{dt}\,. \tag{6.54}$$

Let us suppose that the particle is initially located at the point $z = 0$, and that it moves only by a small amount in comparison to the wavelength $2\pi/\omega$ of the electromagnetic wave. Therefore, to a good approximation, we can assume that the particle is sitting in the uniform, although time-dependent, electromagnetic field obtained by setting $z = 0$ in (6.52). Thus

$$\vec{E} = (E_0\,\cos\omega t, 0, 0)\,, \qquad \vec{B} = (0, E_0\,\cos\omega t, 0)\,, \tag{6.55}$$

and so the Lorentz force equation gives

$$
\begin{aligned}
m\ddot{x} &= eE_0\cos\omega t - e\dot{z}E_0\cos\omega t \approx eE_0\cos\omega t\,, \\
m\ddot{y} &= 0\,, \\
m\ddot{z} &= e\dot{x}\,E_0\cos\omega t\,.
\end{aligned} \tag{6.56}
$$

Note that the approximation in the first line follows from our assumption that the motion of the particle is non-relativistic, so $|\dot{z}| \ll 1$.

With convenient and inessential choices for the constants of integration, we first obtain

$$\dot{x} = \frac{eE_0}{m\omega}\,\sin\omega t\,, \qquad x = -\frac{eE_0}{m\omega^2}\,\cos\omega t\,, \tag{6.57}$$

Substituting into the z equation then gives

$$\ddot{z} = \frac{e^2E_0^2}{m^2\omega}\,\sin\omega t\,\cos\omega t = \frac{e^2E_0^2}{2m^2\omega}\,\sin 2\omega t\,, \tag{6.58}$$

which integrates to give (dropping inessential constants of integration)

$$z = -\frac{e^2E_0^2}{8m^2\omega^3}\,\sin 2\omega t\,. \tag{6.59}$$

The motion in the y direction is purely linear, and since we are not interested in the case where the particle drifts uniformly through space, we can just focus on the solution where y is constant, say $y = 0$.

Thus the interesting motion of the particle in the electromagnetic field is of the form

$$x = \alpha\cos\omega t\,, \qquad z = \beta\sin 2\omega t = 2\beta\sin\omega t\,\cos\omega t\,, \tag{6.60}$$

where

$$\alpha = -\frac{eE_0}{m\omega^2}\,, \qquad \beta = -\frac{e^2E_0^2}{8m^2\omega^3}\,. \tag{6.61}$$

Thus we find

$$z = \frac{2\beta}{\alpha}\, x \sqrt{1 - \frac{x^2}{\alpha^2}}\,. \tag{6.62}$$

This describes a "figure of eight" lying on its side in the (x, z) plane. The assumptions we made in deriving this, namely non-relativistic motion and a small z displacement relative to the wavelength of the electromagnetic wave, can be seen to be satisfied provided the amplitude E_0 of the wave is sufficiently small. Note that the displacement in the z direction is small compared with the displacement in the x direction, since the dimensionless ratio $2\beta/\alpha$ is given by

$$\frac{2\beta}{\alpha} = \frac{eE_0}{4m\omega}\,, \tag{6.63}$$

and we see from the expression for $\dot{x}$ in (6.57) that this ratio is small since $\dot{x}$ is assumed small compared to the speed of light (i.e. 1).

The response of the charge particle to electromagnetic wave provides a model for how the electrons in a receiving antenna behave in the presence of an electromagnetic wave. This shows how the wave is converted into oscilliatory currents in the antenna, which are then amplified and processed into the final output signal in a radio receiver.

6.4   Circular and elliptical polarisation

The electromagnetic wave described in section 6.2 is *linearly polarised.* For example, we could consider the solution with

$$\vec{E}_0 = (0, E_0, 0)\,, \qquad \vec{B}_0 = (0, 0, B_0)\,, \qquad \vec{n} = (1, 0, 0)\,. \tag{6.64}$$

This corresponds to a linearly polarised electromagnetic wave propagating along the x direction.

By taking a linear superposition of waves propagating along a given direction $\vec{n}$, we can obtain circularly polarised, or more generally, elliptically polarised, waves. Let $\vec{e}$ and $\vec{f}$ be two orthogonal unit vectors, that are also both orthogonal to $\vec{n}$:

$$\begin{aligned} \vec{e} \cdot \vec{e} &= 1\,, & \vec{f} \cdot \vec{f} &= 1\,, & \vec{n} \cdot \vec{n} &= 1\,, \\ \vec{e} \cdot \vec{f} &= 0\,, & \vec{n} \cdot \vec{e} &= 0\,, & \vec{n} \cdot \vec{f} &= 0\,. \end{aligned} \tag{6.65}$$

Suppose now we consider a plane wave given by

$$\vec{E} = (E_0\, \vec{e} + \widetilde{E}_0\, \vec{f})\, e^{\mathrm{i}\,(\vec{k} \cdot \vec{r} - \omega t)}\,, \qquad \vec{B} = \vec{n} \times \vec{E}\,, \tag{6.66}$$

where E_0 and $\widetilde{E}_0$ are complex constants. If E_0 and $\widetilde{E}_0$ both have the same phase (i.e. $\widetilde{E}_0/E_0$ is real), then we again have a linearly-polarised electromagnetic wave. If instead the phases of E_0 and $\widetilde{E}_0$ are different, then the wave is in general elliptically polarised.

Consider as an example the case where

$$\widetilde{E}_0 = \pm \mathrm{i}\, E_0\,, \tag{6.67}$$

(with E_0 taken to be real, without loss of generality), for which the electric field will be given by

$$\vec{E} = E_0(\vec{e} \pm \mathrm{i}\, \vec{f})\, e^{\mathrm{i}\,(\vec{k}\cdot\vec{r}-\omega t)}\,. \tag{6.68}$$

Taking the real part, to get the physical electric field, we obtain

$$\vec{E} = E_0\vec{e}\,\cos(\vec{k}\cdot\vec{r}-\omega t) \mp E_0\vec{f}\,\sin(\vec{k}\cdot\vec{r}-\omega t)\,. \tag{6.69}$$

For example, if we choose

$$\vec{n} = (0,0,1)\,, \qquad \vec{e} = (1,0,0)\,, \qquad \vec{f} = (0,1,0)\,, \tag{6.70}$$

then the electric field is given by

$$E_x = E_0\cos\omega(z-t)\,, \qquad E_y = \mp E_0\sin\omega(z-t)\,. \tag{6.71}$$

It is clear from this that the magnitude of the electric field is constant,

$$|\vec{E}| = E_0\,. \tag{6.72}$$

If we fix a value of z, then the $\vec{E}$ vector can be seen to be rotating around the z axis (the direction of motion of the wave). This rotation is *anticlockwise* in the (x,y) plane if we choose the plus sign in (6.67), and *clockwise* if we choose the minus sign instead. These two choices correspond to having a circularly polarised wave of positive or negative helicity respectively. (Positive helicity means the rotation is parallel to the direction of propagation, while negative helicity means the rotation is anti-parallel to the direction of propagation.)

In more general cases, where the magnitudes of E_0 and $\widetilde{E}_0$ are unequal, or where the phase angle between them is not equal to 0 (linear polarisation) or 90 degrees, the electromagnetic wave will be elliptically polarised. Consider, for example, the case where the electric field is given by

$$\vec{E} = (a_1 e^{\mathrm{i}\,\delta_1}, a_2 e^{\mathrm{i}\,\delta_2}, 0)\, e^{\mathrm{i}\,\omega(z-t)}\,, \tag{6.73}$$

with the propagtion direction being $\vec{n} = (0,0,1)$. Then we shall have

$$\vec{B} = \vec{n} \times \vec{E} = (-a_2 e^{\mathrm{i}\,\delta_2}, a_1 e^{\mathrm{i}\,\delta_1}, 0)\, e^{\mathrm{i}\,\omega(z-t)}\,. \tag{6.74}$$

The real constants a_1, a_2, δ_1 and δ_2 determine the nature of this plane wave propagating along the z direction. Of course the overall phase is unimportant, so really it is only the *difference* $\delta_2 - \delta_1$ between the phase angles that is important.

The magnitude and phase information is sometimes expressed in terms of the *Stokes Parameters* (s_0, s_1, s_2, s_3), which are defined by

$$
\begin{aligned}
s_0 &= E_x E_x^* + E_y E_y^* = a_1^2 + a_2^2\,, & s_1 = E_x E_x^* - E_y E_y^* = a_1^2 - a_2^2\,, & \qquad (6.75)\\
s_2 &= 2\Re(E_x^* E_y) = 2a_1 a_2 \cos(\delta_2 - \delta_1)\,, & s_3 = 2\Im(E_x^* E_y) = 2a_1 a_2 \sin(\delta_2 - \delta_1)\,.
\end{aligned}
$$

(The last two involve the real and imaginary parts of $(E_x^* E_y)$ respectively.) The four Stokes parameters are not independent:

$$
s_0^2 = s_1^2 + s_2^2 + s_3^2\,. \qquad (6.76)
$$

The parameter s_0 characterises the intensity of the electromagnetic wave, while s_1 characterises the amount of x polarisation versus y polarisation, with

$$
-s_0 \le s_1 \le s_0\,. \qquad (6.77)
$$

The third independent parameter, which could be taken to be s_2, characterises the phase difference between the x and the y polarised waves. Circular polaristion with $\pm$ helicity corresponds to

$$
s_1 = 0\,, \qquad s_2 = 0\,, \qquad s_3 = \pm s_0\,. \qquad (6.78)
$$

6.5 General superposition of plane waves

So far in the discussion of electromagnetic waves, we have considered the case where there is a single direction of propagation (i.e. a plane wave), and a single frequency (monochromatic). The most general wave-like solutions of the Maxwell equations can be expressed as linear combinations of these basic monochromatic plane-wave solutions.

In order to discuss the general wave solutions, it is helpful to work with the gauge potential $A^\mu = (\phi, \vec{A})$. Recall that we have the freedom to make gauge transformations $A_\mu \to A_\mu + \partial_\mu \lambda$, where λ is an arbitrary function. For the present purposes, of describing wave solutions, a convenient choice of gauge is to set $\phi = 0$. Such a gauge choice would not be very natural when discussing solutions in electrostatics, but in the present case, where we know that the wave solutions are necessarily time-dependent, it is quite helpful. It is know as the *Radiation Gauge*.

Thus, we shall first write a single monochromatic plane wave in terms of the 3-vector potential, as

$$\vec{A} = a\,\vec{e}\,e^{\mathrm{i}\,(\vec{k}\cdot\vec{r}-\omega t)}\,, \tag{6.79}$$

where $\vec{e}$ is a unit *polarisation vector*, and a is a constant. As usual, we must have $|\vec{k}|^2 = \omega^2$. With $\phi = 0$, the electric and magnetic fields will be given by

$$\begin{aligned}
\vec{E} &= -\frac{\partial\vec{A}}{\partial t} = \mathrm{i}\,a\,\omega\,\vec{e}\,e^{\mathrm{i}\,(\vec{k}\cdot\vec{r}-\omega t)}\,, \\
\vec{B} &= \vec{\nabla}\times\vec{A} = \mathrm{i}\,a\,\vec{k}\times\vec{e}\,e^{\mathrm{i}\,(\vec{k}\cdot\vec{r}-\omega t)} = \frac{\vec{k}\times\vec{E}}{\omega}\,.
\end{aligned} \tag{6.80}$$

We can immediately see that $\vec{E}$ and $\vec{B}$ satisfy the wave equation, and that we must impose $\vec{e}\cdot\vec{k} = 0$ in order to satisfy $\vec{\nabla}\cdot\vec{E} = 0$.

We have established, therefore, that (6.79) describes a monochromatic plane wave propagating along the $\vec{k}$ direction, with electric field along $\vec{e}$, provided that $\vec{e}\cdot\vec{k} = 0$ and $|\vec{k}| = \omega$. More precisely, the gauge potential that gives the physical (i.e. real) electric and magnetic fields is given by taking the real part of $\vec{A}$ in (6.79). Thus, when we want to describe the actual physical quantities, we shall write

$$\vec{A} = a\,\vec{e}\,e^{\mathrm{i}\,(\vec{k}\cdot\vec{r}-\omega t)} + a^*\,\vec{e}\,e^{-\mathrm{i}\,(\vec{k}\cdot\vec{r}-\omega t)}\,. \tag{6.81}$$

(We have absorbed a factor of $\frac{1}{2}$ here into a rescaling of a, in order to avoid carrying $\frac{1}{2}$ factors around in all the subsequent equations.) For brevity, we shall usually write the "physical" $\vec{A}$ as

$$\vec{A} = a\,\vec{e}\,e^{\mathrm{i}\,(\vec{k}\cdot\vec{r}-\omega t)} + \text{c.c.}\,, \tag{6.82}$$

where c.c stands for "complex conjugate."

Now consider a general linear superposition of monochromatic plane waves, with different wave-vectors $\vec{k}$, different polarisation vectors $\vec{e}$, and different amplitudes a. We shall therefore label the polarisation vectors and amplitudes as follows:

$$\vec{e} \longrightarrow \vec{e}_\lambda(\vec{k})\,, \qquad a \longrightarrow a_\lambda(\vec{k})\,. \tag{6.83}$$

Here λ is an index which ranges over the values 1 and 2, which labels 2 real orthonormal vectors $\vec{e}_1(\vec{k})$ and $\vec{e}_2(\vec{k})$ that span the 2-plane perpendicular to $\vec{k}$. The general wave solution can then be written as the sum over all such monochromatic plane waves of the form (6.82). Since a continuous range of wave-vectors is allowed, the summation over these will be a 3-dimensional integral. Thus we can write

$$\vec{A} = \sum_{\lambda=1}^{2}\int\frac{d^3\vec{k}}{(2\pi)^3}\left[\vec{e}_\lambda(\vec{k})\,a_\lambda(\vec{k})\,e^{\mathrm{i}\,(\vec{k}\cdot\vec{r}-\omega t)} + \text{c.c.}\right]\,, \tag{6.84}$$

where $\omega = |\vec{k}|$, and

$$\vec{k} \cdot \vec{e}_\lambda(\vec{k}) = 0 , \qquad \vec{e}_\lambda(\vec{k}) \cdot \vec{e}_{\lambda'}(\vec{k}) = \delta_{\lambda\lambda'} . \tag{6.85}$$

For many purposes, it will be convenient to expand $\vec{A}$ in a basis of circularly-polarised monochromatic plane waves, rather than linearly-polarised waves. In this case, we should choose the 2-dimensional basis of polarisation vectors $\vec{\epsilon}_\pm$, related to the previous basis by

$$\vec{\epsilon}_\pm = \frac{1}{\sqrt{2}} \left(\vec{e}_1 \pm \mathrm{i}\, \vec{e}_2 \right) . \tag{6.86}$$

Since we have $\vec{e}_i \cdot \vec{e}_j = \delta_{ij}$, it follows that

$$\vec{\epsilon}_+ \cdot \vec{\epsilon}_+ = 0 , \qquad \vec{\epsilon}_- \cdot \vec{\epsilon}_- = 0 , \qquad \vec{\epsilon}_+ \cdot \vec{\epsilon}_- = 1 . \tag{6.87}$$

Note that $\vec{\epsilon}_\pm{}^* = \vec{\epsilon}_\mp$. We can label the $\vec{\epsilon}_\pm$ basis vectors by $\vec{\epsilon}_\lambda$, where λ is now understood to take the two "values" $+$ and $-$. We then write the general wave solution as

$$\vec{A} = \sum_{\lambda=\pm} \int \frac{d^3\vec{k}}{(2\pi)^3} \left[\vec{\epsilon}_\lambda(\vec{k})\, a_\lambda(\vec{k})\, e^{\mathrm{i}\,(\vec{k}\cdot\vec{r} - \omega t)} + \text{c.c.} \right] , \tag{6.88}$$

Of course, we also have $\vec{k} \cdot \vec{\epsilon}_\lambda = 0$, and $\omega = |\vec{k}|$.

6.5.1 Helicity and energy of circularly-polarised waves

The angular-momentum tensor $M^{\mu\nu}$ for the electromagnetic field is defined by

$$M^{\mu\nu} = \int_{t=\text{const}} (x^\mu T^{\nu\rho} - x^\nu T^{\mu\rho}) d\Sigma_\rho , \tag{6.89}$$

and so the three-dimensional components M^{ij} are

$$
\begin{aligned}
M^{ij} &= \int_{t=\text{const}} (x^i T^{j\rho} - x^j T^{i\rho}) d\Sigma_\rho = \int (x^i T^{j0} - x^j T^{i0}) d^3x , \\
&= \int (x^i S^j - x^j S^i) d^3x .
\end{aligned}
\tag{6.90}
$$

Thus, since $\vec{S} = (\vec{E} \times \vec{B})/(4\pi)$, the three-dimensional angular momentum $L_i = \frac{1}{2}\epsilon_{ijk} M^{jk}$ is given by

$$L_i = \int \epsilon_{ijk} x^j S^k \, d^3x , \tag{6.91}$$

i.e.

$$\vec{L} = \frac{1}{4\pi} \int \vec{r} \times (\vec{E} \times \vec{B}) \, d^3x . \tag{6.92}$$

Now, since $\vec{B} = \vec{\nabla} \times \vec{A}$, we have

$$
\begin{aligned}
[\vec{r} \times (\vec{E} \times \vec{B})]_i &= \epsilon_{ijk}\epsilon_{k\ell m}\, x_j E_\ell B_m\,, \\
&= \epsilon_{ijk}\epsilon_{k\ell m}\epsilon_{mpq}\, x_j E_\ell \partial_p A_q\,, \\
&= \epsilon_{ijk}(\delta_{kp}\delta_{\ell q} - \delta_{kq}\delta_{\ell p})\, x_j E_\ell \partial_p A_q\,, \\
&= \epsilon_{ijk}\, x_j E_\ell \partial_k A_\ell - \epsilon_{ijk}\, x_j E_\ell \partial_\ell A_k\,,
\end{aligned} \tag{6.93}
$$

and so

$$
\begin{aligned}
L_i &= \frac{1}{4\pi}\int (\epsilon_{ijk}\, x_j E_\ell \partial_k A_\ell - \epsilon_{ijk}\, x_j E_\ell \partial_\ell A_k) d^3x\,, \\
&= \frac{1}{4\pi}\int \Big(-\epsilon_{ijk}\, \partial_k(x_j E_\ell)\, A_\ell + \partial_\ell(x_j E_\ell)\, A_k \Big) d^3x\,, \\
&= \frac{1}{4\pi}\int \Big(-\epsilon_{ijk}\, x_j(\partial_k E_\ell)\, A_\ell + \epsilon_{ijk}\, E_j A_k \Big) d^3x\,.
\end{aligned} \tag{6.94}
$$

Note that in performing the integrations by parts here, we have, as usual, assumed that the fields fall off fast enough at infinity that the surface term can be dropped. We have also used the source-free Maxwell equation $\partial_\ell E_\ell = 0$ in getting to the final line. Thus, we conclude that the angular momentum 3-vector can be expressed as

$$
\vec{L} = \frac{1}{4\pi}\int (\vec{E} \times \vec{A} - A_i\, (\vec{r} \times \vec{\nabla})E_i) d^3x\,. \tag{6.95}
$$

The two terms in (6.95) can be interpreted as follows. The second term can be viewed as an "orbital angular momentum," since it clearly depends on the choice of origin. It is rather analogous to an $\vec{r} \times \vec{p}$ contribution to the angular momentum of a system of particles. On the other hand, the first term in (6.95) can be viewed as an "intrinsic spin" term, since it is constructed purely from the electromagnetic fields themselves, and is independent of the choice of origin. We shall calculate this spin contribution,

$$
\vec{L}_{\text{spin}} = \frac{1}{4\pi}\int \vec{E} \times \vec{A}\, d^3x \tag{6.96}
$$

to the angular momentum in the case of the sum over circularly-polarised waves that we introduced in the previous section. Recall that for this sum, the 3-vector potential is given by

$$
\vec{A} = \sum_{\lambda'=\pm}\int \frac{d^3\vec{k}'}{(2\pi)^3}\left[\vec{\epsilon}_{\lambda'}(\vec{k}')\, a_{\lambda'}(\vec{k}')\, e^{\mathrm{i}\,(\vec{k}'\cdot\vec{r}-\omega't)} + \text{c.c.} \right]\,, \tag{6.97}
$$

The electric field is then given by

$$
\vec{E} = -\frac{\partial \vec{A}}{\partial t} = \sum_{\lambda=\pm}\int \frac{d^3\vec{k}}{(2\pi)^3}\left[\mathrm{i}\,\omega\, \vec{\epsilon}_\lambda(\vec{k})\, a_\lambda(\vec{k})\, e^{\mathrm{i}\,(\vec{k}\cdot\vec{r}-\omega t)} + \text{c.c.} \right]\,, \tag{6.98}
$$

103

Note that we have put primes on the summation and integration variables λ and $\vec{k}$ in the expression for $\vec{A}$. This is so that we can take the product $\vec{E} \times \vec{A}$ and not have a clash of "dummy" summation variables, in what will follow below. We have also written the frequency as $\omega' \equiv |\vec{k}'|$ in the expression for $\vec{A}$.

Our interest will be to calculate the time average

$$\langle \vec{L}_{\text{spin}} \rangle \equiv \frac{1}{T} \int_0^T \vec{L}_{\text{spin}} dt \,. \tag{6.99}$$

Since we are considering a wave solution with an entire "chorus" of frequencies now, we define the time average by taking T to infinity. (It is easily seen that this coincides with the previous definition of the time average for a monochromatic wave of frequency ω, where T was taken to be $2\pi/\omega$.) Note that the time average will be zero for any quantity whose time dependence is of the oscilliatory form $e^{i\nu t}$, because we would have

$$\frac{1}{T} \int_0^T e^{i\nu t} dt = \frac{1}{i\nu T}(e^{i\nu T} - 1) \,, \tag{6.100}$$

which clearly goes to zero as T goes to infinity. Since the time dependence of all the quantities we shall consider is precisely of the form $e^{i\nu t}$, it follows that in order to survive the time averaging, it must be that $\nu = 0$. Thus we have $\langle e^{i\nu t} \rangle = 0$ if $\nu \neq 0$ and $\langle e^{i\nu t} \rangle = 1$ if $\nu = 0$.

We are interested in calculating the time average of $\vec{E} \times \vec{A}$, where $\vec{A}$ and $\vec{E}$ are given by (6.97) and (6.98). The quantities ω appearing there are, by definition, positive, since we have defined $\omega \equiv |\vec{k}|$. The only way that we shall get terms in $\vec{E} \times \vec{A}$ that have zero frequency (i.e. $\nu = 0$) is from the product of one of the terms that is explicitly written times one of the "c.c." terms, since these, of course, have the opposite sign for their frequency dependence.

The upshot of this discussion is that when we evaluate the time average of $\vec{E} \times \vec{A}$, with $\vec{A}$ and $\vec{E}$ given by (6.97) and (6.98), the only terms that survive will be coming from the product of the explicitly-written term for $\vec{E}$ times the "c.c." term for $\vec{A}$, plus the "c.c." term for $\vec{E}$ times the explicitly-written term for $\vec{A}$. Furthermore, in order for the products to have zero frequency, and therefore survive the time averaging, it must be that $\omega' = \omega$. We therefore find

$$
\begin{aligned}
\langle \vec{E} \times \vec{A} \rangle \;=\; & \sum_{\lambda\lambda'} \int \frac{d^3\vec{k}}{(2\pi)^3} \frac{d^3\vec{k}'}{(2\pi)^3} i\,\omega \left[\vec{\epsilon}_\lambda(\vec{k}) \times \vec{\epsilon}^{\,*}_{\lambda'}(\vec{k}') a_\lambda(\vec{k}) a^*_{\lambda'}(\vec{k}')\, e^{i(\vec{k}-\vec{k}')\cdot\vec{r}} \right. \\
& \left. - \vec{\epsilon}^{\,*}_\lambda(\vec{k}) \times \vec{\epsilon}_{\lambda'}(\vec{k}') a^*_\lambda(\vec{k}) a_{\lambda'}(\vec{k}')\, e^{-i(\vec{k}-\vec{k}')\cdot\vec{r}} \right] \tag{6.101}
\end{aligned}
$$

We now need to integrate $\langle \vec{E} \times \vec{A} \rangle$ over all 3-space, which we shall write as

$$\int \langle \vec{E} \times \vec{A} \rangle \, d^3 \vec{r}. \tag{6.102}$$

We now make use of the result from the theory of delta functions that

$$\int e^{\mathrm{i}(\vec{k}-\vec{k}')\cdot\vec{r}} \, d^3 \vec{r} = (2\pi)^3 \, \delta^3(\vec{k}-\vec{k}'). \tag{6.103}$$

Therefore, from (6.101) we find

$$\int \langle \vec{E} \times \vec{A} \rangle \, d^3 \vec{r} = \sum_{\lambda\lambda'} \int \frac{d^3\vec{k}}{(2\pi)^3} \, \mathrm{i}\,\omega \left[\vec{\epsilon}_\lambda(\vec{k}) \times \vec{\epsilon}_{\lambda'}^*(\vec{k}) a_\lambda(\vec{k}) a_{\lambda'}^*(\vec{k}) \right.$$
$$\left. - \vec{\epsilon}_\lambda^*(\vec{k}) \times \vec{\epsilon}_{\lambda'}(\vec{k}) a_\lambda^*(\vec{k}) a_{\lambda'}(\vec{k}) \right]. \tag{6.104}$$

Finally, we recall that the polarization vectors $\vec{\epsilon}_\pm(\vec{k})$ span the 2-dimensional space orthogonal to the wave-vector $\vec{k}$. In terms of the original real basis unit vectors $\vec{e}_1(\vec{k})$ and $\vec{e}_2(\vec{k})$ we have

$$\vec{e}_1(\vec{k}) \times \vec{e}_2(\vec{k}) = \frac{\vec{k}}{\omega}, \tag{6.105}$$

and so it follows from (6.86) that

$$\vec{\epsilon}_+(\vec{k}) \times \vec{\epsilon}_+^*(\vec{k}) = -\frac{\mathrm{i}\,\vec{k}}{\omega}, \qquad \vec{\epsilon}_-(\vec{k}) \times \vec{\epsilon}_-^*(\vec{k}) = \frac{\mathrm{i}\,\vec{k}}{\omega}. \tag{6.106}$$

From this, it follows that (6.104) becomes

$$\int \langle \vec{E} \times \vec{A} \rangle \, d^3\vec{r} = 2 \int \frac{d^3\vec{k}}{(2\pi)^3} \, \vec{k} \left[a_+(\vec{k}) a_+^*(\vec{k}) - a_-(\vec{k}) a_-^*(\vec{k}) \right], \tag{6.107}$$

and so we have

$$\langle \vec{L}_{\mathrm{spin}} \rangle = \frac{1}{2\pi} \int \frac{d^3\vec{k}}{(2\pi)^3} \, \vec{k} \left(|a_+(\vec{k})|^2 - |a_-(\vec{k})|^2 \right). \tag{6.108}$$

It can be seen from this result that the modes associated with the coefficients $a_+(\vec{k})$ correspond to circularly-polarised waves of positive helicity; i.e. their spin is *parallel* to the wave-vector $\vec{k}$. Conversely, the modes with coefficients $a_-(\vec{k})$ correspond to circularly-polarised waves of negative helicity; i.e. with spin that is *anti-parallel* to the wave-vector $\vec{k}$.

In a similar fashion, we may evaluate the energy of the general wave solution as a sum over the individual modes. The total energy $\mathcal{E}$ is given by[21]

$$\mathcal{E} = \frac{1}{8\pi} \int (E^2 + B^2) d^3x \longrightarrow \frac{1}{4\pi} \int E^2 d^3x. \tag{6.109}$$

[21]We are being a little bit sloppy here, in invoking the result, shown earlier for a single monochromatic plane wave, that the electric and magnetic fields give equal contributions to the energy. It is certainly not true any longer that $\vec{E}^2 = \vec{B}^2$ for a general superposition of plane waves. However, after integrating over

Since $\vec{E} = -\partial \vec{A}/\partial t$ here, we have

$$
\begin{aligned}
\langle E^2 \rangle &= \sum_{\lambda,\lambda'} \int \frac{d^3 \vec{k}}{(2\pi)^3} \frac{d^3 \vec{k}'}{(2\pi)^3} \, \omega^2 \Big[\vec{\epsilon}_\lambda(\vec{k}) \cdot \vec{\epsilon}_{\lambda'}^*(\vec{k}') \, a_\lambda(\vec{k}) a_{\lambda'}^*(\vec{k}') \, e^{i(\vec{k}-\vec{k}')\cdot\vec{r}} \\
&\qquad\qquad + \vec{\epsilon}_\lambda^*(\vec{k}) \cdot \vec{\epsilon}_{\lambda'}(\vec{k}') \, a_\lambda^*(\vec{k}) a_{\lambda'}(\vec{k}') \, e^{-i(\vec{k}-\vec{k}')\cdot\vec{r}} \Big] \,, \quad (6.110)
\end{aligned}
$$

where again, the time-averaging has picked out only the terms whose total frequency adds to zero. The integration over all space then again gives a three-dimensional delta function $\delta^3(\vec{k} - \vec{k}')$, and so we find

$$
\begin{aligned}
\int \langle E^2 \rangle d^3 \vec{r} &= \sum_{\lambda,\lambda'} \int \frac{d^3 \vec{k}}{(2\pi)^3} \omega^2 \Big[\vec{\epsilon}_\lambda(\vec{k}) \cdot \vec{\epsilon}_{\lambda'}^*(\vec{k}) \, a_\lambda(\vec{k}) a_{\lambda'}^*(\vec{k}) \\
&\qquad\qquad + \vec{\epsilon}_\lambda^*(\vec{k}) \cdot \vec{\epsilon}_{\lambda'}(\vec{k}) \, a_\lambda^*(\vec{k}) a_{\lambda'}(\vec{k}) \Big] \,, \quad (6.111)
\end{aligned}
$$

Finally, using the orthogonality relations (6.87), and the conjugation identity $\vec{\epsilon}_\pm = \vec{\epsilon}_\mp^*$, we obtain

$$
\langle \mathcal{E} \rangle = \frac{1}{2\pi} \int \frac{d^3 \vec{k}}{(2\pi)^3} \, \omega^2 \Big(|a_+(\vec{k})|^2 + |a_-(\vec{k})|^2 \Big) \,. \qquad (6.112)
$$

From the two results (6.108) and (6.112), we see that for a given mode characterised by helicity λ and wave-vector $\vec{k}$, we have

$$
\begin{aligned}
\langle \vec{L}_{\text{spin}} \rangle_{\vec{k},\lambda} &= \frac{1}{2\pi} \vec{k} \, |a_\lambda(\vec{k})|^2 \, (\text{sign}\,\lambda) \,, \\
\langle \mathcal{E} \rangle_{\vec{k},\lambda} &= \frac{1}{2\pi} \, \omega^2 \, |a_\lambda(\vec{k})|^2 \,, \qquad\qquad\qquad (6.113)
\end{aligned}
$$

where $(\text{sign}\,\lambda)$ is $+1$ for $\lambda = +$ and -1 for $\lambda = -$. The helicity σ, which is the component of spin along the direction of the wave-vector $\vec{k}$, is therefore given by

$$
\begin{aligned}
\sigma &= \frac{1}{2\pi} |\vec{k}| \, |a_\lambda(\vec{k})|^2 \, (\text{sign}\,\lambda) \,, \\
&= \frac{1}{2\pi} \, \omega \, |a_\lambda(\vec{k})|^2 \, (\text{sign}\,\lambda) \,, \\
&= \frac{1}{\omega} \, \langle \mathcal{E} \rangle_{\vec{k},\lambda} \, (\text{sign}\,\lambda) \,. \qquad\qquad (6.114)
\end{aligned}
$$

In other words, we have that

$$
\text{energy} = \pm(\text{helicity})\,\omega \,, \qquad\qquad (6.115)
$$

and so we can write

$$\mathcal{E} = |\sigma| \, \omega \,. \tag{6.116}$$

This can be compared with the result in quantum mechanics, that

$$E = \hbar \, \omega \,. \tag{6.117}$$

Planck's constant $\hbar$ has the units of angular momentum, and in fact the basic "unit" of angular momentum for the photon is one unit of $\hbar$. In the transition from classical to quantum physics, the helicity of the electromagnetic field becomes the spin of the photon.

6.6 Gauge invariance and electromagnetic fields

In the previous discussion, we described electromagnetic waves in terms of the gauge potential $A_\mu = (-\phi, \vec{A})$, working in the radiation gauge where $\phi = 0$, i.e. $A_0 = 0$. Since the gauge symmetry of Maxwell's equations is

$$A_\mu \longrightarrow A_\mu + \partial_\mu \lambda \,, \tag{6.118}$$

one might think that all the gauge freedom had been used up when we imposed the condition $\phi = 0$, on the grounds that one arbitrary function (the gauge parameter λ) has been used in order to set one function (the scalar potential ϕ) to zero. This is, in fact, not the case. To see this, recall that for the electromagnetic wave we wrote $\vec{A}$ as a superposition of terms of the form

$$\vec{A} = \vec{c} \, e^{\mathrm{i} \, (\vec{k} \cdot \vec{r} - \omega t)} \,, \tag{6.119}$$

which implied that

$$\vec{E} = -\frac{\partial \vec{A}}{\partial t} = \mathrm{i} \, \omega \, \vec{c} \, e^{\mathrm{i} \, (\vec{k} \cdot \vec{r} - \omega t)} \,. \tag{6.120}$$

From this we have

$$\vec{\nabla} \cdot \vec{E} = -\omega \vec{k} \cdot \vec{c} \, e^{\mathrm{i} \, (\vec{k} \cdot \vec{r} - \omega t)} \,, \tag{6.121}$$

and so the Maxwell equation $\vec{\nabla} \cdot \vec{E} = 0$ implies that $\vec{k} \cdot \vec{c} = 0$, and hence

$$\vec{k} \cdot \vec{A} = 0 \,. \tag{6.122}$$

This means that as well as having $A_0 = -\phi = 0$, we also have a component of $\vec{A}$ vanishing, namely the projection along $\vec{k}$.

To see how this can happen, it is helpful to go back to a Lorentz-covariant gauge choice instead. First, consider the Maxwell field equation, in the absence of source currents:

$$\partial^\mu F_{\mu\nu} = 0 \,. \tag{6.123}$$

107

Since $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$, this implies

$$\partial^\mu\partial_\mu A_\nu - \partial^\mu\partial_\nu A_\mu = 0. \tag{6.124}$$

We now choose the Lorenz gauge condition,

$$\partial^\mu A_\mu = 0. \tag{6.125}$$

The field equation (6.124) then reduces to

$$\partial^\mu\partial_\mu A_\nu = 0, \qquad \text{i.e.} \qquad \Box A_\mu = 0. \tag{6.126}$$

One might again think that all the gauge symmetry had been "used up" in imposing the Lorenz gauge condition (6.125), on the grounds that the arbitrary function λ in the gauge transformation

$$A_\mu \longrightarrow A_\mu + \partial_\mu\lambda \tag{6.127}$$

that allowed one to impose (6.125) would no longer allow any freedom to impose further conditions on A_μ. This is not quite true, however.

To see this, let us suppose we are already in Lorenz gauge, and then try performing a further gauge transformation, as in (6.127), insisting that we must remain in the Lorenz gauge. This means that λ should satisfy

$$\partial^\mu\partial_\mu\lambda = 0, \qquad \text{i.e.} \qquad \Box\lambda = 0. \tag{6.128}$$

Non-trivial such functions λ can of course exist; any solution of the wave equation will work. So the upshot is that even after having imposed the Lorenz gauge condition $\partial_\mu A^\mu = 0$ we can still make a restricted class of further gauge transformations $A_\mu \to A_\mu + \partial_\mu\lambda$, where λ is any solution of $\Box\lambda = 0$, and this will not upset the Lorenz gauge condition.

To see what this implies, let us begin with a general solution of the wave equation (6.126), working in the Lorenz gauge (6.125). We can decompose this solution as a sum over plane waves, where a typical mode in the sum is

$$A_\mu = a_\mu\, e^{\mathrm{i}(\vec{k}\cdot\vec{r} - \omega t)} = a_\mu\, e^{\mathrm{i}\,k_\nu x^\nu} = a_\mu\, e^{\mathrm{i}\,k\cdot x}, \tag{6.129}$$

where a_μ and k_ν are constant. Substituting into the wave equation (6.126) we find

$$0 = \Box A_\mu = \partial^\sigma\partial_\sigma(a_\mu\, e^{\mathrm{i}\,k\cdot x}) = -k^\sigma k_\sigma\, a_\mu\, e^{\mathrm{i}\,k\cdot x}, \tag{6.130}$$

whilst the Lorenz gauge condition (6.125) implies

$$0 = \partial^\mu A_\mu = \partial^\mu(a_\mu\, e^{\mathrm{i}\,k\cdot x}) = \mathrm{i}\,k^\mu a_\mu\, e^{\mathrm{i}\,k\cdot x}. \tag{6.131}$$

108

In other words, k_μ and a_μ must satisfy

$$k^\mu k_\mu = 0\,, \qquad k^\mu a_\mu = 0\,. \tag{6.132}$$

The first of these equations implies that k^μ is a null vector, as we had seen earlier. The second equation implies that 1 of the 4 independent components that a 4-vector a_μ generically has is restricted in this case, so that a_μ has only 3 independent components.

Now we perform the further gauge transformation $A_\mu \rightarrow A_\mu + \partial_\mu \lambda$, where, as discussed above, $\Box \lambda = 0$ so that we keep the gauge-transformed A_μ in Lorenz gauge. Specifically, we shall choose

$$\lambda = \mathrm{i}\, h\, e^{\mathrm{i}\, k \cdot x}\,, \tag{6.133}$$

where h is a constant. (The factor of i is there for later convenience.) Thus we shall have

$$A_\mu \longrightarrow A_\mu - h\, k_\mu\, e^{\mathrm{i}\, k \cdot x}\,. \tag{6.134}$$

With A_μ given by (6.129) this means we shall have

$$a_\mu\, e^{\mathrm{i}\, k \cdot x} \longrightarrow a_\mu\, e^{\mathrm{i}\, k \cdot x} - h\, k_\mu\, e^{\mathrm{i}\, k \cdot x}\,, \tag{6.135}$$

which implies

$$a_\mu \longrightarrow a_\mu - h\, k_\mu\,. \tag{6.136}$$

As a check, we can see that the redefined a_μ indeed still satisfies $k^\mu a_\mu = 0$, as it should, since k^μ is a null vector.

The upshot of this discussion is that the freedom to take the constant h to be anything we like allows us to place a *second* restriction on the components of a_μ. Thus not merely are its ostensible 4 components reduced to 3 by virtue of $k^\mu a_\mu = 0$, but a further component can be eliminated by means of the residual gauge freedom, leaving just 2 independent components in the polarisation vector a_μ. Since the physical degrees of freedom are, by definition, the independent quantities that cannot be changed by making gauge transformations, we see that there are 2 degrees of freedom in the electromagnetic wave, and not 3 as one might naively have supposed.

These 2 physical degrees of freedom can be organised as the $+$ and $-$ helicity states, just as we did in our earlier discussion. These are the circularly-polarised waves rotating anti-clockwise and clockwise, respectively. In other words, these are the states whose spin is either parallel, or anti-parallel, to the direction of propagation. One way of understanding why we have only 2, and not 3, allowed states is that the wave is travelling at the speed of

light, and so it is not possible for it to have a helicity that projects other than fully parallel or anti-parallel to its direction of propagation.

We can make contact with the $\phi = 0$ gauge choice that we made in our previous discussion of electromagnetic waves. Starting in Lorenz gauge, we make use of the residual gauge transformation (6.136) by choosing h so that

$$a_0 - h\,k_0 = 0\,, \qquad \text{i.e.} \qquad h = -\frac{a_0}{\omega}\,. \tag{6.137}$$

this means that after performing the residual gauge transformation we shall have

$$a_0 = 0\,, \tag{6.138}$$

and so, from (6.129), we shall have

$$A_0 = 0\,, \qquad \text{i.e.} \qquad \phi = 0\,. \tag{6.139}$$

The original Lorenz gauge condition (6.125) then reduces to

$$\partial_i A_i = 0\,, \qquad \text{i.e.} \qquad \vec{\nabla} \cdot \vec{A} = 0\,. \tag{6.140}$$

This implies $\vec{k} \cdot \vec{A} = 0$, and so we have reproduced precisely the $\phi = 0$, $\vec{k} \cdot \vec{A} = 0$ gauge conditions that we used previously in our analysis of the general electromagnetic wave solutions.

In D spacetime dimensions, the analogous result can easily be seen to be that the electromagnetic wave has $(D-2)$ degrees of freedom.

6.7 Fourier decomposition of electrostatic fields

We saw earlier in 6.5 that an electromagnetic wave, expressed in the radiation gauge in terms of the 3-vector potential $\vec{A}$, could be decomposed into Fourier modes as in (6.88). For each mode $\vec{A}_{(\vec{k},\lambda)}$ in the sum, we have $\vec{\epsilon}_\lambda(\vec{k}) \cdot \vec{k} = 0$, and so each mode of the electric field $\vec{E}_{(\vec{k},\lambda)} = -\partial \vec{A}_{(\vec{k},\lambda)}/\partial t$ satisfies the transversality condition

$$\vec{k} \cdot \vec{E}_{(\vec{k},\lambda)} = 0\,. \tag{6.141}$$

By constrast, an *electrostatic* field $\vec{E}$ is longitudinal. Consider, for example, a point charge at the origin, whose potential therefore satisfies

$$\nabla^2 \phi = -4\pi e\,\delta^3(\vec{r})\,. \tag{6.142}$$

110

We can express $\phi(\vec{r})$ in terms of its Fourier transform $\Phi(\vec{k})$ as

$$\phi(\vec{r}) = \int \frac{d^3\vec{k}}{(2\pi)^3} \, \Phi(\vec{k}) \, e^{i\,\vec{k}\cdot\vec{r}} \,. \tag{6.143}$$

This is clearly a sum over zero-frequency waves, as one would expect since the fields are static.

It follows from (6.143) that

$$\nabla^2 \phi(\vec{r}) = -\int \frac{d^3\vec{k}}{(2\pi)^3} \, \vec{k}^{\,2} \, \Phi(\vec{k}) \, e^{i\,\vec{k}\cdot\vec{r}} \,. \tag{6.144}$$

We also note that the delta-function in (6.142) can be written as

$$\delta^3(\vec{r}) = \int \frac{d^3\vec{k}}{(2\pi)^3} \, e^{i\,\vec{k}\cdot\vec{r}} \,. \tag{6.145}$$

It follows that if we substitute (6.143) into (6.142) we shall obtain $-\vec{k}^{\,2}\Phi(\vec{k}) = -4\pi e$, and hence

$$\Phi(\vec{k}) = \frac{4\pi e}{k^2} \,. \tag{6.146}$$

The electric field is given by $\vec{E} = -\vec{\nabla}\phi$, and so

$$\vec{E} = -i \int \frac{d^3\vec{k}}{(2\pi)^3} \, \vec{k} \, \Phi(\vec{k}) \, e^{i\,\vec{k}\cdot\vec{r}} \,. \tag{6.147}$$

If we define $\vec{G}(\vec{k})$ to be the Fourier transform of $\vec{E}$, so that

$$\vec{E}(\vec{r}) = \int \frac{d^3\vec{k}}{(2\pi)^3} \, \vec{G}(\vec{k}) \, e^{i\,\vec{k}\cdot\vec{r}} \,, \tag{6.148}$$

then we see that

$$\vec{G}(\vec{k}) = -i\,\vec{k}\,\Phi(\vec{k}) = -\frac{4\pi\,i\,e}{k^2}\,\vec{k} \,. \tag{6.149}$$

Thus we see that $\vec{G}(\vec{k})$ is parallel to $\vec{k}$, which proves that the electrostatic field is *Longitudinal*.

6.8   Waveguides

For our purposes, we shall define a waveguide to be a hollow, perfectly conducting, cylinder, essentially of infinite length. For convenience we shall take the axis of the cylinder to lie along the z direction. The cross-section of the cylinder, in the (x, y) plane, can for now be arbitrary, but it is the same for all values of z. Thus, the cross-section through the cylinder is a closed curve.

We shall consider an electromagnetic wave propagating down the cylinder, with angular frequency ω. It will therefore have z and t dependence of the form

$$e^{\mathrm{i}\,(kz-\omega t)}\,.\tag{6.150}$$

Note that inlike the situation for a monochromatic plane wave in infinite free space, the constants k and ω will not in general be equal here. In other words, the electromagnetic wave propagating down the waveguide does not travel at the speed of light. Note that with all fields having z and t dependence of the form (6.150), we may make the replacements

$$\frac{\partial}{\partial t} \longrightarrow -\mathrm{i}\,\omega\,, \qquad \frac{\partial}{\partial z} \longrightarrow \mathrm{i}\,k\,.\tag{6.151}$$

The source-free Maxwell equations (which hold inside the waveguide), therefore imply

$$\begin{aligned}\vec{\nabla}\cdot\vec{E} &= 0\,, & \vec{\nabla}\times\vec{E} &= \mathrm{i}\,\omega\,\vec{B}\,,\\ \vec{\nabla}\cdot\vec{B} &= 0\,, & \vec{\nabla}\times\vec{B} &= -\mathrm{i}\,\omega\,\vec{E}\,.\end{aligned}\tag{6.152}$$

Because of the assumed form of the z dependence in (6.150), we may write

$$\vec{E}(x,y,z,t) = \vec{E}(x,y)\,e^{\mathrm{i}\,(kz-\omega t)}\,, \qquad \vec{B}(x,y,z,t) = \vec{B}(x,y)\,e^{\mathrm{i}\,(kz-\omega t)}\,.\tag{6.153}$$

It is convenient also to define the unit vector $\vec{m}$ in the z direction (the axis of the waveguide),

$$\vec{m} = (0,0,1)\,,\tag{6.154}$$

and certain transverse quantities, denoted with a $\perp$ subscript, as follows:

$$\begin{aligned}\vec{\nabla}_{\perp} &\equiv \left(\frac{\partial}{\partial x},\frac{\partial}{\partial y},0\right),\\ \vec{E} &\equiv \vec{E}_{\perp} + \vec{m}\,E_z\,, & \vec{B} &\equiv \vec{B}_{\perp} + \vec{m}\,B_z\,.\end{aligned}\tag{6.155}$$

(Note that therefore $\vec{E}_{\perp} = (E_x, E_y, 0)$ and $\vec{B}_{\perp} = (B_x, B_y, 0)$.) From (6.152), the Maxwell equations become

$$\begin{aligned}\vec{\nabla}\cdot\vec{E} = 0: && \vec{\nabla}_{\perp}\cdot\vec{E}_{\perp} &= -\mathrm{i}\,k\,E_z\,,\\ \vec{\nabla}\cdot\vec{B} = 0: && \vec{\nabla}_{\perp}\cdot\vec{B}_{\perp} &= -\mathrm{i}\,k\,B_z\,,\end{aligned}$$

$$\begin{aligned}\vec{\nabla}\times\vec{E} = \mathrm{i}\,\omega\,\vec{B}: && \mathrm{i}\,k\vec{E}_{\perp} + \mathrm{i}\,\omega\,\vec{m}\times\vec{B}_{\perp} &= \vec{\nabla}_{\perp}E_z\,,\\ && \vec{m}\cdot(\vec{\nabla}_{\perp}\times\vec{E}_{\perp}) &= \mathrm{i}\,\omega\,B_z\,,\end{aligned}$$

$$\begin{aligned}\vec{\nabla}\times\vec{B} = -\mathrm{i}\,\omega\,\vec{E}: && \mathrm{i}\,k\vec{B}_{\perp} - \mathrm{i}\,\omega\,\vec{m}\times\vec{E}_{\perp} &= \vec{\nabla}_{\perp}B_z\,,\\ && \vec{m}\cdot(\vec{\nabla}_{\perp}\times\vec{B}_{\perp}) &= -\mathrm{i}\,\omega\,E_z\,.\end{aligned}\tag{6.156}$$

Note that the cross product of any pair of transverse vectors, $\vec{U}_\perp \times \vec{V}_\perp$, lies purely in the z direction, i.e. parallel to $\vec{m}$. In components, the last four lines in (6.156) are:

$$
\begin{aligned}
\mathrm{i}\,kE_x - \mathrm{i}\,\omega B_y &= \partial_x E_z\,, \qquad \mathrm{i}\,kE_y + \mathrm{i}\,\omega B_x = \partial_y E_z\,, \\
\partial_x E_y - \partial_y E_x &= \mathrm{i}\,\omega B_z\,, \\
\mathrm{i}\,kB_x + \mathrm{i}\,\omega E_y &= \partial_x B_z\,, \qquad \mathrm{i}\,kB_y - \mathrm{i}\,\omega E_x = \partial_y B_z\,, \\
\partial_x B_y - \partial_y B_x &= -\mathrm{i}\,\omega E_z\,,
\end{aligned}
\tag{6.157}
$$

where $\partial_x = \partial/\partial x$ and $\partial_y = \partial/\partial y$. Note that the four equations comprising lines 1 and 3 can be solved algebraically for E_x, E_y, B_x and B_y in terms of the first derivatives of E_z and B_z. We shall come back to this later.

6.8.1 TEM modes

There are various types of modes that can be considered. First, we may dispose of an "uninteresting" possibility, called *TEM modes*. The acronym stands for "transverse electric and magnetic," meaning that both $\vec{E}$ and $\vec{B}$ have only transverse components, with no longitudinal component along z:

$$
E_z = 0\,, \qquad B_z = 0\,.
\tag{6.158}
$$

From the equations in (6.156) for $\vec{E}_\perp$, we see that

$$
\vec{\nabla}_\perp \cdot \vec{E}_\perp = 0\,, \qquad \vec{\nabla}_\perp \times \vec{E}_\perp = 0\,.
\tag{6.159}
$$

These are the equations for electrostatics in the 2-dimensional (x, y) plane. The second equation implies we can write $\vec{E}_\perp = -\vec{\nabla}_\perp \phi$, and then the first equation implies that the electrostatic potential ϕ satisfies the 2-dimensional Laplace equation

$$
\nabla_\perp^2 \phi = \frac{\partial^2 \phi}{\partial x^2} + \frac{\partial^2 \phi}{\partial y^2} = 0\,.
\tag{6.160}
$$

Since the cross-section of the waveguide in the (x, y) plane is a closed curve, at a fixed potential (since it is a conductor), we can deduce that ϕ is constant everywhere inside the conductor:

$$
0 = \int dxdy\, \phi \nabla_\perp^2 \phi = -\int dxdy\, |\vec{\nabla}_\perp \phi|^2\,,
\tag{6.161}
$$

which implies $\vec{\nabla}_\perp \phi = 0$ inside the waveguide, and hence $\phi = $ constant and so $\vec{E} = 0$. Similar considerations imply $\vec{B} = 0$ for the TEM mode also.[22]

---

[22]If the waveguide were replaced by coaxial conducting cylinders then TEM modes *could* exist in the gap between the innner and outer cylinder, since the potentials on the two cylinder need not be equal.

6.8.2 TE and TM modes

In order to have non-trivial modes propagating in the waveguide, we must relax the TEM assumption. There are two basic types of non-trivial modes we may consider, where either $\vec{E}$ or $\vec{B}$ (but not both) are taken to be transverse. These are called TE modes and TM modes respectively. A general electromagnetic wave propagating in a waveguide can be decomposed as a sum of such modes.

To analyse these modes, we first need to consider the boundary conditions at the conducting surface of the cylinder. The component of $\vec{E}$ parallel to the surface must vanish (seen by integrating $\vec{E}$ around a loop comprising a line segment just inside the waveguide, and closed by a line segment just inside the conductor, where $\vec{E} = 0$ by definition). Then, if we define $\vec{n}$ to be the unit normal vector at the surface, we may say that $\vec{n} \times \vec{E} = 0$. Next, taking the scalar product of $\vec{n}$ with the $\vec{\nabla} \times \vec{E} = i\,\omega\,\vec{B}$ Maxwell equation, we get

$$i\,\omega\,\vec{n} \cdot \vec{B} = \vec{n} \cdot (\vec{\nabla} \times \vec{E}) = -\vec{\nabla} \cdot (\vec{n} \times \vec{E}) = 0\,. \tag{6.162}$$

Thus, we have

$$\vec{n} \times \vec{E} = 0\,, \qquad \vec{n} \cdot \vec{B} = 0 \tag{6.163}$$

on the surface of the waveguide. We may restate these boundary conditions as

$$E_z\Big|_S = 0\,, \qquad \vec{n} \cdot \vec{B}_\perp\Big|_S = 0\,, \tag{6.164}$$

where S denotes the surface of the cylindrical waveguide.

The two boundary conditions above imply also that

$$\vec{n} \cdot \vec{\nabla}_\perp B_z\Big|_S = 0\,. \tag{6.165}$$

This follows by taking the scalar product of $\vec{n}$ with the penultimate equation in (6.156):

$$
\begin{aligned}
\vec{n} \cdot \vec{\nabla}_\perp B_z &= i\,k\,\vec{n} \cdot \vec{B}_\perp - i\,\omega\,\vec{n} \cdot (\vec{m} \times \vec{E}_\perp)\,, \\
&= i\,k\,\vec{n} \cdot \vec{B}_\perp + i\,\omega\,\vec{m} \cdot (\vec{n} \times \vec{E}_\perp)\,,
\end{aligned} \tag{6.166}
$$

and then restricting to the surface S of the cylinder. The condition (6.165) may be rewritten as

$$\frac{\partial B_z}{\partial n}\Big|_S = 0\,, \tag{6.167}$$

where $\partial/\partial n \equiv \vec{n} \cdot \vec{\nabla}$ is the normal derivative.

With the assumption (6.150), the wave equations for $\vec{E}$ and $\vec{B}$ (i.e. $(\nabla^2 - \partial^2/\partial t^2)\vec{E} = 0$ and $(\nabla^2 - \partial^2/\partial t^2)\vec{B} = 0$) become

$$\nabla_\perp^2 \vec{E} + (\omega^2 - k^2)\,\vec{E} = 0\,, \qquad \nabla_\perp^2 \vec{B} + (\omega^2 - k^2)\,\vec{B} = 0\,, \tag{6.168}$$

114

where $\nabla^2_\perp = \partial^2/\partial x^2 + \partial^2/\partial y^2$ is the 2-dimensional Laplacian. The four equations appearing in the first and third lines of (6.157) can be solved for E_x, E_y, B_x and B_y in terms of E_z and B_z, giving

$$
\begin{aligned}
E_x &= \frac{\mathrm{i}}{\omega^2 - k^2}\left(\omega\,\partial_y B_z + k\,\partial_x E_z\right), \\
E_y &= \frac{\mathrm{i}}{\omega^2 - k^2}\left(-\omega\,\partial_x B_z + k\,\partial_y E_z\right), \\
B_x &= \frac{\mathrm{i}}{\omega^2 - k^2}\left(-\omega\,\partial_y E_z + k\,\partial_x B_z\right), \\
B_y &= \frac{\mathrm{i}}{\omega^2 - k^2}\left(\omega\,\partial_x E_z + k\,\partial_y B_z\right).
\end{aligned}
\tag{6.169}
$$

This means that we can concentrate on solving for E_z and B_z, which must satisfy (see (6.168))

$$
\nabla^2_\perp E_z + (\omega^2 - k^2)\,E_z = 0\,, \quad \nabla^2_\perp B_z + (\omega^2 - k^2)\,B_z = 0\,.
\tag{6.170}
$$

Having solved for E_z and B_z, substitution into (6.169) gives the expressions for the remaining field components E_x, E_y, B_x and B_y.

As mentioned earlier, we can divide the discussion into a consideration of two different categories of wave solution in the waveguide. These are

$$
\textbf{TE waves}: \quad E_z = 0\,, \quad \text{and} \quad \left.\frac{\partial B_z}{\partial n}\right|_S = 0\,,
$$
$$
\vec{B}_\perp = \frac{\mathrm{i}\,k}{\omega^2 - k^2}\,\vec{\nabla}B_z\,, \qquad \vec{E} = \vec{E}_\perp = -\frac{\omega}{k}\,\vec{m}\times\vec{B}_\perp\,,
\tag{6.171}
$$

$$
\textbf{TM waves}: \quad B_z = 0\,, \quad \text{and} \quad \left.E_z\right|_S = 0\,,
$$
$$
\vec{E}_\perp = \frac{\mathrm{i}\,k}{\omega^2 - k^2}\,\vec{\nabla}E_z\,, \qquad \vec{B} = \vec{B}_\perp = \frac{\omega}{k}\,\vec{m}\times\vec{E}_\perp\,.
\tag{6.172}
$$

Note that the vanishing of E_z or B_z in the two cases means *by definition* that this field component vanishes *everywhere* inside the waveguide, and not just on the cylindrical conductor. Note also that the second condition in each case is just the residual content of the boundary conditions in (6.164) and (6.165), after having imposed the transversality condition $E_z = 0$ or $B_z = 0$ respectively. The second line in each of the TE and TM cases gives the results from (6.169), written now in a slightly more compact way. In each case, the basic wave solution is given by solving the 2-dimensional Helmholtz equation

$$
\frac{\partial^2\psi}{\partial x^2} + \frac{\partial^2\psi}{\partial y^2} + \Omega^2\psi = 0\,,
\tag{6.173}
$$

where

$$
\Omega^2 \equiv \omega^2 - k^2\,,
\tag{6.174}
$$

115

and ψ is equal to B_z or E_z in the case of TE or TM waves respectively. We also have the boundary conditions:

$$\textbf{TE waves}: \qquad \frac{\partial \psi}{\partial n}\bigg|_S = 0\,, \qquad \psi = B_z\,, \tag{6.175}$$

$$\textbf{TM waves}: \qquad \psi\bigg|_S = 0\,, \qquad \psi = E_z\,. \tag{6.176}$$

Equation (6.173), together with the boundary condition (6.175) or (6.176), defines an eigenfunction/eigenvalue problem. Since the the cross-section of the waveguide is a closed loop in the (x, y) plane, the equation (6.173) is to be solved in a compact closed region, and so the eigenvalue spectrum for Ω^2 will be discrete; there will be a semi-infinite number of eigenvalues, unbounded above, discretely separated from each other.

Notice that for each of the allowed eigenvalues Ω^2, there is a corresponding minimum angular frequency of wave that can propagate down the waveguide. This follows from eqn (6.174), which shows that we must have

$$\omega^2 \geq \Omega^2 \tag{6.177}$$

in order for k to be real. If ω^2 is less than Ω^2 then k is imaginary, which means that the "wave" dies off exponentially, like $e^{-|k|\,z}$, as it travels down the waveguide. Such waves are called *evanescent waves*.

Consider, as an example, TM waves propagating down a waveguide with rectangular cross-section:

$$0 \leq x \leq a\,, \qquad 0 \leq y \leq b\,. \tag{6.178}$$

For TM waves, we must satisfy the boundary condition that ψ vanishes on the edges of the rectangle. It follows from an elementary calculation, in which one separates variables in (6.173) by writing $\psi(x, y) = X(x)Y(y)$, that the eigenfunctions and eigenvalues, labelled by integers (m, n), are given by[23]

$$\begin{aligned} \psi_{mn} &= e_{mn} \sin\left(\frac{m\pi x}{a}\right) \sin\left(\frac{n\pi y}{b}\right)\,, \qquad m \geq 1\,, \quad n \geq 1\,, \\ \Omega_{mn}^2 &= \frac{m^2\pi^2}{a^2} + \frac{n^2\pi^2}{b^2}\,. \end{aligned} \tag{6.179}$$

The wave-number k and the angular frequency ω for the (m, n) mode are then related by

$$k^2 = \omega^2 - \Omega_{mn}^2\,. \tag{6.180}$$

---

[23]If we were instead solving for TE modes, we would have the boundary condition $\partial\psi/\partial n = 0$ on the edges of the rectangle, rather than $\psi = 0$ on the edges. This would give different eigenfunctions, involving cosines rather than sines.

The absolute lowest bound on the angular frequency that can propagate down the waveguide is clearly given, in this example, by $\Omega_{1,1}$. In other words, the lowest angular frequency of TM wave that can propagate down the rectangular waveguide is given by

$$\omega_{\min} = \pi\sqrt{\frac{1}{a^2} + \frac{1}{b^2}}\,. \tag{6.181}$$

In view of the relation (6.174) between the angular frequency and the wave-number, we see that the *phase velocity* v_{ph} and the *group velocity* v_{gr} are given by

$$
\begin{aligned}
v_{\mathrm{ph}} &\equiv \frac{\omega}{k} = \left(1 - \frac{\Omega^2}{\omega^2}\right)^{-1/2}, \\
v_{\mathrm{gr}} &\equiv \frac{d\omega}{dk} = \left(1 - \frac{\Omega^2}{\omega^2}\right)^{1/2}.
\end{aligned} \tag{6.182}
$$

Note that because of the particular form of the *dispersion relation*, i.e. the equation (6.174) relating ω to k, it is the case here that

$$v_{\mathrm{ph}}\, v_{\mathrm{gr}} = 1\,. \tag{6.183}$$

Since we must have $\omega \geq \Omega$ for wave propagation, we see that while the group velocity satisfies

$$v_{\mathrm{gr}} = \left(1 - \frac{\Omega^2}{\omega^2}\right)^{1/2} \leq 1\,, \tag{6.184}$$

the phase velocity satisfies

$$v_{\mathrm{ph}} = \left(1 - \frac{\Omega^2}{\omega^2}\right)^{-1/2} \geq 1\,. \tag{6.185}$$

There is nothing wrong with this, even though it means the phase velocity exceeds the speed of light, since nothing material, and no signal, is transferred faster than the speed of light. In fact, as we shall now verify, energy and information travel at the group velocity v_{gr}, which is always less than or equal to the speed of light.

Note that the group velocity approaches the speed of light (from below) as ω goes to infinity. To be more precise, the group velocity approaches the speed of light as ω becomes large compared to the eigenvalue Ω associated with the mode of propagation under discussion. An example where this limit is (easily) approached is if you look through a length of metal drainpipe. Electromagnetic waves in the visible spectrum have a frequency vastly greater than the lowest TM or TE modes of the drainpipe, and they propagate through the pipe as if it wasn't there. The story would be different if one tried to channel waves from the magnetron in a microwave oven down the drainpipe.

Let us now investigate the flow of energy down the waveguide. This is obtained by working out the time average of the Poynting flux,

$$\langle \vec{S} \rangle = \frac{1}{8\pi} \vec{E} \times \vec{B}^*\,. \tag{6.186}$$

117

Note that here the fields $\vec{E}$ and $\vec{B}$ are taken to be complex, and we are using the result discussed earlier about taking time averages of quadratic products of the physical $\vec{E}$ and $\vec{B}$ fields.

If we consider TM modes, then we shall have

$$
\begin{aligned}
\vec{E}_\perp &= \frac{\mathrm{i}\,k}{\Omega^2}\,\vec{\nabla}\psi\,, \qquad E_z = \psi\,, \\
\vec{B}_\perp &= \frac{\omega}{k}\,\vec{m}\times\vec{E}_\perp = \frac{\mathrm{i}\,\omega}{\Omega^2}\,\vec{m}\times\vec{\nabla}\psi\,, \qquad B_z = 0\,.
\end{aligned}
\tag{6.187}
$$

(Recall that $\vec{m} = (0,0,1)$.) Note that the expressions for $\vec{E}$ and $\vec{B}$ can be condensed down to

$$
\vec{E} = \frac{\mathrm{i}\,k}{\Omega^2}\,\vec{\nabla}\psi + \vec{m}\,\psi\,, \qquad \vec{B} = \frac{\mathrm{i}\,\omega}{\Omega^2}\,\vec{m}\times\vec{\nabla}\psi\,.
\tag{6.188}
$$

We therefore have

$$
\vec{E}\times\vec{B}^* = \left(\frac{\mathrm{i}\,k}{\Omega^2}\,\vec{\nabla}\psi + \vec{m}\,\psi\right)\times\left(-\frac{\mathrm{i}\,\omega}{\Omega^2}\,\vec{m}\times\vec{\nabla}\psi^*\right).
\tag{6.189}
$$

Using the vector identity $\vec{A}\times(\vec{B}\times\vec{C}) = (\vec{A}\cdot\vec{C})\,\vec{B} - (\vec{A}\cdot\vec{B})\,\vec{C}$, we then find

$$
\vec{E}\times\vec{B}^* = \frac{\omega k}{\Omega^4}\,(\vec{\nabla}\psi\cdot\vec{\nabla}\psi^*)\,\vec{m} + \frac{\mathrm{i}\,\omega}{\Omega^2}\,\psi\,\vec{\nabla}\psi^*\,,
\tag{6.190}
$$

since $\vec{m}\cdot\vec{\nabla}\psi = 0$. Along the z direction (i.e. along $\vec{m}$), we therefore have

$$
\langle S_z\rangle = \frac{\omega k}{8\pi\Omega^4}\,(\vec{\nabla}\psi\cdot\vec{\nabla}\psi^*) = \frac{\omega k}{8\pi\Omega^4}\,|\vec{\nabla}\psi|^2\,.
\tag{6.191}
$$

(The second term in (6.190) describes the circulation of energy within the cross-sectional plane of the waveguide. This can be seen from the fact that $\vec{\nabla}\psi(x,y)$ lies purely in the (x,y) plane, with no projection along the z direction.)

The total transmitted power P is obtained by integrating $\langle S_z\rangle$ over the cross-sectional area Σ of the waveguide. This gives

$$
\begin{aligned}
P &= \int_\Sigma dx\,dy\,\langle S_z\rangle = \frac{\omega k}{8\pi\Omega^4}\int_\Sigma dx\,dy\,\vec{\nabla}\psi^*\cdot\vec{\nabla}\psi\,, \\
&= \frac{\omega k}{8\pi\Omega^4}\int_\Sigma dx\,dy\,\left(\vec{\nabla}\cdot(\psi^*\,\vec{\nabla}\psi) - \psi^*\,\nabla^2\psi\right), \\
&= \frac{\omega k}{8\pi\Omega^4}\oint_C \psi^*\frac{\partial\psi}{\partial n}\,d\ell - \frac{\omega k}{8\pi\Omega^4}\int_\Sigma dx\,dy\,\psi^*\,\nabla^2\psi\,, \\
&= -\frac{\omega k}{8\pi\Omega^4}\int_\Sigma dx\,dy\,\psi^*\,\nabla^2\psi = \frac{\omega k}{8\pi\Omega^2}\int_\Sigma dx\,dy\,\psi^*\,\psi\,,
\end{aligned}
\tag{6.192}
$$

and so we have

$$
P = \frac{\omega k}{8\pi\Omega^2}\int_\Sigma dx\,dy\,|\psi|^2\,.
\tag{6.193}
$$

118

Note that in (6.192), the boundary term over the closed loop C that forms the boundary of the waveguide in the (x, y) plane gives zero because ψ vanishes everywhere on the cylinder. (Recall that we are considering the example of TM modes here.) The remaining term was then simplified by using (6.173).

We may also work out the total energy per unit length of the waveguide. The total time-averaged energy density is given by

$$
\begin{aligned}
\langle W \rangle &= \frac{1}{8\pi} \vec{E} \cdot \vec{E}^* = \frac{1}{8\pi} \left(\frac{\mathrm{i}\,k}{\Omega^2} \vec{\nabla}\psi + \vec{m}\,\psi \right) \cdot \left(-\frac{\mathrm{i}\,k}{\Omega^2} \vec{\nabla}\psi^* + \vec{m}\,\psi^* \right), \\
&= \frac{k^2}{8\pi\Omega^4} \vec{\nabla}\psi^* \cdot \vec{\nabla}\psi + \frac{1}{8\pi} \psi\psi^*.
\end{aligned}
\tag{6.194}
$$

The energy per unit length U is then obtained by integrating $\langle W \rangle$ over the cross-sectional area, which gives

$$
\begin{aligned}
U &= \int_{\Sigma} dx dy \, \langle W \rangle = \frac{k^2}{8\pi\Omega^4} \int_{\Sigma} dx dy \, \vec{\nabla}\psi^* \cdot \vec{\nabla}\psi + \frac{1}{8\pi} \int_{\Sigma} dx dy \, |\psi|^2, \\
&= \frac{k^2}{8\pi\Omega^2} \int_{\Sigma} dx dy \, |\psi|^2 + \frac{1}{8\pi} \int_{\Sigma} dx dy \, |\psi|^2,
\end{aligned}
\tag{6.195}
$$

where we have again integrated by parts in the first term, dropped the boundary term because ψ vanishes on the cylinder, and used (6.173) to simplify the result. Thus we find

$$
U = \frac{\omega^2}{8\pi\Omega^2} \int_{\Sigma} dx dy \, |\psi|^2.
\tag{6.196}
$$

Having obtained the expression (6.193) for the power P passing through the waveguide, and the expression (6.196) for the energy per unit length in the waveguide, we may note that

$$
P = \frac{k}{\omega} U = \frac{1}{v_{\mathrm{ph}}} U = v_{\mathrm{gr}} U.
\tag{6.197}
$$

This demonstrates that the energy flows down the waveguide at the group velocity v_{gr}.

Since the group velocity is given in terms of the eigenvalue Ω^2 of the particular mode of propagation, and the angular frequency ω of the wave by the expression in (6.182), this means that if a given frequency is able to propagate in more than one mode (with different values of Ω^2), then the travel time for the signal in the waveguide will be different depending upon which of the modes is excited. In practice, this means that the size and geometry of the waveguide cross-section is usually chosen so that the frequency that is to be propagated is only capable of exciting the lowest mode. Otherwise, if two or more modes can be excited, one is liable to end up with a jumbled-up mess coming out of the far end of the waveguide, with the same signal arriving multiple times as the faster or slower propagating versions in

different modes make it through to the far end. If we label the eigenvalues for the various modes by Ω_1^2, Ω_2^2, Ω_3^2, etc., and order them so that

$$\Omega_1^2 \leq \Omega_2^2 \leq \Omega_3^2 \leq \cdots, \tag{6.198}$$

then the size and geometry of the waveguide should be chosen so that the angular frequency ω that one wants to propagate down the waveguide is such that

$$\Omega_1^2 < \omega^2 < \Omega_2^2 \leq \Omega_3^2 \leq \cdots. \tag{6.199}$$

This then ensures that the signal can only propagate in what is called the *Dominant Mode*, namely the one with the lowest eigenvalue, Ω_1^2.

We gave an explicit example previously of a waveguide with a rectangular cross section. Other examples where one can solving explicitly for the eigenfunctions and eigenvalues are rather few and far between. A simple example that does admit explicit solutions is when the cross section is an isosceles right triangle, that is, a triangle with one internal angle being $90°$ and the other two each being $45°$. This case can be understood by first taking the square-section specialisation of the previous rectangular waveguide by setting $b = a$. Consider TM modes. The eigenvalues Ω_{mn}^2 in (6.179) now have the property that $\Omega_{mn}^2 = \Omega_{nm}^2$, and so this means that we can construct eigenfunctions of the form

$$\Psi_{mn} = \sin\left(\frac{m\pi x}{a}\right) \sin\left(\frac{n\pi y}{a}\right) - \sin\left(\frac{n\pi x}{a}\right) \sin\left(\frac{m\pi y}{a}\right) \tag{6.200}$$

which have the property of vanishing on the diagonal line $x = y$ as well as on the sides of the square. (The sum of any two eigenfunctions *with the same eigenvalue* is also an eigenfunction.) These are eigenfunctions of the Helmholtz equation

$$\nabla_\perp^2 \Psi_{mn} + \Omega_{mn}^2 \Psi_{mn} = 0, \qquad \Omega_{mn}^2 = \frac{\pi^2 (m^2 + n^2)}{a^2}. \tag{6.201}$$

Since these Ψ_{mn} obey the Dirichlet boundary condition $\Psi_{mn} = 0$ on the three sides $y = 0$; $x = a$; and $x = y$ of the isosceles right triangle, they provide the solutions for the TM modes in this triangular geometry. Note that it is essential, for this construction, that the sides a and b of the original rectangle be equal: The combinations $\psi_{mn} - \psi_{nm}$ of the original rectangular wavefunctions in (6.179) would certainly vanish on the diagonal $ay = bx$ even if $b \neq a$, but these would not be eigenfunctions, since ψ_{mn} and ψ_{nm} are eigenfunctons with different eigenvalues.

As far as I am aware, the only other case of a triangular-section waveguide that is explicitly solvable is that of an equilateral triangle. The solutions for the wavefunctions

are rather elegant symmetric functions with a discrete 120° rotational symmetry. (See homeworks for an example.) There is also a minor variant of this case, where the waveguide has a cross-section which is the triangle formed by bisecting the equilateral triangle with a line from one vertex to the midpoint of the opposite side.

One other explicitly solvable example arises when the cross-section of the waveguide is circular, with radius a. In this case one uses polar coordinates in the (x, y) plane, writing

$$x = r \cos\theta \,, \qquad y = r \sin\theta \,. \tag{6.202}$$

The Helmholtz equation (6.173) is now

$$\frac{1}{r} \frac{\partial}{\partial r} \Big(r \frac{\partial \psi}{\partial r} \Big) + \frac{1}{r^2} \frac{\partial^2 \psi}{\partial \theta^2} + \Omega^2 \, \psi = 0 \,. \tag{6.203}$$

This can be solved by looking for separated solutions of the form

$$\psi(r, \theta) = R(r) \, \Theta(\theta) \,, \tag{6.204}$$

leading to

$$r^2 \, R''(r) + r \, R'(r) + (\Omega^2 \, r^2 - \lambda) \, R(r) = 0 \,, \qquad \Theta''(\theta) + \lambda \, \Theta(\theta) = 0 \,. \tag{6.205}$$

The periodicity requirement $\Theta(\theta + 2\pi) = \Theta(\theta)$ for the angular functions means that the separation constant λ, which arises in the solutions $\Theta(\theta) = \exp(\mathrm{i} \sqrt{\lambda} \, \theta)$, must be of the form $\lambda = m^2$ where m is an integer. The radial equation is the Bessel equation, with solutions $J_m(\Omega \, r)$ and $Y_m(\Omega \, r)$. (See, for example, my PHYS 603 lecture notes for a discussion of some relevant properties of the Bessel functions.) Only the $J_m(\Omega \, r)$ Bessel functions provide acceptable solutions, since the $Y_m(\Omega \, r)$ Bessel functions are singular on the axis $r = 0$. Thus the wavefunctions for the circular waveguide have the form

$$\psi = c \, J_m(\Omega \, r) \, e^{\mathrm{i} \, m \, \theta} \,, \tag{6.206}$$

where c is a constant.

For TM modes, we require $\psi(a, \theta) = 0$, which implies that $\Omega \, a$ must be equal to one of the zeros of the J_m Bessel function. Thus we have

$$\Omega_{mn} = \frac{\alpha_{mn}}{a} \,, \qquad \psi_{mn} = c_{mn} \, J_m\Big(\frac{\alpha_{mn} \, r}{a}\Big) \, e^{\mathrm{i} \, m \, \theta} \,, \tag{6.207}$$

where α_{mn} is the n'th zero of J_m, so $J_m(\alpha_{mn}) = 0$.

For TE modes, where the wavefunction ψ must instead satisfy $\partial \psi / \partial n = 0$ at $r = a$, the eigenvalues Ω^2 and wavefunctions are now given by

$$\Omega_{mn} = \frac{\beta_{mn}}{a} \,, \qquad \psi_{mn} = c_{mn} \, J_m\Big(\frac{\beta_{mn} \, r}{a}\Big) \, e^{\mathrm{i} \, m \, \theta} \,, \tag{6.208}$$

where β_{mn} is the n'th zero of J'_m, the derivative of J_m. The values of the first few zeros of $J_m(x)$ and $J'_m(x)$ are listed in the table below, for $J_m(x)$ with $m = 0$, 1, 2 and 3.

$n =$	1	2	3	4
$\alpha_{0n} =$	2.405	5.520	8.654	11.79
$\alpha_{1n} =$	3.832	7.016	10.17	13.32
$\alpha_{2n} =$	5.136	8.417	11.62	14.80
$\alpha_{3n} =$	6.380	9.761	13.02	16.22

$n =$	1	2	3	4
$\beta_{0n} =$	3.832	7.016	10.174	13.32
$\beta_{1n} =$	1.841	5.331	8.536	11.71
$\beta_{2n} =$	3.054	6.706	9.969	13.17
$\beta_{3n} =$	4.201	8.015	14.586	17.789

The values for the α_{mn} and β_{mn} increase monotonically, both horizontally and vertically, if the tables are extended beyond the m and n ranges listed here. In those cases where the same approximate value is appearing in both tables, the corresponding α and β are in fact exactly equal (this is because $J'_0(x) = -J_1(x)$).

The plots of the first few $J_m(x)$ Bessel functions are displayed in Figure 1 below.
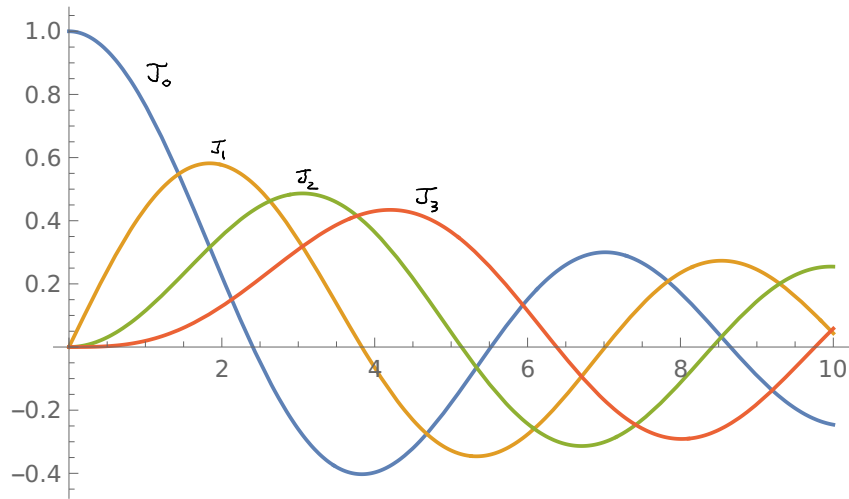
Figure 1: The Bessel functions $J_0(x)$, $J_1(x)$, $J_2(x)$ and $J_3(x)$.

The first few $Y_m(x)$ Bessel functions are shown in Figure 2. They are all singular at $x = 0$, so as mentioned above, they would not play a role in constructing the Helmholtz eigenfunctions in a hollow cylindrical waveguide. However, if one were instead wanting to construct the eigenfunctions in a *coaxial* waveguide, where the waves travel in the hollow region between two concentric cylinders of radii a and b, then both the J_m and Y_m Bessel

functions would be involved. There would now be two boundary conditions to be satisfied, at $r = a$ (the radius of the inner cylinder) and at $r = b$ (the radius of the outer cylinder). The Y_m and the J_m Bessel functions are all non-singular in this region between the two cyclinders; hence both sets now enter in the solutions for the wavefunctions.

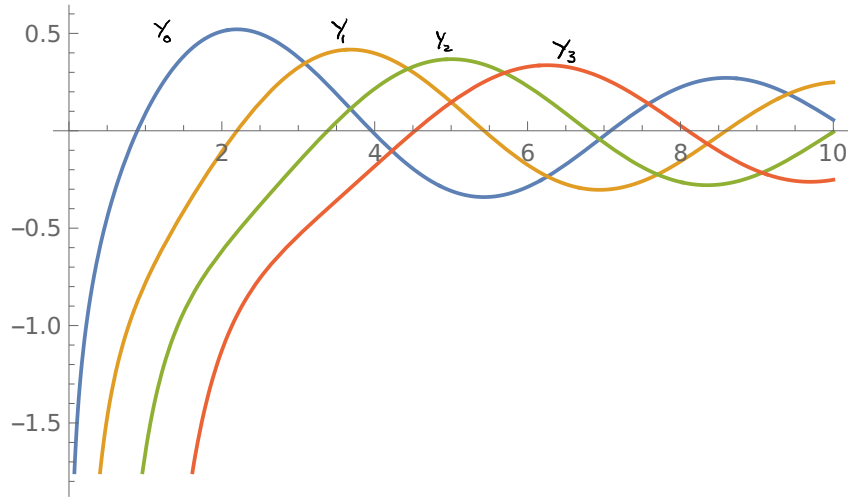Figure 2: The Bessel functions $Y_0(x)$, $Y_1(x)$, $Y_2(x)$ and $Y_3(x)$.

6.9   Resonant cavities

A resonant cavity is a hollow, closed conducting "container," inside which is an electromagnetic wave. A simple example would be to take a length of waveguide of the sort we have considered in section 6.8, and turn it into a closed cavity by attaching conducting plates at each end of the cylinder. Let us suppose that the length of the cavity is d.

Consider, as an example, TM modes in the cavity. We solve the same 2-dimensional Helmholtz equation (6.173) as before,

$$\frac{\partial^2 \psi}{\partial x^2} + \frac{\partial^2 \psi}{\partial y^2} + \Omega^2 \, \psi = 0 \,, \tag{6.209}$$

subject again to the TM boundary condition that ψ must vanish on the surface of the cylinder. The $\vec{E}$ and $\vec{B}$ fields are given, as before, by

$$\begin{aligned}
\vec{E}_\perp &= \frac{\mathrm{i}\,k}{\Omega^2} \, e^{\mathrm{i}\,(\kappa z - \omega t)} \, \vec{\nabla}\psi \,, \qquad E_z = \psi \, e^{\mathrm{i}\,(\kappa z - \omega t)} \,, \\
\vec{B}_\perp &= \frac{\omega}{k} \, \vec{m} \times \vec{E}_\perp \,,
\end{aligned} \tag{6.210}$$

123

where $\vec{m} = (0, 0, 1)$. Now, however, we have the additional boundary conditions that $\vec{E}_\perp$ must vanish on the two conducting plates, which we shall take to be at $z = 0$ and $z = d$. This is because the component of $\vec{E}$ parallel to a conductor must vanish at the conducting surface.

In order to arrange that $\vec{E}_\perp$ vanish, for all t, at $z = 0$ and $z = d$, it must be that there is a superposition of right-moving and left-moving waves. (These correspond to z and t dependences $e^{\mathrm{i}(\pm\kappa z - \omega t)}$ respectively.) Thus we need to take the combination that makes a standing wave,

$$\vec{E}_\perp = -\frac{k}{\Omega^2} \sin kz\, e^{-\mathrm{i}\omega t}\, \vec{\nabla}\psi\,, \tag{6.211}$$

in order to have $\vec{E}_\perp = 0$ at $z = 0$. Furthermore, in order to have also that $\vec{E}_\perp = 0$ at $z = d$, it must be that the wave-number k is now quantised, according to

$$k = \frac{p\pi}{d}\,, \tag{6.212}$$

where p is an integer. Note that we also have

$$E_z = \psi \cos kz\, e^{-\mathrm{i}\omega t}\,. \tag{6.213}$$

Recall that in the waveguide, we had already found that $\Omega^2 \equiv \omega^2 - k^2$ was quantised, being restricted to a semi-infinite discrete set of eigenvalues for the 2-dimensional Helmoholtz equation. This still allowed k and ω to take continuous values, subject to the constraint (dispersion relation)

$$\omega^2 = \Omega^2 + k^2\,. \tag{6.214}$$

In the resonant cavity we now have the further restriction that k is quantised, according to (6.212). This means that the spectrum of allowed frequencies ω is now discrete, and given by

$$\omega^2 = \Omega^2 + \frac{p^2\pi^2}{d^2}\,. \tag{6.215}$$

If, for example, we consider the previous example of TM modes in a rectangular waveguide whose cross-section has sides of lengths a and b, but now with the added end-caps at $z = 0$ and $z = d$, then Ω^2 is given by (6.179), and so the resonant frequencies in the cavity are given by

$$\omega^2 = \pi^2 \Big(\frac{m^2}{a^2} + \frac{n^2}{b^2} + \frac{p^2}{d^2} \Big)\,, \tag{6.216}$$

for positive integers (m, n, p).

A cylindrical resonant cavity with circular cross-section can be analysed in an analogous way, starting from the circular cross-section wavguide discussed earlier, and capping it off at $z = 0$ and $z = d$ with conducting end-plates.

One can also look at resonant cavities that are not simply obtained by capping off a length of waveguide. A classic example is a spherical resonator, formed by the interior of a hollow conducting sphere. Simple examples of resonant modes in the spherical cavity are explored in one of the forthcoming homework problems. Variants of the spherical cavity, such as a cavity formed between two conducting spherical shells, are also of interest physically. Such a set-up provides a model for the behaviour of radio waves trapped between the surface of the earth and the ionosphere, and it gives rise to the intriguing *Schumann Resonances*, which are extremely low frequency radio waves that encircle the earth.

7 Fields Due to Accelerating Charges

7.1 Retarded potentials

If we solve the Bianchi identity by writing $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$, the remaining Maxwell equation (i.e. the field equation)

$$\partial_\mu F^{\mu\nu} = -4\pi J^\nu \tag{7.1}$$

becomes

$$\partial_\mu \partial^\mu A^\nu - \partial_\mu \partial^\nu A^\mu = -4\pi J^\nu \,. \tag{7.2}$$

If we choose to work in the Lorenz gauge,

$$\partial_\mu A^\mu = 0 \,, \tag{7.3}$$

then (7.2) becomes simply

$$\Box A^\mu = -4\pi J^\mu \,. \tag{7.4}$$

Since $A^\mu = (\phi, \vec{A})$ and $J^\mu = (\rho, \vec{J})$, this means we shall have

$$\Box \phi = -4\pi \, \rho \,, \qquad \Box \vec{A} = -4\pi \, \vec{J}, \tag{7.5}$$

or, in the three-dimensional language,

$$\nabla^2 \phi - \frac{\partial^2 \phi}{\partial t^2} = -4\pi \, \rho \,, \qquad \nabla^2 \vec{A} - \frac{\partial^2 \vec{A}}{\partial t^2} = -4\pi \, \vec{J}. \tag{7.6}$$

(We are assuming here, when writing the equation for $|vac\rangle A$, that Cartesian coordinates and vectors are being used.)

In general, we can write the solutions to (7.6) as the sums of a *particular integral* of the inhomogeneous equation (i.e. the one with the source term on the right-hand side) plus the general solution of the homogeneous equation (the one with the right-hand side set to

zero). Our interest now will be in finding the particular integral. Solving this problem in the case of static sources and fields will be very familiar from electrostatics and magnetostatics. Now, however, we wish to solve for the particular integral in the case where there is time dependence too. Consider the equation for ϕ first.

First consider the situation where there is just an infinitesimal amount of charge $\delta e(t)$ in an infinitesimal volume located at the origin. Note that we are allowing it to be time dependent. Of course it would not be possible to have a single element of charge in isolation that varied with time, since this would be forbidden by charge conservation. However, one should think of this charge element as being part of a charge distribution, whose *total* charge remains fixed but where the individual charge elements can vary within it as a function of time. We shall superpose the contributions from the distinct charge elements later on in the calculation. The charge density for our time-dependent charge $\delta e(t)$ at the origin will be

$$\rho(\vec{r}, t)) = \delta e(t)\, \delta^3(\vec{r})\,, \tag{7.7}$$

where $\vec{r}$ is the position vector from the origin to the observation point. We therefore wish to solve

$$\nabla^2 \phi - \frac{\partial^2 \phi}{\partial t^2} = -4\pi\, \delta e(t)\, \delta^3(\vec{r})\,. \tag{7.8}$$

(It would probably be better, logically speaking, to use $\delta\phi$ rather than ϕ here, since we are working for now with the infinitesimal potential due to the infinitesimal charge element δe. It would be a bit cumbersome to use $\delta\phi$ in all the formulae below, so we shall just use ϕ, hoping that it is clear what is intended.) When $\vec{r} \neq 0$, we have simply $\nabla^2 \phi - \frac{\partial^2 \phi}{\partial t^2} = 0$.

Clearly, because of the spherical symmetry around the origin, ϕ depends on $\vec{r}$ only through its magnitude $r \equiv |\vec{r}|$, and so $\phi = \phi(r, t)$. Now, with $\vec{r} = (x_1, x_2, x_3)$, we have $r^2 = x_i x_i$ and so $\partial_i r = x_i/r$. Consequently, we shall have

$$\partial_i \phi = \frac{x_i}{r}\, \phi'\,, \tag{7.9}$$

where $\phi' \equiv \partial\phi/\partial r$, and then

$$\nabla^2 \phi = \partial_i \partial_i \phi = \phi'' + \frac{2}{r}\, \phi'\,. \tag{7.10}$$

Letting $\Phi = r\, \phi$, we have

$$\phi' = \frac{1}{r}\, \Phi' - \frac{1}{r^2}\, \Phi\,, \qquad \phi'' = \frac{1}{r}\, \Phi'' - \frac{2}{r^2}\, \Phi' + \frac{2}{r^3}\, \Phi\,. \tag{7.11}$$

This means that for $\vec{r} \neq 0$, we shall have

$$\frac{\partial^2 \Phi}{\partial r^2} - \frac{\partial^2 \Phi}{\partial t^2} = 0\,, \qquad \text{i.e.} \quad \left(\frac{\partial}{\partial t} - \frac{\partial}{\partial r}\right)\left(\frac{\partial}{\partial t} + \frac{\partial}{\partial r}\right)\Phi = 0\,. \tag{7.12}$$

126

The general solution to this equation is

$$\Phi(r, t) = f_1(t - r) + f_2(t + r),\tag{7.13}$$

where f_1 and f_2 are arbitrary functions.

The solution with f_1 is called the *retarded* solution, and the solution with f_2 is called the *advanced* solution. The reason for this terminology is that in the retarded solution, the "effect" occurs after the "cause," in the sense that the profile of the function f_1 propagates *outwards* from the origin where the charge $de(t)$ is located. By contrast, in the advanced solution the effect precedes the cause; the disturbance propagates inwards as time increases. The advanced solution is *acausal*, and therefore unphysical, and so we shall keep only the causal solution, i.e. the retarded solution. The upshot is that for $r \neq 0$, the solution is

$$\phi(r, t) = \frac{1}{r} \Phi(t - r),\tag{7.14}$$

where Φ is an as-yet arbitrary function of its argument, which we shall determine shortly.

We clearly expect that ϕ will go to infinity as r approaches zero, since the charge (albeit infinitesimal) is located there. Consequently, it will be the case that the derivatives $\partial/\partial r$ will dominate over the time derivatives $\partial/\partial t$ near to $r = 0$, and so in that region we can write

$$\nabla^2 \phi \approx -4\pi \delta e(t)\, \delta^3(\vec{r})\,.\tag{7.15}$$

This therefore has the usual solution that is familiar from electrostatics, namely

$$\phi \approx \frac{\delta e(t)}{r}\,,\tag{7.16}$$

or, in other words,

$$\Phi(t - r) \approx \delta e(t)\tag{7.17}$$

when r is negligibly small. In the limit as $r \to 0$ we therefore have the exact statement $\Phi(t) = \delta e(t)$. Thus the function Φ is fully, and exactly, determined. Since Φ is already established, in general, to depend on t and r only through the combination $t - r$, we can therefore immediately write down the exact solution valid for all r, namely

$$\Phi(t - r) = \delta e(t - r)\,.\tag{7.18}$$

From (7.14), we therefore have that

$$\phi(\vec{r}, t) = \frac{\delta e(t - r)}{r}\,.\tag{7.19}$$

127

At this point, having completed the discussion of the potential due to the infinitesimal charge element $\delta e(t)$ at the origin, it might be helpful to rewrite (7.19) as

$$\delta\phi(\vec{r},t) = \frac{\delta e(t-r)}{r}\,, \tag{7.20}$$

to emphasise that so far we have found the infinitesimal contribution to the total potential ϕ coming just from $\delta e(t)$ at the origin.

This solution (7.20) is valid for the particular case of an infinitesimal charge $\delta e(t)$ located at $r = 0$. For a general time-dependent charge distribution $\rho(\vec{r},t)$, we just exploit the linearity of the Maxwell equations and sum up the contributions from all the charges in the distribution. This therefore gives

$$\phi(\vec{r},t) = \int \frac{\rho(\vec{r}',t-R)}{R}\,d^3\vec{r}'\,, \tag{7.21}$$

where $\vec{R} \equiv \vec{r} - \vec{r}'$ and $R = |\vec{R}| = |\vec{r} - \vec{r}'|$. We leave it as an elementary exercise to check that if one just takes $\rho(\vec{r},t) = \delta e(t)\,\delta(\vec{r})$ as in (7.7), eqn (7.21) gives back (7.19).

The solution (7.21) of the inhomogeneous equation is the one that is "forced" by the source term, in the sense that it vanishes if the source charge density ρ vanishes. The general solution is given by this particular integral plus an arbitrary solution of the homogeneous equation $\Box\phi = 0$. The solution (7.21) can be written as

$$\phi(\vec{r},t) = \int \frac{\rho(\vec{r}',t-|\vec{r}-\vec{r}'|)}{|\vec{r}-\vec{r}'|}\,d^3\vec{r}'\,. \tag{7.22}$$

In an identical fashion, we can see that the solution for the 3-vector potential $\vec{A}$ in the presence of a 3-vector current source $\vec{J}(\vec{r},t)$ will be

$$\vec{A}(\vec{r},t) = \int \frac{\vec{J}(\vec{r}',t-|\vec{r}-\vec{r}'|)}{|\vec{r}-\vec{r}'|}\,d^3\vec{r}'\,, \tag{7.23}$$

where we are assuming that Cartesian coordinates and vectors are being used.

The solutions for $\phi(\vec{r},t)$ and $\vec{A}(\vec{r},t)$ that we have obtained here are called the *Retarded Potentials*. The analogous "advanced potentials" would correspond to having $t + |\vec{r} - \vec{r}'|$ instead of $t - |\vec{r} - \vec{r}'|$ as the time argument of the charge and current densities inside the integrals. It is clear that the retarded potentials are the physically sensible ones, in that the potentials at the present time t depend upon the charge and current densities at times $\leq t$. In the advanced potentials, by contrast, the potentials at the current time t would be influenced by what the charge and current densities *will be* in the future. This would be unphysical, since it would violate causality.

Since the procedure by which we arrived at the retarded potential solutions(7.22) and (7.23) may have seemed slightly "handwaving," it is perhaps worthwhile to go back and check that they are indeed correct. This can be done straightforwardly, simply by substituting them into the original wave equations (7.6). One finds that they do indeed yield exact solutions of the equations. We leave this as an exercise for the reader.

7.2 Liénard-Wiechert potentials

We now turn to a discussion of the electromagnetic fields produced by a point charge e moving along an arbitrary path $\vec{r} = \vec{r}_0(t)$. We already considered a special case of this in section 5.3, where we worked out the fields produced by a charge in uniform motion (i.e. moving at constant velocity). In that case, we could work out the electromagnetic fields by using the trick of transforming to the Lorentz frame in which the particle was at rest, doing the very simple calculation of the fields in that frame, and then transforming back to the frame where the particle was in uniform motion.

Now, we are going to study the more general case where the particle can be accelerating; i.e. where its velocity is not uniform. This means that there does not exist an inertial frame in which the particle is at rest for all time, and so we cannot use the previous trick.

It is worth emphasising that even though the particle is accelerating, this does not mean that we cannot solve the problem using special relativity. The point is that we shall only ever study the fields from the viewpoint of an observer who *is* in an inertial frame, and so for this observer, the laws of special relativity apply. Only if we wanted to study the problem from the viewpoint of an observer in an *accelerating* frame, such as the rest-frame of the particle, would we need to use the laws of general relativity.

One way to obtain the potentials due to the moving charge is to make use of the expressions for the retarded potentials that we derived in the previous subsection, substituting in the appropriate functions for the charge density and current density of the point charge. We shall carry out that calculation shortly. Since this is essentially just an exercise in mathematics we shall first present a different derivation, based more on the application of physical principles. This derivation highlights an interesting point about causality.

Note that although we cannot use special relativity to study the problem in the *rest frame* of the accelerating particle, we can, and sometimes will, make use of an *instantaneous rest frame*. This is an inertial frame whose velocity just happens to match exactly the velocity of the particle at a particular instant of time. Since the particle is accelerating, then a moment later the particle will no longer be at rest in this frame. We could, if we wished,

then choose an "updated" instantaneous rest frame, and use special relativity to study the problem (for an instant) in the new inertial frame. We shall find it expedient at times to make use of the concept of an instantaneous rest frame, in order to simply intermediate calculations. Ultimately, of course, we do not want to restrict ourselves to having to hop onto a new instantaneous rest frame every time we discuss the problem, and so the goal is to obtain results that are valid in *any* inertial frame.

Now, on with the problem. We can expect, on the grounds of causality, that the electromagnetic fields we observe at the spacetime point P specified by the cordinates $(\vec{r}, t)$ should depend only on the the position and state of motion of the particle at earlier points in its path that are causally connected to the observation point. That is to say, that only those points on the path for which information could reach the observation point P by travelling at less than or equal to the speed of light should influence the fields observed at $(\vec{r}, t)$.

In fact, it turns out that a much stronger statement is true: for each spacetime observation point $(\vec{r}, t)$ there is only one point on the particle's path that influences the observed fields, namely the point whose distance from P is such that a light beam emitted at the so-called "retarded time" time t' when the particle was there reaches the observation point exactly at the observation time t. In other words, the information governing the fields at the observation point P propagates there at *exactly* the speed of light; no faster (obviously), but also no slower. This feature can in fact be seen from the form of the retarded potential derived previously, as we shall see later. It is worth remarking that it is a feature that depends upon the spacetime dimension. In particular, and by contrast, in an *odd* spacetime dimension the potential due to a moving charge depends upon the entire causally-connected past history of the particle's motion.

We begin by defining the *retarded time t'*, as measured in the chosen inertial frame. It is useful to define

$$\vec{R}(t') \equiv \vec{r} - \vec{r}_0(t') \, . \tag{7.24}$$

This is the radius vector from the location $\vec{r}_0(t')$ of the charge at the time t' to the observation point r. The time t' is then determined by

$$t - t' = R(t') \, , \qquad \text{where} \qquad R(t') = |\vec{R}(t')| \, . \tag{7.25}$$

It can be shown that there is one solution for t', for each choice of t, provided that the particle is moving (as it must) with a velocity $\vec{v}(t') = d\vec{r}_0(t')/dt'$ that is less than the speed of light, i.e. $|\vec{v}| < 1$.

In the Lorentz frame where the particle is at rest at the particular instant t' (it would be

more precise to say that our observer at $(\vec{r}, t)$ is using a Lorentz frame in which the particle *was* at rest at the retarded time $t' = t - R(t')$), the potential at $\vec{r}$ at time t will, according to the claim above, be given by

$$\phi = \frac{e}{R(t')}, \qquad \vec{A} = 0.$$ (7.26)

We can determine the 4-vector potential A^μ in an arbitrary Lorentz frame simply by inventing a 4-vector expression that reduces to (7.26) under the specialisation that the velocity $\vec{v} \equiv d\vec{r}'/dt'$ of the charge is zero at time t'. (This is just a quick way of getting to the answer that we could alternatively and equivalently work out by explicitly performing the Lorentz transformation to the new inertial frame.)

Let the 4-velocity of the charge, in the observer's inertial frame, be U^μ. If the charge is at rest, its 4-velocity will be

$$U^\mu = (1, \vec{0}).$$ (7.27)

Thus to write a 4-vector expression for $A^\mu = (\phi, \vec{A})$ that reduces to (7.26) if U^μ is given by (7.27), we just have to find a scalar f such that

$$A^\mu = f \, U^\mu,$$ (7.28)

with f becoming $e/R(t')$ in the special case above. Let us define the 4-vector

$$R^\mu = \left(t - t', \vec{r} - \vec{r}_0(t') \right) = \left(t - t', \vec{R}(t') \right).$$ (7.29)

(This is clearly a 4-vector, because $(t, \vec{r})$ is a 4-vector, and $(t', \vec{r}_0(t'))$, the spacetime coordinates of the particle, is a 4-vector.) Then, we can write f as the scalar

$$f = \frac{e}{(-U^\nu \, R_\nu)}, \qquad \text{and so} \qquad A^\mu = -\frac{e U^\mu}{(U^\nu \, R_\nu)},$$ (7.30)

since clearly if U^μ is given by (7.27), we shall have $-U^\nu \, R_\nu = -R_0 = R^0 = t - t' = R(t')$.

Having written A^μ as a 4-vector expression that reduces to (7.26) under the specialisation (7.27), we know that it must be the correct expression in any Lorentz frame. Now, we have

$$U^\mu = (\gamma, \gamma \, \vec{v}), \qquad \text{where} \qquad \gamma = \frac{1}{\sqrt{1 - v^2}},$$ (7.31)

and so we see that

$$
\begin{aligned}
\phi(\vec{r}, t) &= A^0 = \frac{e\gamma}{(t - t')\gamma - \gamma \, \vec{v} \cdot \vec{R}} = \frac{e}{t - t' - \vec{v} \cdot \vec{R}} = \frac{e}{R - \vec{v} \cdot \vec{R}}, \\
\vec{A}(\vec{r}, t) &= \frac{e\gamma \, \vec{v}}{(t - t')\gamma - \gamma \, \vec{v} \cdot \vec{R}} = \frac{e \, \vec{v}}{R - \vec{v} \cdot \vec{R}} = \phi(\vec{r}, t) \, \vec{v}.
\end{aligned}
$$ (7.32)

To summarise, we have concluded that the gauge potentials for a charge e moving along the path $\vec{r} = \vec{r}_0(t')$, as seen from the point $\vec{r}$ at time t, are given by

$$\phi(\vec{r}, t) = \frac{e}{R - \vec{v} \cdot \vec{R}}, \qquad \vec{A}(\vec{r}, t) = \frac{e\,\vec{v}}{R - \vec{v} \cdot \vec{R}}, \tag{7.33}$$

where all quantities on the right-hand sides are evaluated at the time t', i.e. $\vec{R}$ means $\vec{R}(t')$ and $\vec{v}$ means $d\vec{r}_0(t')/dt'$, with

$$\vec{R}(t') = \vec{r} - \vec{r}_0(t'), \tag{7.34}$$

and t' is determined by solving the equation

$$R(t') = t - t', \qquad \text{where} \qquad R(t') \equiv |\vec{R}(t')|. \tag{7.35}$$

These potentials are known as the Liénard-Wiechert potentials.

The next step will be to calculate the electric and magnetic fields from the Liénard-Wiechert potentials. However, before doing so, it is worthwhile to pause to justify the argument we used above, in which we assumed that the potentials at the observation point $\vec{r}$ at time t are determined by the state of the particle at the earlier point on its path that was exactly a light travel time earlier. As mentioned previously, causality be used to argue that the potentials at $(\vec{r}, t)$ could, *a priori*, depend on the entire past history of the particle that is in causal contact with the spacetime point $(\vec{r}, t)$. Thus, although one can certainly say that any part of its history that lies "outside the light cone" cannot affect the potentials at $(\vec{r}, t)$, one cannot say, based on causality alone, that only the instant when the particle was a light-travel distance away could be relevant for determining the potentials at $(\vec{r}, t)$. As a matter of fact if one repeats the analogous calculation in a general D-dimensional spacetime, then it turns out that only when D is *even* (such as our familiar 4-dimensional world) does the effect propagate only "on the light-cone." In an odd spacetime dimension, the analogous Liénard-Wiechert potentials really *do* depend on the entire previous causally-connected history of the particle.

A complete derivation of the Liénard-Wiechert potentials can be given as follows. We take as the starting point the expressions (7.22) and (7.23) for the retarded potentials due to a time-dependent charge and current source. These expressions can themselves be regarded as solid and rigorous, since one only has to verify by direct substitution into (7.6) that they are indeed correct. Consider first the retarded potential for ϕ, given in (7.22). We can rewrite this as a 4-dimensional integral by introducing a delta-function in the time variable, so that

$$\phi(\vec{r}, t) = \int \int \frac{\rho(\vec{r}', t'')}{|\vec{r} - \vec{r}'|}\, \delta(t'' - t + |\vec{r} - \vec{r}'|)\, dt''\, d^3\vec{r}'. \tag{7.36}$$

If you carry out the t'' integration you immediately get back to the expression (7.22); instead, we shall be performing the three spatial integrations first.

The charge density for a point charge e moving along the path $\vec{r} = \vec{r}_0(t)$ is given by

$$\rho(\vec{r}, t) = e\,\delta^3(\vec{r} - \vec{r}_0(t))\,. \tag{7.37}$$

This means that we shall have

$$\phi(\vec{r}, t) = \int \int \frac{e\,\delta^3(\vec{r}' - \vec{r}_0(t''))}{|\vec{r} - \vec{r}'|}\,\delta(t'' - t + |\vec{r} - \vec{r}'|)\,dt''\,d^3\vec{r}'\,. \tag{7.38}$$

We now make use of the 3-dimensional spatial delta function in order to perform the spatial integrations, giving

$$\phi(\vec{r}, t) = \int \frac{e}{|\vec{r} - \vec{r}_0(t'')|}\,\delta(t'' - t + |\vec{r} - \vec{r}_0(t'')|)\,dt''\,. \tag{7.39}$$

To evaluate the time integral, we need to make use of a basic result about the Dirac delta-function, namely that if a function $f(x)$ has as zero at $x = x_0$, then[24]

$$\delta(f(x)) = \delta(x - x_0)\left|\frac{df}{dx}\right|_{x=x_0}^{-1}\,, \tag{7.40}$$

where df/dx is evaluated at $x = x_0$.[25]

To evaluate (7.39), we note that

$$\begin{aligned}
\frac{\partial}{\partial t''}\left(t'' - t + |\vec{r} - \vec{r}_0(t'')|\right) &= 1 + \frac{\partial}{\partial t''}\left((\vec{r} - \vec{r}_0(t'')) \cdot (\vec{r} - \vec{r}_0(t''))\right)^{1/2}\,, \\
&= 1 + \left((\vec{r} - \vec{r}_0(t'')) \cdot (\vec{r} - \vec{r}_0(t''))\right)^{-1/2}(\vec{r} - \vec{r}_0(t'')) \cdot \frac{\partial(-\vec{r}_0(t''))}{\partial t''}\,, \\
&= 1 - \frac{\vec{v} \cdot (\vec{r} - \vec{r}_0(t''))}{|\vec{r} - \vec{r}_0(t'')|}\,, \\
&= 1 - \frac{\vec{v} \cdot \vec{R}(t'')}{R(t'')}\,, \tag{7.41}
\end{aligned}$$

---

[24]To prove this, consider the integral $I = \int dx h(x)\delta(f(x))$ for an arbitrary function $h(x)$. Next, change variable to $z = f(x)$, so $dx = dz/(df/dx)$. Then we have

$$I = \int dz h(x)\frac{\delta(z)}{|df/dx|} = h(x_0)/|df/dx|\int \delta(z)dz = h(x_0)/|df/dx|\,,$$

where df/dx is evaluated at $x = x_0$. Thus we have

$$I = h(x_0)/|df/dx|_{x_0} = \int dx h(x)\frac{\delta(x - x_0)}{|df/dx|_{x_0}}\,,$$

which proves (7.40). (The reason for the absolute-value on $|df/dz|$ is that it is to be understood that the direction of the limits of the z integration should be the standard one (negative to positive). If the gradient of f is negative at $x = x_0$ then one has to insert a minus sign to achieve this. This is therefore handled by the absolute-value sign.)

[25]The result given here is valid if $f(x)$ vanishes only at the point $x = x_0$. If $f(x)$ has more than one zero, then there will be a sum of terms of the type given in (7.40). In our case, there is in fact only one zero, as was mentioned earlier when it was stated that once the path of the particle is specified, there exists a unique retarded time t' associated with every spacetime observation point $(\vec{r}, t)$.

where $\vec{v} = d\vec{r}_0(t'')/dt''$. Following the rule (7.40) for handling a "delta-function of a function," we therefore take the function in the integrand of (7.39) that multiplies the delta-function, evaluate it at the time t for which the argument of the delta-function vanishes, and divide by the absolute value of the derivative of the argument of the delta-function. This therefore gives

$$\phi(\vec{r}, t) = \frac{e}{R(t') - \vec{v} \cdot \vec{R}(t')} \,, \tag{7.42}$$

where t' is the solution of $t - t' = R(t')$, and so we have reproduced the previous expression for the Liénard-Wiechert potential for ϕ in (7.33). The derivation for $\vec{A}$ is very similar.

As mentioned previously, in higher *even-dimensional* spacetimes the analogous Liénard-Wiechert potentials again depend only on what was happening to the charged particle at the retarded time t'. However, the form of the potentials is more complicated in higher even-dimensional spacetimes, and instead of depending only on the position and velocity of the particle, as in eqns (7.33), they now depend also on higher derivatives of the velocity. For example, as is explored in homework 11, in six spacetime dimensions the Liénard-Wiechert potential $\phi(\vec{r}, t)$ takes the form

$$\phi(\vec{r}, t) = \frac{e\,(1 - v^2)}{\left[R(t') - \vec{v} \cdot \vec{R}(t')\right]^3} + \frac{e\,\vec{a} \cdot \vec{R}(t')}{\left[R(t') - \vec{v} \cdot \vec{R}(t')\right]^3} \,, \tag{7.43}$$

where the quantities t', $\vec{R}(t')$, $R(t')$ and $\vec{v}$ have the same meanings as in 4 spacetime dimensions except that now $\vec{R}(t')$ and $\vec{v}$ are vectors in the 5 spatial dimensions instead of 3. The acceleration $\vec{a} = d\vec{v}(t')/dt'$ of the particle also enters in the expression for the Liénard-Wiechert potential.

An example of a retarded potential in an odd-dimensional spacetime arises in homework 10, where the potential due to an instantaneous charge at the origin in $3 = 2 + 1$ spacetime dimensions is constructed. It can be seen that the potential resulting from this pulse of charge lingers indefinitely, gradually decaying but remaining non-zero at all finite future times.

7.3 Electric and magnetic fields of a moving charge

Having obtained the Liénard-Wiechert potentials ϕ and $\vec{A}$ of a moving charge, the next step is to calculate the associated electric and magnetic fields,

$$\vec{E} = -\vec{\nabla}\phi - \frac{\partial \vec{A}}{\partial t} \,, \qquad \vec{B} = \vec{\nabla} \times \vec{A} \,. \tag{7.44}$$

To do this, we shall need the following results. First, we note that

$$\frac{\partial R}{\partial t} = \frac{\partial R}{\partial t'} \frac{\partial t'}{\partial t} \,, \tag{7.45}$$

and so, since $R^2 = R_i R_i$ we have

$$\frac{\partial R}{\partial t'} = \frac{R_i}{R}\frac{\partial R_i}{\partial t'} = -\frac{v_i(t')\,R_i}{R} = -\frac{\vec{v}\cdot\vec{R}}{R}\,. \tag{7.46}$$

(Recall that $\vec{R}$ means $\vec{R}(t')$, and that it is given by (7.34).) Equation (7.45) therefore becomes

$$\frac{\partial R}{\partial t} = -\frac{\vec{v}\cdot\vec{R}}{R}\frac{\partial t'}{\partial t}\,, \tag{7.47}$$

and so, since we have from (7.35) that $R(t') = t - t'$, it follows that

$$1 - \frac{\partial t'}{\partial t} = -\frac{\vec{v}\cdot\vec{R}}{R}\frac{\partial t'}{\partial t}\,. \tag{7.48}$$

Solving for $\partial t'/\partial t$, we therefore have the results that

$$\frac{\partial t'}{\partial t} = \left(1 - \frac{\vec{v}\cdot\vec{R}}{R}\right)^{-1}, \tag{7.49}$$

$$\frac{\partial R}{\partial t} = -\frac{\vec{v}\cdot\vec{R}}{R - \vec{v}\cdot\vec{R}}\,. \tag{7.50}$$

Some other expressions we shall also need are as follows. First, from $t - t' = R(t')$ it follows that $\partial_i t' = -\partial_i R(t')$. Now $\vec{R}(t') = \vec{r} - \vec{r}_0(t')$, and so

$$R^2 = (x_j - x_j^0(t'))(x_j - x_j^0(t'))\,. \tag{7.51}$$

From this, by acting with ∂_i, we obtain

$$\begin{aligned} 2R\partial_i R &= 2(\delta_{ij} - \partial_i x_j^0(t'))(x_j - x_j^0(t'))\,, \\ &= 2R_i - 2\frac{\partial x_j^0(t')}{\partial t'}\frac{\partial t'}{\partial x_i}(x_j - x_j^0(t'))\,, \\ &= 2R_i - 2\vec{v}\cdot\vec{R}\,\partial_i t'\,. \end{aligned} \tag{7.52}$$

From this and $\partial_i t' = -\partial_i R(t')$ it follows that

$$\partial_i t' = -\frac{R_i}{R - \vec{v}\cdot\vec{R}}\,, \qquad \partial_i R = \frac{R_i}{R - \vec{v}\cdot\vec{R}}\,. \tag{7.53}$$

Further results that follow straightforwardly are

$$\begin{aligned} \partial_i R_j &= \partial_i(x_j - x_j^0(t')) = \delta_{ij} - \frac{\partial x_j^0(t')}{\partial t'}\partial_i t' = \delta_{ij} + \frac{v_j R_i}{R - \vec{v}\cdot\vec{R}}\,, \\ \partial_i v_j &= \frac{\partial v_j}{\partial t'}\partial_i t' = -\frac{\dot{v}_j R_i}{R - \vec{v}\cdot\vec{R}}\,, \\ \frac{\partial v_i}{\partial t} &= \frac{\partial v_i}{\partial t'}\frac{\partial t'}{\partial t} = \frac{\dot{v}_i R}{R - \vec{v}\cdot\vec{R}}\,, \\ \frac{\partial R}{\partial t} &= -\frac{\vec{v}\cdot\vec{R}}{R - \vec{v}\cdot\vec{R}}\,, \\ \frac{\partial \vec{R}}{\partial t} &= \frac{\partial \vec{R}}{\partial t'}\frac{\partial t'}{\partial t} = -\vec{v}\frac{\partial t'}{\partial t} = -\frac{\vec{v}\,R}{R - \vec{v}\cdot\vec{R}}\,. \end{aligned} \tag{7.54}$$

135

Note that $\dot{v}_i$ means $\partial v_i / \partial t'$; we shall define the acceleration $\vec{a}$ of the particle by

$$\vec{a} \equiv \frac{\partial \vec{v}}{\partial t'} . \tag{7.55}$$

We are now ready to evaluate the electric and magnetic fields. From (7.33) and the results above, we have

$$
\begin{aligned}
E_i &= -\partial_i \phi - \frac{\partial A_i}{\partial t} , \\
&= \frac{e}{(R - \vec{v} \cdot \vec{R})^2} \left(\partial_i R - \partial_i (v_j R_j) \right) - \frac{e}{R - \vec{v} \cdot \vec{R}} \frac{\partial v_i}{\partial t} + \frac{e v_i}{(R - \vec{v} \cdot \vec{R})^2} \left(\frac{\partial R}{\partial t} - \frac{\partial (\vec{v} \cdot \vec{R})}{\partial t} \right), \\
&= \frac{e}{(R - \vec{v} \cdot \vec{R})^3} \left\{ R_i - v_i (R - \vec{v} \cdot \vec{R}) - v^2 R_i + \vec{a} \cdot \vec{R} \, R_i - a_i \, R \, (R - \vec{v} \cdot \vec{R}) \right. \\
&\qquad\qquad \left. - v_i \vec{v} \cdot \vec{R} - v_i \vec{a} \cdot \vec{R} \, R + v^2 v_i R \right\}, \\
&= \frac{e(1 - v^2)(R_i - v_i \, R)}{(R - \vec{v} \cdot \vec{R})^3} + \frac{e[\vec{a} \cdot \vec{R} \, (R_i - v_i \, R) - a_i (R - \vec{v} \cdot \vec{R}) \, R]}{(R - \vec{v} \cdot \vec{R})^3} . \tag{7.56}
\end{aligned}
$$

This can be rewritten as

$$\vec{E} = \frac{e(1 - v^2)(\vec{R} - \vec{v} \, R)}{(R - \vec{v} \cdot \vec{R})^3} + \frac{e \vec{R} \times [(\vec{R} - \vec{v} \, R) \times \vec{a}]}{(R - \vec{v} \cdot \vec{R})^3} . \tag{7.57}$$

An analogous calculation of $\vec{B}$ shows that it can be written as

$$\vec{B} = \frac{\vec{R} \times \vec{E}}{R} . \tag{7.58}$$

Note that this means that $\vec{B}$ is perpendicular to $\vec{E}$.

The first term in (7.57) is independent of the acceleration $\vec{a}$, and so it represents a contribution that is present even if the charge is in uniform motion. It is easily seen that at large distance, where $R \to \infty$, it falls off like $1/R^2$. If the charge is moving with uniform velocity $\vec{v}$ then we shall have

$$\vec{r}_0(t) = \vec{r}_0(t') + (t - t') \, \vec{v}, \tag{7.59}$$

and so

$$
\begin{aligned}
\vec{R}(t') - \vec{v} \, R(t') &= \vec{r} - \vec{r}_0(t') - (t - t') \, \vec{v}, \\
&= \vec{r} - \vec{r}_0(t) + (t - t') \, \vec{v} - (t - t') \, \vec{v}, \\
&= \vec{r} - \vec{r}_0(t) \equiv \vec{R}_t . \tag{7.60}
\end{aligned}
$$

In other words, in this case of uniform motion, $\vec{R}(t') - \vec{v} \, R(t')$ is equal to the vector $\vec{R}_t \equiv \vec{r} - \vec{r}_0(t)$ that gives the line joining the location of the charge, *at the time the observation*

136

is made, to the location of the observation point. We shall also then have

$$
\begin{aligned}
R(t') - \vec{v} \cdot \vec{R}(t') &= R(t') - v^2 R(t') - \vec{v} \cdot \vec{R}_t , \\
&= (1 - v^2) R(t') - \vec{v} \cdot \vec{R}_t .
\end{aligned}
\tag{7.61}
$$

If we now introduce the angle θ between $\vec{v}$ and $\vec{R}_t$, we shall have $\vec{v} \cdot \vec{R}_t = v\, R_t \cos\theta$. Since, as we saw above, $\vec{R}(t') = \vec{v}\, R(t') + \vec{R}_t$, we obtain, by squaring,

$$
R^2(t') = v^2 R^2(t') + 2 v R_t\, R(t') \cos\theta + R_t^2 ,
\tag{7.62}
$$

and this quadratic equation for $R(t')$ can be solved to give

$$
R(t') = \frac{v R_t \, \cos\theta + R_t \sqrt{1 - v^2 \sin^2\theta}}{1 - v^2} .
\tag{7.63}
$$

Equation (7.61) then gives

$$
R(t') - \vec{v} \cdot \vec{R}(t') = v R_t \, \cos\theta + R_t \sqrt{1 - v^2 \sin^2\theta} - v R_t \, \cos\theta = R_t \sqrt{1 - v^2 \sin^2\theta} .
\tag{7.64}
$$

For a uniformly moving charge we therefore obtain the result

$$
\vec{E} = \frac{e \vec{R}_t}{R_t^3} \frac{1 - v^2}{(1 - v^2 \sin^2\theta)^{3/2}} ,
\tag{7.65}
$$

which has reproduced the result (5.41) that we had previously obtained by boosting from the rest frame of the charged particle.

The second term in (7.57) is proportional to $\vec{a}$, and so it occurs only for an accelerating charge. At large distance, this term falls off like $1/R$, in other words, much less rapidly than the $1/R^2$ fall-off of the first term in (7.57). In fact the $1/R$ fall-off of the acceleration term is characteristic of electromagnetic radiation, as we shall now discuss.

7.4 Radiation by accelerated charges

A charge at rest generates a purely electric field, and if it is in uniform motion it generates both $\vec{E}$ and $\vec{B}$ fields. In neither case, of course, does it radiate any energy. However, if the charge is *accelerating*, then it actually emits electromagnetic radiation.

The easiest case to consider is when the velocity of the charge is small compared with the speed of light. In this case the acceleration term in (7.57) is approximated by

$$
\vec{E} = \frac{e \vec{R} \times (\vec{R} \times \vec{a})}{R^3} = \frac{e \vec{n} \times (\vec{n} \times \vec{a})}{R} ,
\tag{7.66}
$$

where

$$
\vec{n} \equiv \frac{\vec{R}}{R} .
\tag{7.67}
$$

137

Note that $\vec{n} \cdot \vec{E} = 0$, and that $\vec{E}$ is also perpendicular to $\vec{n} \times \vec{a}$.

From (7.58) we shall also have

$$\vec{B} = \vec{n} \times \vec{E} \,. \tag{7.68}$$

As usual, all quantities here in the expressions for $\vec{E}$ and $\vec{B}$ are evaluated at the retarded time t'.

The energy flux, given by the Poynting vector, is given by

$$\vec{S} = \frac{1}{4\pi}\vec{E} \times \vec{B} = \frac{1}{4\pi}\,\vec{E} \times (\vec{n} \times \vec{E}) = \frac{1}{4\pi}E^2\,\vec{n} - \frac{1}{4\pi}(\vec{n} \cdot \vec{E})\,\vec{E} \,, \tag{7.69}$$

and so, since $\vec{n} \cdot \vec{E} = 0$ we have

$$\vec{S} = \frac{1}{4\pi}\,E^2\,\vec{n} \,. \tag{7.70}$$

Let us define θ to be the angle between the unit vector $\vec{n}$ and the acceleration $\vec{a}$. Then we shall have

$$\vec{E} = \frac{e\vec{n} \times (\vec{n} \times \vec{a})}{R} = \frac{e}{R}\left((\vec{n} \cdot \vec{a})\,\vec{n} - \vec{a}\right) = \frac{e}{R}\,(a\,\vec{n}\,\cos\theta - \vec{a}) \,, \tag{7.71}$$

and so

$$E^2 = \frac{e^2}{R^2}\,(a^2\cos^2\theta - 2a^2\cos^2\theta + a^2) = \frac{e^2a^2\sin^2\theta}{R^2} \,, \tag{7.72}$$

implying that the energy flux is

$$\vec{S} = \frac{e^2a^2\sin^2\theta}{4\pi R^2}\,\vec{n} \,. \tag{7.73}$$

The area element $d\vec{\Sigma}$ can be written as

$$d\vec{\Sigma} = R^2\,\vec{n}\,d\Omega \,, \tag{7.74}$$

where $d\Omega = \sin\theta\,d\theta d\varphi$ is the area element on the unit-radius sphere (i.e. the solid angle element). The power radiated into the area element $d\vec{\Sigma}$ is $dP = \vec{S} \cdot d\vec{\Sigma} = R^2\vec{n} \cdot \vec{S}\,d\Omega$, and so we find that

$$\frac{dP}{d\Omega} = \frac{e^2a^2}{4\pi}\,\sin^2\theta \tag{7.75}$$

is the power radiated per unit solid angle.

The total power radiated in all directions is given by

$$\begin{aligned} P &= \int \frac{dP}{d\Omega}\,d\Omega = \frac{e^2a^2}{4\pi} \int_0^\pi \sin^3\theta\,d\theta \int_0^{2\pi} d\varphi \,, \\ &= \tfrac{1}{2}e^2a^2 \int_0^\pi \sin^3\theta\,d\theta = \tfrac{1}{2}e^2a^2 \int_{-1}^1 (1-c^2)dc = \tfrac{2}{3}e^2a^2 \,, \end{aligned} \tag{7.76}$$

where, to evaluate the θ integral we change variable to $c = \cos\theta$. The expression

$$P = \tfrac{2}{3}e^2a^2 \tag{7.77}$$

is known as the *Larmor Formula* for a non-relativistic accelerating charge.

The Larmor formula can be generalised to the relativistic result fairly easily. In principle, we could simply repeat the argument given above, but without making the approximation that v is small compared to 1 (the speed of light). Note that in terms of the unit vector $\vec{n} = \vec{R}/R$, the expression (7.57) for the electric field becomes

$$\vec{E} = \frac{e(1-v^2)(\vec{n}-\vec{v})}{R^2\,(1-\vec{n}\cdot\vec{v})^3} + \frac{e\vec{n}\times[(\vec{n}-\vec{v})\times\vec{a}]}{R\,(1-\vec{n}\cdot\vec{v})^3}\,. \tag{7.78}$$

We can, in fact, obtain the relativisitic Larmor formula by a simple trick. First, we note from (7.78) that since $\vec{S} = (\vec{E}\times\vec{B})/(4\pi)$ and $\vec{B} = \vec{n}\times\vec{E}$, the energy flux from the acceleration term must be quadratic in the acceleration $\vec{a}$. We can also note that the total radiated power P is a Lorentz scalar (since it is energy per unit time, and each of these quantities transforms as the 0 component of a 4-vector). Thus, the task is to find a Lorentz-invariant expression for P that reduces to the non-relativisitic Larmor result (7.77) in the limit when v goes to zero.

First, we note that the non-relativistic Larmor formula (7.77) can be written as

$$P = \tfrac{2}{3}e^2 a^2 = \frac{2e^2}{3m^2}\left(\frac{d\vec{p}}{dt}\right)^2, \tag{7.79}$$

since in this non-relativistic limit we have $\vec{p} = m\,\vec{v}$. There is only one Lorentz-invariant quantity, quadratic in $\vec{a}$, that reduces to (7.79) in the limit that v goes to zero. It is given by

$$P = \frac{2e^2}{3m^2}\frac{dp^\mu}{d\tau}\frac{dp_\mu}{d\tau}\,, \tag{7.80}$$

where p^μ is the 4-momentum of the particle and τ is the proper time along its path. Noting that $p^\mu = m(\gamma, \gamma\vec{v})$, we see that

$$\frac{dp^\mu}{d\tau} = \gamma\frac{dp^\mu}{dt} = m\gamma\left(\gamma^3\vec{v}\cdot\vec{a}, \gamma^3(\vec{v}\cdot\vec{a})\,\vec{v} + \gamma\vec{a}\right), \tag{7.81}$$

and so

$$\begin{aligned}\frac{dp^\mu}{d\tau}\frac{dp_\mu}{d\tau} &= m^2\gamma^2[-\gamma^6(\vec{v}\cdot\vec{a})^2 + \gamma^6 v^2(\vec{v}\cdot\vec{a})^2 + 2\gamma^4(\vec{v}\cdot\vec{a})^2 + \gamma^2 a^2]\,, \\ &= m^2\gamma^2[\gamma^4(\vec{v}\cdot\vec{a})^2 + \gamma^2 a^2]\,. \end{aligned} \tag{7.82}$$

Now consider the quantity

$$\begin{aligned}a^2 - (\vec{v}\times\vec{a})^2 &= a^2 - \epsilon_{ijk}\epsilon_{i\ell m}v_j a_k v_\ell a_m\,, \\ &= a^2 - v^2 a^2 + (\vec{v}\cdot\vec{a})^2 = \frac{a^2}{\gamma^2} + (\vec{v}\cdot\vec{a})^2\,, \end{aligned} \tag{7.83}$$

which shows that we can write

$$\frac{dp^\mu}{d\tau}\frac{dp_\mu}{d\tau} = m^2\gamma^6\Big(\frac{a^2}{\gamma^2} + (\vec{v}\cdot\vec{a})^2\Big) = m^2\gamma^6\left[a^2 - (\vec{v}\times\vec{a})^2\right]. \tag{7.84}$$

Thus we see that the scalar P given in (7.80) is given by

$$P = \tfrac{2}{3}e^2\gamma^6[a^2 - (\vec{v}\times\vec{a})^2]. \tag{7.85}$$

This indeed reduces to the non-relativistic Larmor formula (7.77) if the velocity $\vec{v}$ is sent to zero. For the reasons we described above, it must therefore be the correct fully-relativistic Larmor result for the total power radiated by an accelerating charge.

7.5 Applications of Larmor formula

7.5.1 Linear accelerator

In a linear accelerator, a charged massive particle is accelerated along a straight-line trajectory, and so its velocity $\vec{v}$ and acceleration $\vec{a}$ are parallel. Defining $p = |\vec{p}| = m\gamma|\vec{v}|$, we have

$$\frac{dp}{dt} = m\gamma\frac{dv}{dt} + mv\frac{d\gamma}{dt}, \tag{7.86}$$

where $v = |\vec{v}|$ and $\gamma = (1 - v^2)^{-1/2}$. Clearly we have

$$v\frac{dv}{dt} = \vec{v}\cdot\frac{d\vec{v}}{dt} = \vec{v}\cdot\vec{a} = va, \qquad \frac{d\gamma}{dt} = \gamma^3\vec{v}\cdot\frac{d\vec{v}}{dt} = \gamma^3 va, \tag{7.87}$$

and so

$$\frac{dp}{dt} = m\gamma^3 a. \tag{7.88}$$

With $\vec{v}$ and $\vec{a}$ parallel, the relativisitic Larmor formula (7.85) gives $P = \tfrac{2}{3}e^2\gamma^6 a^2$, and so we have

$$P = \frac{2e^2}{3m^2}\Big(\frac{dp}{dt}\Big)^2. \tag{7.89}$$

The expression (7.89) gives the power that is radiated by the charge as it is accelerated along a straight line trajectory. In a particle accelerator, the goal, obviously, is to accelerate the particles to as high a velocity as possible. Equation (7.89) describes the the power that is lost through radiation when the particle is being accelerated. The energy $\mathcal{E}$ of the particle is related to its rest mass m and 3-momentum $\vec{p}$ by the standard formula

$$\mathcal{E}^2 = p^2 + m^2. \tag{7.90}$$

The rate of change of energy with distance x travelled, $d\mathcal{E}/dx$, is therefore given by differentiating (7.90) with respect to x, giving

$$\mathcal{E}\frac{d\mathcal{E}}{dx} = p\frac{dp}{dx}, \tag{7.91}$$

140

and so we have

$$\frac{d\mathcal{E}}{dx} = \frac{p}{\mathcal{E}}\frac{dp}{dx} = \frac{m\gamma v}{m\gamma}\frac{dp}{dx} = v\frac{dp}{dx} = \frac{dx}{dt}\frac{dp}{dx} = \frac{dp}{dt}. \tag{7.92}$$

This means that (7.89) can be rewritten as

$$P = \frac{2e^2}{3m^2}\left(\frac{d\mathcal{E}}{dx}\right)^2. \tag{7.93}$$

The "energy-loss factor" of the accelerator can be judged by taking the ratio of the power radiated divided by the power supplied. By energy conservation, the power supplied is equal to the rate of change of energy of the particle, $d\mathcal{E}/dt$. Thus we have

$$\begin{aligned}\frac{\text{Power radiated}}{\text{Power supplied}} &= \frac{P}{(d\mathcal{E}/dt)} = \frac{P}{(d\mathcal{E}/dx)}\frac{dt}{dx} = \frac{P}{v\,(d\mathcal{E}/dx)}, \\ &= \frac{2e^2}{3m^2 v}\frac{d\mathcal{E}}{dx}. \end{aligned} \tag{7.94}$$

In the relativistic limit, where v is very close to the speed of light (as is typically achieved in a powerful linear accelerator), we therefore have

$$\frac{\text{Power radiated}}{\text{Power supplied}} \approx \frac{2e^2}{3m^2}\frac{d\mathcal{E}}{dx}. \tag{7.95}$$

A typical electron linear accelerator achieves an energy input of about 10 MeV per metre, and this translates into an energy-loss factor of about 10^{-13}. In other words, very little of the applied power being used to accelerate the electron is lost through Larmor radiation.

7.5.2 Circular accelerator

The situation is very different in the case of a circular accelerator, since the transverse acceleration necessary to keep the particle in a circular orbit is typically very much larger than the linear acceleration discussed above. In other words, the direction of the 3-momemtum $\vec{p}$ is changing rapidly, while, by contrast, the energy, and hence the magnitude of $\vec{p}$, is relatively slowly-changing. In fact the change in $|\vec{p}|$ per revolution is rather small, and we can study the power loss by assuming that the particle is in an orbit of fixed angular frequency ω. This means that we shall have

$$\left|\frac{d\vec{p}}{dt}\right| = \omega\,|\vec{p}|, \tag{7.96}$$

and so

$$\left|\frac{d\vec{p}}{d\tau}\right| = \gamma\omega\,|\vec{p}|, \tag{7.97}$$

where $d\tau = dt/\gamma$ is the proper-time interval. Since the energy is constant in this approximation, we therefore have

$$\frac{dp^0}{d\tau} = 0, \qquad \text{and so} \qquad \frac{dp^\mu}{d\tau}\frac{dp_\mu}{d\tau} = \left(\frac{d\vec{p}}{d\tau}\right)^2 = \gamma^2\omega^2 p^2. \tag{7.98}$$

Using equation (7.80) for the Larmor power radiation, we therefore have

$$P = \frac{2e^2}{3m^2}\gamma^2\omega^2 p^2 = \tfrac{2}{3}e^2\gamma^4\omega^2 v^2 \,. \tag{7.99}$$

If the radius of the accelerator is R then the angular and linear velocities of the particle are related by $\omega = v/R$ and so the power loss is given by

$$P = \frac{2e^2\gamma^4 v^4}{3R^2} \,. \tag{7.100}$$

The radiative energy loss per revolution, $\Delta\mathcal{E}$, is given by the product of P with the period of the orbit, namely

$$\Delta\mathcal{E} = \frac{2\pi R P}{v} = \frac{4\pi e^2\gamma^4 v^3}{3R} \,. \tag{7.101}$$

Consider as an example a 10 GeV electron synchrotron, for which the radius R is about 100 metres. Since the rest mass of the electron is about 0.5 MeV, this means that $\gamma = (1-v^2)^{-1/2}$ for the 10 GeV electron is about $\gamma = 2000$. Plugging in the numbers, this implies an energy loss of about 10 MeV per revolution, or about 0.1% of the energy of the particle. Bearing in mind that the time taken to complete an orbit is very small (the electron is travelling at nearly the speed of light), it is necessary to supply energy at a very high rate in order to replenish the radiative loss. It also implies that there will be a considerable amount of radiation being emitted by the accelerator.

The losses due to this synchrotron radiation are much less dramatic for a proton accelerator such as the LHC at CERN. This is in part because the radius of the accelerator is much larger (about 4.3 Km), and also because the energy carried by the accelerated proton is very much larger. At the LHC the protons acquire an energy of about 7 TeV, and since their rest mass is about 1 MeV this means $\gamma \approx 7000$. According to a rough estimate (numbers need to be checked!) this would seem to imply an energy loss of about 35 MeV per revolution, which is only about 0.0005% of the proton energy.

7.6 Angular distribution of the radiated power

We saw previously that for a non-relativistic charged particle whose acceleration $\vec{a}$ makes an angle θ with respect to the position vector $\vec{R}$, the angular distribution of the radiated power is given by (see (7.75))

$$\frac{dP}{d\Omega} = \frac{e^2 a^2}{4\pi}\,\sin^2\theta \,. \tag{7.102}$$

In the general (i.e. relativistic) case, where the velocity $\vec{v}$ is not small, we have, from (7.78), that the large-R radiation-field term is

$$\vec{E} = \frac{e\vec{n}\times[(\vec{n}-\vec{v})\times\vec{a}]}{R\,(1-\vec{n}\cdot\vec{v})^3}\,, \qquad \vec{B} = \vec{n}\times\vec{E}\,. \tag{7.103}$$

The Poynting vector is therefore given by

$$
\begin{aligned}
\vec{S} &= \frac{1}{4\pi}\left(\vec{E}\times\vec{B}\right) = \frac{1}{4\pi}[\vec{E}\times(\vec{n}\times\vec{E})]\,,\\
&= \frac{1}{4\pi}\vec{n}\,E^2\,,
\end{aligned}
\tag{7.104}
$$

since $\vec{n}\cdot\vec{E} = 0$. Thus $\vec{S}$ is in the radial direction (parallel to $\vec{R}(t')$, and we have

$$
\vec{n}\cdot\vec{S} = \frac{e^2}{4\pi R^2}\left|\frac{\vec{n}\times[(\vec{n}-\vec{v})\times\vec{a}]}{(1-\vec{n}\cdot\vec{v})^3}\right|^2\,,
\tag{7.105}
$$

where as usual all quantities on the right-hand side are evaluated at the retarded time t' calculated from the equation $t - t' = R(t')$, with $\vec{R}(t') = \vec{r} - \vec{r}_0(t')$. It is conventional to denote the quantity in (7.105) by $[\vec{n}\cdot\vec{S}]_{\text{ret.}}$, to indicate that it is evaluated at the retarded time t'. Since $d\Sigma = \vec{n}\,R^2\,d\Omega$, we shall have

$$
\frac{dP(t)}{d\Omega} = [\vec{n}\cdot\vec{S}R^2]_{\text{ret}}\,.
\tag{7.106}
$$

The associated energy radiated during the time interval from $t = T_1$ to $t = T_2$ is therefore given by

$$
\frac{d\mathcal{E}}{d\Omega} = \int_{T_1}^{T_2}[R^2\,\vec{n}\cdot\vec{S}]_{\text{ret.}}dt\,,
\tag{7.107}
$$

Defining the corresponding retarded times $t' = T_i'$, the integral can therefore be rewritten as[26]

$$
\frac{d\mathcal{E}}{d\Omega} = \int_{T_1'}^{T_2'}[R^2\,\vec{n}\cdot\vec{S}]_{\text{ret.}}\frac{dt}{dt'}\,dt'\,.
\tag{7.108}
$$

The quantity $[R^2\,\vec{n}\cdot\vec{S}]_{\text{ret.}}(dt/dt')$ is the power radiated per unit solid angle, as measured with respect to the charge's retarded time t', and so we have the result that

$$
\frac{dP(t')}{d\Omega} = [R^2\,\vec{n}\cdot\vec{S}]_{\text{ret.}}\frac{dt}{dt'} = (1-\vec{n}\cdot\vec{v})[R^2\,\vec{n}\cdot\vec{S}]_{\text{ret.}}\,.
\tag{7.109}
$$

(Note that we used the result (7.49) here.)

7.6.1 Angular power distribution for linear acceleration

As an example, consider the situation when the charge is accelerated uniformly for only a short time, so that $\vec{v}$ as well as $\vec{a}$ are approximately constant during the time interval of the

---

[26]All that is really being said here is that we can relate the previously-defined quantity $dP(t)/d\Omega$ (power per unit solid angle as measured by the observer at time t) to $dP(t')/d\Omega$ (power per unit solid angle as measured at the particle, at retarded time t') by $dP/d\Omega\,dt = dP/d\Omega\,(dt/dt')\,dt' \equiv dP(t')/d\Omega\,dt'$.

acceleration. This means that $\vec{n}$ and R are approximately constant, and so from (7.105) and (7.109) we obtain the angular distribution

$$\frac{dP(t')}{d\Omega} = \frac{e^2}{4\pi} \frac{|\vec{n} \times [(\vec{n} - \vec{v}) \times \vec{a}]|^2}{(1 - \vec{n} \cdot \vec{v})^5} \, . \tag{7.110}$$

If we now suppose that the acceleration is linear, i.e. that $\vec{v}$ and $\vec{a}$ are parallel, then we obtain

$$\frac{dP(t')}{d\Omega} = \frac{e^2 a^2}{4\pi} \frac{\sin^2 \theta}{(1 - v \cos \theta)^5} \, , \tag{7.111}$$

where as before we define θ to be the angle between $\vec{a}$ and $\vec{n}$.

When $|v| << 1$, the expression (7.111) clearly reduces to the non-relativistic result given in (7.75). In this limit, the angular radiated power distribution is described by a figure-of-eight, oriented perpendicularly to the direction of the acceleration. As the velocity becomes larger, the two lobes of the figure-of-eight start to tilt forwards, along the direction of the acceleration. This is illustrated for the non-relativistic and relativisitic cases in Figures 1 and 2 below. In each case, the acceleration is to the right along the horizontal axis.
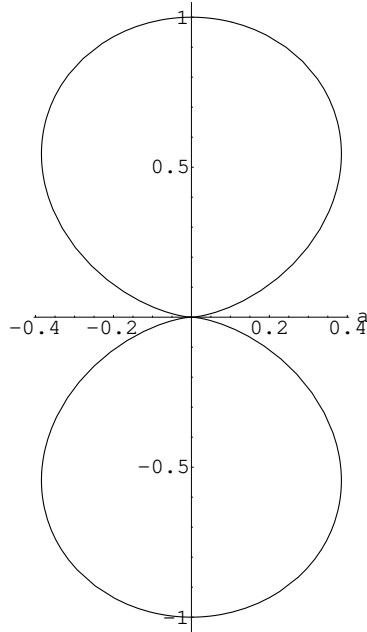

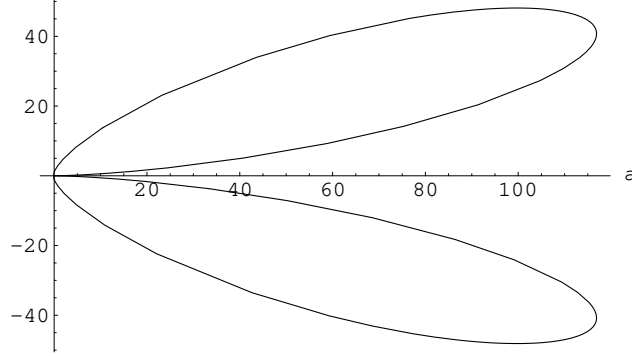Figure 3: The angular power distribution in the non-relativistic case

Figure 4: The angular power distribution in the relativistic case ($v = 4/5$)

The angle at which the radiated power is largest is found by solving $d(dP/d\Omega)/d\theta = 0$. This gives

$$2(1 - v\cos\theta)\cos\theta - 5v\sin^2\theta = 0\,,\tag{7.112}$$

and hence

$$\theta_{\text{max.}} = \arccos\left(\frac{\sqrt{1 + 15v^2} - 1}{3v}\right).\tag{7.113}$$

In the case of a highly relativistic particle, for which v is very close to the speed of light, the velocity itself is not a very convenient parameter, and instead we can more usefully characterise it by $\gamma = (1 - v^2)^{-1/2}$, which becomes very large in the relativistic limit. Thus, substituting $v = \sqrt{1 - \gamma^{-2}}$ into (7.113), we obtain

$$\theta_{\text{max.}} = \arccos\left(\frac{4\sqrt{1 - \frac{15}{16}\gamma^{-2}} - 1}{3\sqrt{1 - \gamma^{-2}}}\right).\tag{7.114}$$

At large γ we can expand the argument as a power series in γ^{-2}, finding that

$$\theta_{\text{max.}} \approx \arccos(1 - \tfrac{1}{8}\gamma^{-2})\,.\tag{7.115}$$

This implies that $\theta_{\text{max.}}$ is close to 0 when γ is very large. In this regime we have $\cos\theta_{\text{max.}} \approx 1 - \tfrac{1}{2}\theta_{\text{max.}}^2$, and so in the highly relativistic case we have

$$\theta_{\text{max.}} \approx \frac{1}{2\gamma}\,.\tag{7.116}$$

We see that the lobes of the angular power distribution tilt forward sharply, so that they are directed nearly parallel to the direction of acceleration of the particle.

Continuing with the highly-relativistic limit, we may consider the profile of the angular power distribution for all small angles θ. Substituting

$$v = \sqrt{1 - \gamma^{-2}}\,, \qquad \sin\theta \approx \theta\,, \qquad \cos\theta \approx 1 - \tfrac{1}{2}\theta^2\tag{7.117}$$

145

into (7.111), and expanding in inverse powers of γ, we find that

$$\frac{dP(t')}{d\Omega} \approx \frac{e^2 a^2 \theta^2}{4\pi \left(1 - \sqrt{1 - \gamma^{-2}}(1 - \frac{1}{2}\theta^2)\right)^5} \approx \frac{8e^2 a^2 \theta^2}{\pi(\gamma^{-2} + \theta^2)^5} , \qquad (7.118)$$

which can be written as

$$\frac{dP(t')}{d\Omega} \approx \frac{8e^2 a^2 \gamma^8}{\pi} \frac{(\gamma\theta)^2}{[1 + (\gamma\theta)^2]^5} . \qquad (7.119)$$

This shows that indeed there are two lobes, of characteristic width $\Delta\theta \sim 1/\gamma$, on each side of $\theta = 0$. The radiated power is zero in the exactly forward direction $\theta = 0$.

We can straightforwardly integrate our result (7.111) for the angular power distribution for a linearly-accelerated particle, to find the total radiated power. We obtain

$$P = \int \frac{dP(t')}{d\Omega} d\Omega = \frac{e^2 a^2}{4\pi} 2\pi \int_0^\pi \frac{\sin^2\theta}{(1 - v\cos\theta)^5} \sin\theta \, d\theta = \frac{1}{2}e^2 a^2 \int_{-1}^1 \frac{(1 - c^2)dc}{(1 - vc)^5} , \quad (7.120)$$

where $c = \cos\theta$. The integral is elementary, giving the result

$$P = \frac{2}{3}e^2 \gamma^6 a^2 . \qquad (7.121)$$

This can be seen to be in agreement with our earlier result (7.85), under the specialisation that $\vec{a}$ and $\vec{v}$ are parallel.

7.6.2 Angular power distribution for circular motion

For a second example, consider the situation of a charge that is in uniform circular motion. For these purposes, we need only assume that it is instantaneously in such motion; the complete path of the particle could be something more complicated than a circle, but such that at some instant it can be described by a circular motion.

Circular motion implies that the velocity $\vec{v}$ and the acceleration $\vec{a}$ are perpendicular. At the instant under consideration, we may choose a system of Cartesian axes oriented so that the velocity $\vec{v}$ lies along the z direction, and the acceleration lies along the x direction. The unit vector $\vec{n} = \vec{R}/R$ can then be parameterised by spherical polar coordinates (θ, φ) defined in the usual way; i.e. θ measures the angle between $\vec{n}$ and the z axis, and φ is the azimuthal angle, measured from the x axis, of the projection of $\vec{n}$ onto the (x, y) plane. Thus we shall have

$$\vec{n} = (\sin\theta\,\cos\varphi, \sin\theta\,\sin\varphi, \cos\theta) , \qquad \vec{v} = (0, 0, v) , \qquad \vec{a} = (a, 0, 0) . \qquad (7.122)$$

Of course, in particular, we have $\vec{n} \cdot \vec{v} = v\cos\theta$.

146

From (7.105) and (7.109), we have the general expression

$$\frac{dP(t')}{d\Omega} = \frac{e^2}{4\pi} \frac{\left|\vec{n} \times [(\vec{n} - \vec{v}) \times \vec{a}]\right|^2}{(1 - \vec{n} \cdot \vec{v})^5}, \tag{7.123}$$

for the angular distribution of the radiated power. Using the fact that $\vec{v} \cdot \vec{a} = 0$ in the case of circular motion, we have

$$
\begin{aligned}
|\vec{n} \times [(\vec{n} - \vec{v}) \times \vec{a}]|^2 &= |(\vec{n} \cdot \vec{a})(\vec{n} - \vec{v}) - (1 - \vec{n} \cdot \vec{v})\vec{a}|^2, \\
&= (\vec{n} \cdot \vec{a})^2(1 - 2\vec{n} \cdot \vec{v} + v^2) + (1 - \vec{n} \cdot \vec{v})^2 a^2 - 2(\vec{n} \cdot \vec{a})^2(1 - \vec{n} \cdot \vec{v}), \\
&= -(\vec{n} \cdot \vec{a})^2(1 - v^2) + (1 - \vec{n} \cdot \vec{v})^2 a^2, \\
&= (1 - v\cos\theta)^2 a^2 - \gamma^{-2} a^2 \sin^2\theta \cos^2\varphi, \tag{7.124}
\end{aligned}
$$

and so for instantaneous circular motion we have

$$\frac{dP(t')}{d\Omega} = \frac{e^2 a^2}{4\pi(1 - v\cos\theta)^3}\left[1 - \frac{\sin^2\theta \cos^2\varphi}{\gamma^2 (1 - v\cos\theta)^2}\right]. \tag{7.125}$$

We see that as v tends to 1, the angular distribution is peaked in the forward direction i.e. in the direction of the velocity $\vec{v}$, meaning that θ is close to 0.

The total power is obtained by integrating $\frac{dP(t')}{d\Omega}$ over all solid angles:

$$
\begin{aligned}
P(t') &= \int \frac{dP(t')}{d\Omega} d\Omega = \int_0^{2\pi} d\varphi \int_0^\pi \sin\theta d\theta \frac{dP(t')}{d\Omega}, \\
&= \int_0^{2\pi} d\varphi \int_0^\pi \sin\theta d\theta \frac{e^2 a^2}{4\pi(1 - v\cos\theta)^3}\left[1 - \frac{\sin^2\theta \cos^2\varphi}{\gamma^2 (1 - v\cos\theta)^2}\right], \\
&= \int_0^\pi \sin\theta d\theta \frac{e^2 a^2}{2(1 - v\cos\theta)^3}\left[1 - \frac{\sin^2\theta}{2\gamma^2 (1 - v\cos\theta)^2}\right], \\
&= \int_{-1}^1 \frac{e^2 a^2}{2(1 - vc)^3}\left[1 - \frac{1 - c^2}{2\gamma^2(1 - vc)^2}\right] dc, \tag{7.126}
\end{aligned}
$$

where $c = \cos\theta$. After performing the integration, we obtain

$$P(t') = \tfrac{2}{3} e^2 \gamma^4 a^2. \tag{7.127}$$

This expression can be compared with the general result (7.85), specialised to the case where $\vec{v}$ and $\vec{a}$ are perpendicular. Noting that then

$$(\vec{v} \times \vec{a})^2 = \epsilon_{ijk}\epsilon_{i\ell m}v_j a_k v_\ell a_m = v_j v_j a_k a_k - v_j a_j v_k a_k = v_j v_j a_k a_k = v^2 a^2, \tag{7.128}$$

we see that (7.85) indeed agrees with (7.127) in this case.

The total power radiated in the case of linear acceleration, with its γ^6 factor as in (7.121), is larger by a factor of γ^2 than the total power radiated in the case of circular

motion, *provided we take the acceleration a to be the same in the two cases.* However, this is not always the most relevant comparison to make. Another way to make the comparison is to take the magnitude of the applied force, $|d\vec{p}/dt|$, to be the same in the two cases. For circular motion we have that v is constant, and so

$$\frac{d\vec{p}}{dt} = m\gamma \frac{d\vec{v}}{dt} = m\,\gamma\,\vec{a}\,. \tag{7.129}$$

Thus for circular motion, we have from (7.127) that

$$P(t') = \frac{2e^2\gamma^2}{3m^2} \left|\frac{d\vec{p}}{dt}\right|^2\,. \tag{7.130}$$

By contrast, for linear acceleration, where $\vec{v}$ is parallel to $\vec{a}$, we have

$$\frac{d\vec{p}}{dt} = m\gamma\vec{a} + m\gamma^3(\vec{v}\cdot\vec{a})\vec{v} = m\,\gamma^3\,\vec{a}\,, \tag{7.131}$$

and so this gives

$$P(t') = \frac{2e^2}{3m^2} \left|\frac{d\vec{p}}{dt}\right|^2\,. \tag{7.132}$$

Thus if we hold $|d\vec{p}/dt|$ fixed when comparing the two, we see that it is the particle in circular motion whose radiated power is larger than that of the linearly-accelerated particle, by a factor of γ^2.

7.7 Frequency distribution of radiated energy

In this section, we shall discuss the spectrum of frequencies of the electromagnetic radiation emitted by an accelerating charge. The basic technique for doing this will be to perform a Fourier transform of the time dependence of the radiated power.

In general, we have

$$\frac{dP(t)}{d\Omega} = [R^2\,\vec{n}\cdot\vec{S}]_{\text{ret}} = \frac{1}{4\pi}\left|[R\vec{E}]_{\text{ret}}\right|^2\,. \tag{7.133}$$

Let

$$\vec{G}(t) = \frac{1}{\sqrt{4\pi}}\,[R\vec{E}]_{\text{ret}}\,, \tag{7.134}$$

so that we shall have

$$\frac{dP(t)}{d\Omega} = |\vec{G}(t)|^2\,. \tag{7.135}$$

Note that here $dP(t)/d\Omega$ is expressed in the *observer's* time t, and not the retarded time t'. This is because our goal here will be to determine the frequency spectrum of the electromagnetic radiation as measured by the observer.

148

Suppose that the acceleration of the charge occurs only for a finite period of time, so that the total energy emitted is finite. We shall assume that the observation point is far enough away from the charge that the distance travelled by the charge during the period when it is accelerating subtends only a small angle as seen by the observer.

The total energy radiated per unit solid angle is given by

$$\frac{dW}{d\Omega} = \int_{-\infty}^{\infty} \frac{dP}{d\Omega}\, dt = \int_{-\infty}^{\infty} |\vec{G}(t)|^2 dt \,. \tag{7.136}$$

We now define the Fourier transform $\vec{g}(\omega)$ of $\vec{G}(t)$:

$$\vec{g}(\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \vec{G}(t)\, e^{\mathrm{i}\omega t}\, dt \,. \tag{7.137}$$

In the usual way, the inverse transform is then

$$\vec{G}(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \vec{g}(\omega)\, e^{-\mathrm{i}\omega t}\, d\omega \,. \tag{7.138}$$

It follows that

$$\frac{dW}{d\Omega} = \int_{-\infty}^{\infty} |\vec{G}(t)|^2 dt = \frac{1}{2\pi} \int_{-\infty}^{\infty} dt \int_{-\infty}^{\infty} d\omega \int_{-\infty}^{\infty} d\omega'\, \vec{g}^{\,*}(\omega') \cdot \vec{g}(\omega)\, e^{\mathrm{i}\,(\omega'-\omega)t} \,. \tag{7.139}$$

The t integration can be performed, using

$$\int_{-\infty}^{\infty} dt\, e^{\mathrm{i}\,(\omega'-\omega)t} = 2\pi\delta(\omega' - \omega) \,, \tag{7.140}$$

and so

$$\frac{dW}{d\Omega} = \int_{-\infty}^{\infty} d\omega \int_{-\infty}^{\infty} d\omega'\, \vec{g}^{\,*}(\omega') \cdot \vec{g}(\omega)\, \delta(\omega' - \omega) = \int_{-\infty}^{\infty} d\omega\, \vec{g}^{\,*}(\omega) \cdot \vec{g}(\omega) \,, \tag{7.141}$$

i.e.

$$\frac{dW}{d\Omega} = \int_{-\infty}^{\infty} d\omega\, |\vec{g}(\omega)|^2 \,. \tag{7.142}$$

(The result that (7.136) can be expressed as (7.142) is known as *Parseval's Theorem* in Fourier transform theory.)

We can re-express (7.142) as

$$\frac{dW}{d\Omega} = \int_{0}^{\infty} d\omega\, \frac{d^2 I(\omega, \vec{n})}{d\omega d\Omega} \,, \tag{7.143}$$

where

$$\frac{d^2 I(\omega, \vec{n})}{d\omega d\Omega} = |\vec{g}(\omega)|^2 + |\vec{g}(-\omega)|^2 \tag{7.144}$$

is the energy emitted per unit solid angle per unit frequency interval. If $\vec{G}(t) = [R\vec{E}]_{\mathrm{ret}}/\sqrt{4\pi}$ is real, then

$$\vec{g}(-\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} dt\, \vec{G}(t)\, e^{-\mathrm{i}\omega t} = \vec{g}^{\,*}(\omega) \,, \tag{7.145}$$

149

and then

$$\frac{d^2 I(\omega, \vec{n})}{d\omega d\Omega} = 2|\vec{g}(\omega)|^2 \,. \tag{7.146}$$

Using the expression for $\vec{E}$ in (7.103), the Fourier transform $\vec{g}(\omega)$, given by (7.137) with (7.134), is

$$\vec{g}(\omega) = \frac{e}{2\sqrt{2}\,\pi} \int_{-\infty}^{\infty} e^{\mathrm{i}\omega t} \left[\frac{\vec{n} \times [(\vec{n} - \vec{v}\,) \times \vec{a}\,]}{(1 - \vec{n}\cdot\vec{v}\,)^3}\right]_{\mathrm{ret}} dt \,, \tag{7.147}$$

where as usual, the subscript "ret" is a reminder that the quantity is evaluated at the retarded time t'. Since

$$dt = \frac{dt}{dt'}\,dt' = (1 - \vec{n}\cdot\vec{v}\,)\,dt' \,, \tag{7.148}$$

we therefore have

$$\vec{g}(\omega) = \frac{e}{2\sqrt{2}\,\pi} \int_{-\infty}^{\infty} e^{\mathrm{i}\omega\left(t' + R(t')\right)}\, \frac{\vec{n} \times [(\vec{n} - \vec{v}\,) \times \vec{a}\,]}{(1 - \vec{n}\cdot\vec{v}\,)^2}\, dt' \,. \tag{7.149}$$

(We have now dropped the "ret" reminder, since everything inside the integrand now depends on the retarded time t'.)

We are assuming that the observation point is far away from the accelerating charge, and that the period over which the acceleration occurs is short enough that the the vector $\vec{n} = \vec{R}(t')/R(t')$ is approximately constant during this time interval. It is convenient to choose the origin to be near to the particle during its period of acceleration. With the observer being far away, at position vector $\vec{r}$, it follows from $\vec{R}(t') = \vec{r} - \vec{r}_0(t')$ that to a good approximation we have

$$R^2(t') \approx r^2 - 2\vec{r}\cdot\vec{r}_0(t') \,, \tag{7.150}$$

and so

$$R(t') \approx r\left(1 - \frac{2\vec{r}\cdot\vec{r}_0(t')}{r^2}\right)^{1/2} \approx r - \frac{\vec{r}\cdot\vec{r}_0(t')}{r} \,. \tag{7.151}$$

Furthermore, we can also approximate $\vec{n} \equiv \vec{R}(t')/R(t')$ by $\vec{r}/r$, and so

$$R(t') \approx r - \vec{n}\cdot\vec{r}_0(t') \,. \tag{7.152}$$

Substituting this into (7.149), there will be a phase factor $e^{\mathrm{i}\omega r}$ that can be taken outside the integral, since it is independent of t'. This overall phase factor is unimportant (it will cancel out when we calculate $|\vec{g}(\omega)|^2$, and so we may drop it and write

$$\vec{g}(\omega) = \frac{e}{2\sqrt{2}\,\pi} \int_{-\infty}^{\infty} e^{\mathrm{i}\omega\left(t' - \vec{n}\cdot\vec{r}_0(t')\right)}\, \frac{\vec{n} \times [(\vec{n} - \vec{v}\,) \times \vec{a}\,]}{(1 - \vec{n}\cdot\vec{v}\,)^2}\, dt' \,. \tag{7.153}$$

From (7.146) we therefore have

$$\frac{d^2 I(\omega, \vec{n})}{d\omega d\Omega} = \frac{e^2}{4\pi^2} \left|\int_{-\infty}^{\infty} e^{\mathrm{i}\omega\left(t' - \vec{n}\cdot\vec{r}_0(t')\right)}\, \frac{\vec{n} \times [(\vec{n} - \vec{v}\,) \times \vec{a}\,]}{(1 - \vec{n}\cdot\vec{v}\,)^2}\, dt'\right|^2 \,, \tag{7.154}$$

150

as the energy per unit solid angle per unit frequency interval.

The integral can be neatened up by observing that we can write

$$\frac{\vec{n} \times [(\vec{n} - \vec{v}) \times \vec{a}]}{(1 - \vec{n} \cdot \vec{v})^2} = \frac{d}{dt'} \left[\frac{\vec{n} \times (\vec{n} \times \vec{v})}{1 - \vec{n} \cdot \vec{v}} \right], \tag{7.155}$$

under the assumption that $\vec{n}$ is a constant. This can be seen be distributing the derivative, to obtain

$$
\begin{aligned}
\frac{d}{dt'} \left[\frac{\vec{n} \times (\vec{n} \times \vec{v})}{1 - \vec{n} \cdot \vec{v}} \right] &= \frac{\vec{n} \times (\vec{n} \times \vec{a})}{1 - \vec{n} \cdot \vec{v}} + \frac{\vec{n} \times (\vec{n} \times \vec{v})(\vec{n} \cdot \vec{a})}{(1 - \vec{n} \cdot \vec{v})^2}, \\
&= \frac{(1 - \vec{n} \cdot \vec{v})\big(\vec{n}(\vec{n} \cdot \vec{a}) - \vec{a}\big) + (\vec{n}(\vec{n} \cdot \vec{v}) - \vec{v})(\vec{n} \cdot \vec{a})}{(1 - \vec{n} \cdot \vec{v})^2}, \\
&= \frac{(\vec{n} \cdot \vec{a})(\vec{n} - \vec{v}) - (1 - \vec{n} \cdot \vec{v})\vec{a}}{(1 - \vec{n} \cdot \vec{v})^2}, \\
&= \frac{\vec{n} \times [(\vec{n} - \vec{v}) \times \vec{a}]}{(1 - \vec{n} \cdot \vec{v})^2}. \tag{7.156}
\end{aligned}
$$

This allows us to integrate (7.154) by parts, to give

$$\frac{d^2 I(\omega, \vec{n})}{d\omega d\Omega} = \frac{e^2}{4\pi^2} \left| -\int_{-\infty}^{\infty} \frac{\vec{n} \times (\vec{n} \times \vec{v})}{1 - \vec{n} \cdot \vec{v}} \frac{d}{dt'} e^{i\omega\left(t' - \vec{n}\cdot\vec{r}_0(t')\right)} \, dt' \right|^2, \tag{7.157}$$

and hence

$$\frac{d^2 I(\omega, \vec{n})}{d\omega d\Omega} = \frac{e^2 \omega^2}{4\pi^2} \left| \int_{-\infty}^{\infty} \vec{n} \times (\vec{n} \times \vec{v}) \, e^{i\omega\left(t' - \vec{n}\cdot\vec{r}_0(t')\right)} \, dt' \right|^2, \tag{7.158}$$

It should be remarked here that the effect of having integrated by parts is that the acceleration $\vec{a}$ no longer appears in the expression (7.158). Prior to the integration by parts, the fact that we were taking the acceleration to be non-zero for only a finite time interval ensured that the integration over all t' from $-\infty$ to ∞ would be cut down to an integration over only the finite time interval during which $\vec{a}$ was non-zero. After the integration by parts, the integrand in (7.158) no longer vanishes outside the time interval of the non-zero acceleration, and so one might worry about issues of convergence, and the validity of having dropped the boundary terms at $t' = \pm\infty$ coming from the integration by parts. In fact, it can be verified that all is well, and any problem with convergence can be handled by introducing a convergence factor $e^{-\epsilon|t'|}$, and then sending ϵ to zero.

We shall make use of the result (7.158) in two applications. In the first, we shall calculate the frequency spectrum for a relativistic particle in instantaneous circular motion.

7.8   Frequency spectrum for relativistic circular motion

Consider a particle which, at some instant, is following a circular arc of radius ρ. We shall choose axes so that the arc lies in the (x, y) plane, and choose the origin so that at $t = 0$

the particle is located at the origin, $x = y = 0$. Without loss of generality, we may choose the unit vector $\vec{n}$ (which points in the direction of the observation point) to lie in the (x, z) plane. We shall, for notational convenience, drop the prime from the time t', so from now on t will denote the retarded time.

In fact, we shall make the assumption that the particle is moving highly relativistically. As we saw earlier, this means that the radiation is concentrated into very narrow beams in the direction of the velocity vector, and hence we need only consider a small arc of the trajectory.

The position vector of the particle at time t will be given by

$$\vec{r}_0 = \left(\rho \sin \frac{vt}{\rho}, \rho \cos \frac{vt}{\rho} - \rho, 0 \right), \tag{7.159}$$

where $v = |\vec{v}|$ is its speed. Since $\vec{v} = d\vec{r}_0(t)/dt$, we shall have

$$\vec{v} = \left(v \cos \frac{vt}{\rho}, -v \sin \frac{vt}{\rho}, 0 \right). \tag{7.160}$$

We may parameterise the unit vector $\vec{n}$, which we are taking to lie in the (x, z) plane, in terms of the angle θ between $\vec{n}$ and the x axis:

$$\vec{n} = (\cos \theta, 0, \sin \theta). \tag{7.161}$$

We then have

$$\vec{n} \times (\vec{n} \times \vec{v}) = (\vec{n} \cdot \vec{v}) \, \vec{n} - \vec{v} = \left(- v \sin^2 \theta \, \cos \frac{vt}{\rho}, -v \sin \frac{vt}{\rho}, v \sin \theta \, \cos \theta \, \cos \frac{vt}{\rho} \right). \tag{7.162}$$

We shall write this as

$$\vec{n} \times (\vec{n} \times \vec{v}) = -v \sin \frac{vt}{\rho} \, \vec{e}_\parallel + v \sin \theta \, \cos \frac{vt}{\rho} \, \vec{e}_\perp, \tag{7.163}$$

where

$$\vec{e}_\parallel = (0, 1, 0) \qquad \text{and} \qquad \vec{e}_\perp = \vec{n} \times \vec{e}_\parallel = (-\sin \theta, 0, \cos \theta). \tag{7.164}$$

We shall consider a particle whose velocity is highly-relativistic. It will be recalled from our earlier discussions that for such a particle, the electromagnetic radiation will be more or less completely concentrated in the range of angles θ very close to 0. Thus, to a good approximation we shall have $\vec{e}_\perp \approx (0, 0, 1)$, which is the unit normal to the plane of the circular motion. In what follows, we shall make approximations that are valid for small θ, and also for small t. We shall also assume that v is very close to 1 (the speed of light).

From (7.159) and (7.161), we find

$$
\begin{aligned}
t - \vec{n} \cdot \vec{r}_0(t) &= t - \rho \cos\theta \, \sin\frac{vt}{\rho} \approx t - \rho(1 - \tfrac{1}{2}\theta^2)\Big[\frac{vt}{\rho} - \tfrac{1}{6}\Big(\frac{vt}{\rho}\Big)^3\Big], \\
&\approx (1 - v)t + \tfrac{1}{2}\theta^2 vt + \frac{v^3 t^3}{6\rho^2}, \\
&\approx \tfrac{1}{2}(1 + v)(1 - v)t + \tfrac{1}{2}\theta^2 \, t + \frac{t^3}{6\rho^2}, \\
&= \frac{t}{2\gamma^2} + \tfrac{1}{2}\theta^2 \, t + \frac{t^3}{6\rho^2}.
\end{aligned}
\tag{7.165}
$$

From (7.163), we find

$$
\vec{n} \times (\vec{n} \times \vec{v}) \approx -\frac{t}{\rho}\, \vec{e}_\parallel + \theta\, \vec{e}_\perp.
\tag{7.166}
$$

We therefore find from (7.158) that

$$
\begin{aligned}
\frac{d^2 I}{d\omega d\Omega} &\approx \frac{e^2\omega^2}{4\pi^2}\Big| - g_\parallel(\omega)\, \vec{e}_\parallel + g_\perp(\omega)\, \vec{e}_\perp \Big|^2, \\
&= \frac{e^2\omega^2}{4\pi^2}\Big(|g_\parallel(\omega)|^2 + |g_\perp(\omega)|^2\Big),
\end{aligned}
\tag{7.167}
$$

where

$$
\begin{aligned}
g_\parallel(\omega) &= \frac{1}{\rho}\int_{-\infty}^{\infty} t\, e^{\mathrm{i}\omega[(\gamma^{-2}+\theta^2)t + \frac{1}{3}t^3\rho^{-2}]/2}\, dt, \\
g_\perp(\omega) &= \theta\int_{-\infty}^{\infty} e^{\mathrm{i}\omega[(\gamma^{-2}+\theta^2)t + \frac{1}{3}t^3\rho^{-2}]/2}\, dt.
\end{aligned}
\tag{7.168}
$$

Letting

$$
u = \frac{t}{\rho}\,(\gamma^{-2} + \theta^2)^{-1/2}, \qquad \xi = \tfrac{1}{3}\omega\rho(\gamma^{-2} + \theta^2)^{3/2},
\tag{7.169}
$$

leads to

$$
\begin{aligned}
g_\parallel(\omega) &= \rho(\gamma^{-2} + \theta^2)\int_{-\infty}^{\infty} u e^{3\mathrm{i}\xi(u+u^3/3)/2}\, du, \\
g_\perp(\omega) &= \rho\theta(\gamma^{-2} + \theta^2)^{1/2}\int_{-\infty}^{\infty} e^{3\mathrm{i}\xi(u+u^3/3)/2}\, du.
\end{aligned}
\tag{7.170}
$$

These integrals are related to Airy integrals, or modified Bessel functions:

$$
\int_{0}^{\infty} u\sin[3\xi(u+u^3/3)/2]\, du = \frac{1}{\sqrt{3}}\, K_{2/3}(\xi), \qquad \int_{0}^{\infty}\cos[3\xi(u+u^3/3)/2]\, du = \frac{1}{\sqrt{3}}\, K_{1/3}(\xi),
\tag{7.171}
$$

and so we have

$$
\frac{d^2 I}{d\omega d\Omega} \approx \frac{e^2\omega^2\rho^2}{3\pi^2}\,(\gamma^{-2} + \theta^2)^2\Big[\big(K_{2/3}(\xi)\big)^2 + \frac{\theta^2}{\gamma^{-2} + \theta^2}\,\big(K_{1/3}(\xi)\big)^2\Big].
\tag{7.172}
$$

The asymptotic forms of the modified Bessel functions $K_\nu(x)$, for small x and large x, are

$$K_\nu(x) \longrightarrow \tfrac{1}{2}\Gamma(\nu)\left(\frac{2}{x}\right)^\nu; \qquad x \longrightarrow 0,$$

$$K_\nu(x) \longrightarrow \sqrt{\frac{\pi}{2x}}\,e^{-x}; \qquad x \longrightarrow \infty. \tag{7.173}$$

It therefore follows from (7.172) that $d^2I/(d\omega d\Omega)$ falls off rapidly when ξ becomes large. Bearing in mind that γ^{-2} is small (since the velocity of the particle is very near to the speed of light), and that θ has been assumed to be small, we see from (7.169) that there is a regime where ξ can be large, whilst still fulfilling our assumptions, if $\omega\rho$ is large enough. The value of ξ can then become very large if θ increases sufficiently (whilst still being small compared to 1), and so the radiation is indeed concentrated around very small angles θ.

If ω becomes sufficiently large that $\omega\rho\gamma^{-3}$ is much greater than 1 then ξ will be very large even if $\theta = 0$. Thus, there is an effective high-frequency cut-off for all angles. It is convenient to define a "cut-off" frequency ω_c for which $\xi = 1$ at $\theta = 0$:

$$\omega_c = \frac{3\gamma^3}{\rho} = \frac{3}{\rho}\left(\frac{\mathcal{E}}{m}\right)^3. \tag{7.174}$$

If the particle is following a uniform periodic circular orbit, with angular frequency $\omega_0 = v/\rho \approx 1/\rho$, then we shall have

$$\omega_c = 3\omega_0\left(\frac{\mathcal{E}}{m}\right)^3. \tag{7.175}$$

The radiation in this case of a charged particle in a highly relativistic circular orbit is known as "Synchrotron Radiation."

Consider the frequency spectrum of the radiation in the orbital plane, $\theta = 0$. In the two regimes $\omega << \omega_c$ and $\omega >> \omega_c$ we shall therefore have

$$\omega << \omega_c: \quad \frac{d^2I}{d\omega d\Omega}\Big|_{\theta=0} \approx e^2\left(\frac{\Gamma(2/3)}{\pi}\right)^2\left(\frac{3}{4}\right)^{1/3}(\omega\rho)^{2/3},$$

$$\omega >> \omega_c: \quad \frac{d^2I}{d\omega d\Omega}\Big|_{\theta=0} \approx \frac{3e^2\gamma^2}{2\pi}\frac{\omega}{\omega_c}e^{-2\omega/\omega_c}. \tag{7.176}$$

This shows that the power per unit solid angle per unit frequency increases from 0 like $\omega^{2/3}$ for small ω, reaches a peak around $\omega = \omega_c$, and then falls off exponentially rapidly one ω is significantly greater than ω_c.

It is clear that one could continue with the investigation of the properties of the synchrotron radiation in considerably more depth. For example, would could consider the detailed angular distibution of the radiation as a function of θ, and one could consider the total power per unit frequency interval, obtained by integrating over all solid angles:

$$\frac{dI}{d\omega} = \int \frac{d^2I}{d\omega d\Omega}\,d\Omega. \tag{7.177}$$

A discussion of further details along these lines can be found in almost any of the advanced electrodynamics textbooks.

7.9 Frequency spectrum for periodic motion

Suppose that the motion of the charged particle is exactly periodic, with period $T = 2\pi/\omega_0$, where ω_0 is the angular frequency of the particle's motion. This implies that the emitted radiation will have time dependence of the form

$$H(t) = \sum_{n=-\infty}^{\infty} b_n \, e^{-i \, n \omega_0 \, t} \,, \tag{7.178}$$

so its Fourier transform $h(\omega)$ is zero except when ω is an integer multiple of ω_0, and for these values it is proportional to a delta function:

$$\begin{aligned}
h(\omega) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{i\omega t} \, H(t) dt = \frac{1}{\sqrt{2\pi}} \sum_n b_n \int_{-\infty}^{\infty} e^{i \, (\omega - n\omega_0)t} \, dt \,, \\
&= \sqrt{2\pi} \sum_n b_n \, \delta(\omega - n\omega_0) \,. \tag{7.179}
\end{aligned}$$

(When unspecified, the summation will be assumed to be over all n, positive and negative.)

In fact, it is more appropriate to work with Fourier series, rather than Fourier transforms, in this situation with a discrete frequency spectrum. Going back to section 7.7, we therefore now expand $\vec{G}(t) = \frac{1}{\sqrt{4\pi}} [R \vec{E}]_{ret}$ in the Fourier series

$$\vec{G}(t) = \sum_n \vec{a}_n \, e^{-i \, n \, \omega_0 \, t} \,. \tag{7.180}$$

Multiplying by $e^{i \, m \, \omega_0 \, t}$ and integrating over the period $T = 2\pi/\omega_0$ gives

$$\frac{1}{T} \int_0^T e^{i m \omega_0 t} \vec{G}(t) dt = \frac{1}{T} \sum_{n \geq 1} \vec{a}_n \int_0^T e^{i (m-n)\omega_0 t} dt = \vec{a}_m \,, \tag{7.181}$$

since the integral of $e^{i \, (m-n)\omega_0 t}$ vanishes unless $n = m$:

$$\frac{1}{T} \int_0^T e^{i \, (m-n)\omega_0 \, t} dt = \delta_{m,n} \,. \tag{7.182}$$

Thus the coefficients $\vec{a}_n$ in the Fourier series (7.180) are given by

$$\vec{a}_n = \frac{1}{T} \int_0^T e^{i \, n\omega_0 \, t} \vec{G}(t) dt \,. \tag{7.183}$$

The analogue of Parseval's theorem for the case of the discrete Fourier series is now given by considering

$$\frac{1}{T} \int_0^T |\vec{G}(t)|^2 dt = \frac{1}{T} \int_0^T \sum_{m,n} \vec{a}_n \cdot \vec{a}_m^* \, e^{i \, (m-n)\omega_0 \, t} dt = \sum_{n=-\infty}^{\infty} |\vec{a}_n|^2 \,. \tag{7.184}$$

The time average of the power per unit solid angle is therefore given by

$$\left\langle \frac{dP}{d\Omega} \right\rangle = \frac{1}{T} \int_0^T \frac{dP}{d\Omega} \, dt = \frac{1}{T} \int_0^T |\vec{G}(t)|^2 dt = \sum_{n=-\infty}^{\infty} |\vec{a}_n|^2 \, . \tag{7.185}$$

In a manner analogous to the earlier discussion of the continuum case, we may express this as a sum of terms associated with frequencies $\omega_n = n\,\omega_0$ for $n \geq 1$ (i.e. fundamental plus harmonics). Thus we write

$$\left\langle \frac{dP}{d\Omega} \right\rangle = \sum_{n \geq 1} \frac{dP_n}{d\Omega} \, , \tag{7.186}$$

from which it follows that

$$\frac{dP_n}{d\Omega} = |\vec{a}_n|^2 + |\vec{a}_{-n}|^2 \, . \tag{7.187}$$

(We do not need to consider the zero-frequency $n = 0$ mode in the sum, since this would correspond to a static component to the field, which will not arise here.) If $\vec{G}(t)$ is real, it follows that $\vec{a}_n^* = \vec{a}_{-n}$, and so

$$\frac{dP_n}{d\Omega} = 2|\vec{a}_n|^2 \, . \tag{7.188}$$

The expression $dP_n/d\Omega$ has the interpretation of being the time-averaged power per unit solid angle in the n'th mode.

It is now a straightforward matter, using (7.183), to obtain an expression for $|\vec{a}_n|^2$ in terms of the integral of the retarded electric field. The steps follow exactly in parallel with those we described in section 7.7, except that the integral $\int_{-\infty}^{\infty} dt$ is now replaced by $T^{-1} \int_0^T dt$. The upshot is that the expression (7.158) for $d^2I/(d\omega d\Omega)$ is replaced by[27]

$$\frac{dP_n}{d\Omega} = \frac{e^2 n^2 \omega_0^4}{(2\pi)^3} \left| \int_0^T \vec{n} \times (\vec{n} \times \vec{v}) \, e^{i\,n\omega_0 \left(t' - \vec{n} \cdot \vec{r}_0(t') \right)} \, dt' \right|^2 \, , \tag{7.189}$$

where $T = 2\pi/\omega_0$. This gives the expression for the time-averaged power per unit solid angle in the n'th Fourier mode.

Since we are assuming the observer (at $\vec{r}$) is far away from the particle, and since the integral in (7.189) is taken over the finite time interval $T = 2\pi/\omega_0$, it follows that to a good approximation we can freely take the unit vector $\vec{n}$ outside the integral. Thus we may make the replacement

$$\int_0^T \vec{n} \times (\vec{n} \times \vec{v}) \, e^{i\,n\omega_0 \left(t' - \vec{n} \cdot \vec{r}_0(t') \right)} \, dt' \longrightarrow \vec{n} \times \left(\vec{n} \times \int_0^T \vec{v} \, e^{i\,n\omega_0 \left(t' - \vec{n} \cdot \vec{r}_0(t') \right)} \right) \, dt' \, . \tag{7.190}$$

Now, for any vector $\vec{V}$, we have that

$$|\vec{n} \times (\vec{n} \times \vec{V})|^2 = |\vec{n} \cdot \vec{V} - \vec{V}|^2 = V^2 - (\vec{n} \cdot \vec{V})^2 \, , \tag{7.191}$$

---

[27]The integer n labelling the modes is not to be confused with the unit vector $\vec{n}$, of course!

and on the other hand we also have

$$|\vec{n} \times \vec{V}|^2 = (\vec{n} \times \vec{V}) \cdot (\vec{n} \times \vec{V}) = \vec{n} \cdot [\vec{V} \times (\vec{n} \times \vec{V})] = \vec{n} \cdot [V^2 \vec{n} - (\vec{n} \cdot \vec{V})\vec{V}] = V^2 - (\vec{n} \cdot \vec{V})^2 \,. \quad (7.192)$$

Thus $|\vec{n} \times (\vec{n} \times \vec{V})|^2 = |\vec{n} \times \vec{V}|^2$, and so we can re-express (7.189) as

$$\frac{dP_n}{d\Omega} = \frac{e^2 n^2 \omega_0^4}{(2\pi)^3} \Big| \int_0^T \vec{v} \times \vec{n} \, e^{\mathrm{i} n\omega_0 \left(t' - \vec{n}\cdot\vec{r}_0(t')\right)} \, dt' \Big|^2 \,, \quad (7.193)$$

where $T = 2\pi/\omega_0$ and ω_0 is the angular frequency of the periodic motion.

7.10 Cerenkov radiation

So far, all the situations we have considered have involved electromagnetic fields in a vacuum, i.e. in the absence of any dielectric or magnetically permeable media. In this section, we shall take a brief foray into a situation where there is a dielectric medium.

It will be recalled that if a medium has permittivity ϵ and permeability μ, then electromagnetic waves in the medium will propagate with speed $\tilde{c} = 1/\sqrt{\epsilon\,\mu}$. This means in general that the "speed of light" in the medium will be less than the speed of light in vacuum. A consequence of this is that a charged particle, such as an electron, can travel faster than the local speed of light inside the medium. This leads to an interesting effect, known as *Cerenkov Radiation*. In practice, the types of media of interest are those that are optically transparent, such as glass or water, and these have magnetic permeability μ very nearly equal to 1, while the electric permittivity ϵ can be quite significantly greater than 1. Thus for the purposes of our discussion, we shall assume that $\mu = 1$ and that the local speed of light is reduced because ϵ is significantly greater than 1.

We shall make use of the result (7.158) for the radiated power spectrum, in order to study the Cerenkov radiation. First, we shall need to introduce the dielectric constant into the formula. This can be done by a simple scaling argument. We shall also, just for the purposes of this section, restore the explicit symbol c for the speed of light. This can be done by sending

$$t \longrightarrow ct \,, \qquad \omega \longrightarrow \frac{\omega}{c} \,. \quad (7.194)$$

(Of course any other quantity that involves time will also need to be rescaled appropriately too. This is just dimensional analysis.)

Having first restored the explicit appearance of the speed of light, we shall next make further rescalings of the fields in order to introduce a non-unit dielectric constant ϵ. Referring back to the discussion in section 2.1, it can be seen that this can be done by means of

the rescalings

$$\rho \longrightarrow \frac{\rho}{\sqrt{\epsilon}}, \quad \vec{E} \longrightarrow \sqrt{\epsilon}\,\vec{E}, \quad \vec{B} \longrightarrow \frac{\vec{B}}{\sqrt{\epsilon}}, \quad c \longrightarrow \frac{c}{\sqrt{\epsilon}}. \tag{7.195}$$

Of course the scaling of the charge density ρ implies that we must also rescale the charge e of the particle, according to

$$e \longrightarrow \frac{e}{\sqrt{\epsilon}}. \tag{7.196}$$

Note that c continues to mean the speed of light in vacuum. The speed of light inside the dielectric medium is given by

$$\tilde{c} = \frac{c}{\sqrt{\epsilon}}. \tag{7.197}$$

The expression (7.158) for the radiated power per unit solid angle per unit frequency interval now becomes

$$\frac{d^2 I(\omega, \vec{n})}{d\omega d\Omega} = \frac{e^2 \omega^2 \sqrt{\epsilon}}{4\pi^2 c^3} \left| \int_{-\infty}^{\infty} \vec{n} \times (\vec{n} \times \vec{v})\, e^{\mathrm{i}\,\omega(t' - \sqrt{\epsilon}\,\vec{n}\cdot\vec{r}_0(t')/c)}\, dt' \right|^2, \tag{7.198}$$

For a charge moving at constant velocity $\vec{v}$, we shall have

$$\vec{r}_0(t') = \vec{v}\,t', \tag{7.199}$$

and so (7.198) gives

$$\frac{d^2 I(\omega, \vec{n})}{d\omega d\Omega} = \frac{e^2 \omega^2 \sqrt{\epsilon}}{4\pi^2 c^3} |\vec{n} \times \vec{v}|^2 \left| \int_{-\infty}^{\infty} e^{\mathrm{i}\,\omega t'(1 - \sqrt{\epsilon}\,\vec{n}\cdot\vec{v}/c)}\, dt' \right|^2, \tag{7.200}$$

since $|\vec{n} \times (\vec{n} \times \vec{v})|^2 = |(\vec{n} \cdot \vec{v})\,\vec{n} - \vec{v}|^2 = v^2 - (\vec{n} \cdot \vec{v})^2 = |\vec{n} \times \vec{v}|^2$.

The integration over t' produces a delta-function.[28] Defining θ to be the angle between $\vec{n}$ and $\vec{v}$, so that $\vec{n} \cdot \vec{v} = v \cos\theta$, we therefore have

$$\frac{d^2 I(\omega, \vec{n})}{d\omega d\Omega} = \frac{e^2 \omega^2 \sqrt{\epsilon}}{c^3}\, v^2\, \sin^2\theta\, |\delta(\omega(1 - \sqrt{\epsilon}\,(v/c)\cos\theta))|^2, \tag{7.201}$$

and so (using $\delta(ax) = \delta(x)/a$)

$$\frac{d^2 I(\omega, \vec{n})}{d\omega d\Omega} = \frac{e^2 \sqrt{\epsilon}}{c^3}\, v^2\, \sin^2\theta\, |\delta(1 - \sqrt{\epsilon}\,(v/c)\cos\theta)|^2. \tag{7.202}$$

This expression shows that all the radiation is emitted at a single angle θ_c, known as the *Cerenkov Angle*, given by

$$\cos\theta_c = \frac{c}{v\sqrt{\epsilon}}. \tag{7.203}$$

---

[28]The occurrence of the delta-function is because of the unphysical assumption that the particle has been moving in the medium forever. Below, we shall obtain a more realistic expression by supposing that the particle travels through a slab of medium of finite thickness.

Note that in terms of $\tilde{c}$, the speed of light in the medium, as given in (7.197), we have

$$\cos\theta_c = \frac{\tilde{c}}{v}. \qquad (7.204)$$

This makes clear that the phenomenon of Cerenkov radiation occurs only if $v > \tilde{c}$, i.e. if the charged particle is moving through the medium at a velocity that is greater than the local velocity of light in the medium. In fact one can understand the Cerenkov radiation as a kind of "shock wave," very like the acoustic shock wave that occurs when an aircraft is travelling faster than the speed of sound. The Cerenkov angle θ_c is given by a very simple geometric construction, shown in Figure 3 below. The circles show the light-fronts of light emitted by the particle. Since the particle is travelling faster than the speed of light in the medium, it "outruns" the circles, leaving a trail of light-fronts tangent to the angled line in the figure. This is the light-front of the Cerenkov radiation.
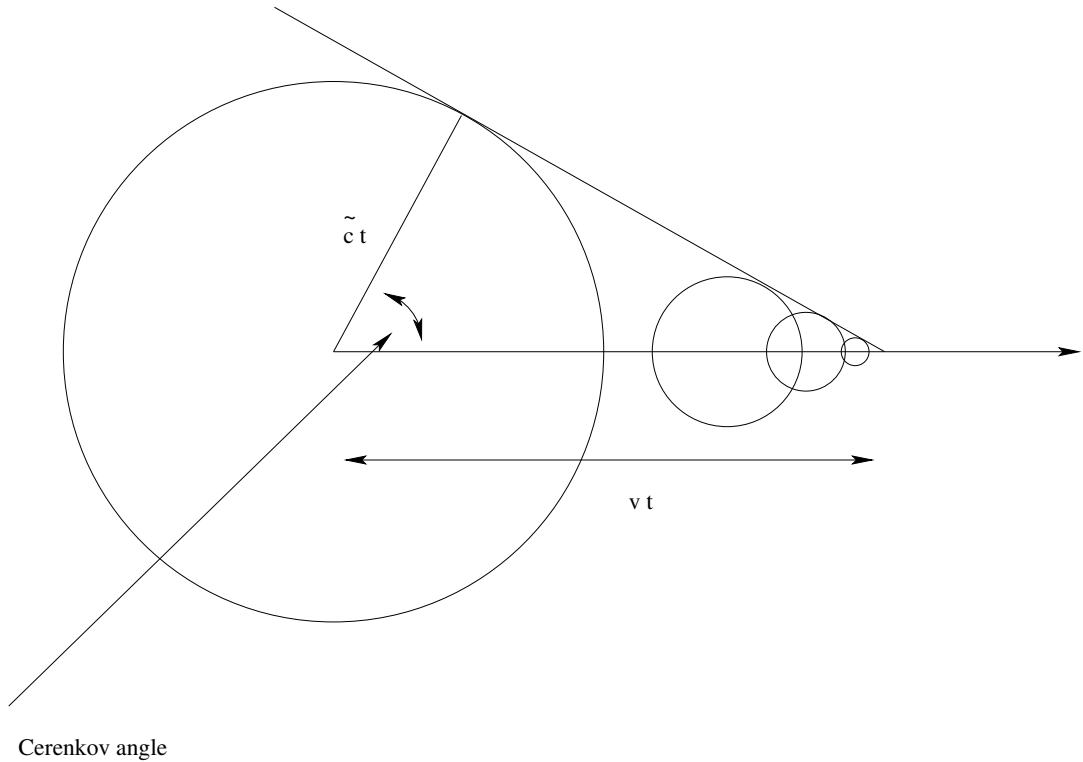

Cerenkov angle

Figure 5: The Cerenkov angle θ_c is given by $\cos\theta_c = (\tilde{c}t)/(vt) = \tilde{c}/v$.

As mentioned above, the squared delta-function in (7.202) is the result of making the

159

unrealistic assumption that the particle has been ploughing through the medium for ever, at a speed greater than the local speed of light. A more realistic situation would be to consider a charged particle entering a thin slab of dielectric medium, such that it enters at time $t' = -T$ and exits at $t' = +T$. The expression (7.200) is then replaced by

$$\frac{d^2 I(\omega, \vec{n})}{d\omega d\Omega} = \frac{e^2 \omega^2 \sqrt{\epsilon}}{4\pi^2 c^3} |\vec{n} \times \vec{v}|^2 \left| \int_{-T}^{T} e^{i\omega t'(1 - \sqrt{\epsilon}\,\vec{n}\cdot\vec{v}/c)}\, dt' \right|^2, \tag{7.205}$$

which, using $\int_{-T}^{T} dt e^{ibt} = 2b^{-1} \sin bT$, therefore implies that

$$\frac{d^2 I(\omega, \vec{n})}{d\omega d\Omega} = \frac{e^2 \omega^2 \sqrt{\epsilon}\, v^2 T^2 \sin^2 \theta}{\pi^2 c^3} \left(\frac{\sin[\omega T(1 - \sqrt{\epsilon}\,(v/c)\cos\theta)]}{\omega T(1 - \sqrt{\epsilon}\,(v/c)\cos\theta} \right)^2. \tag{7.206}$$

This is sharply peaked around the Cerenkov angle θ_c given by (7.203).

Integrating over all angles we obtain the total energy per unit frequency interval

$$\frac{dI}{d\omega} = \int \frac{d^2 I}{d\omega d\Omega}\, d\Omega \approx \frac{2e^2 \omega^2 \sqrt{\epsilon}\, v^2 T^2 \sin^2 \theta_c}{\pi c^3} \int_0^{\pi} \left(\frac{\sin[\omega T(1 - \sqrt{\epsilon}\,(v/c)\cos\theta)]}{\omega T(1 - \sqrt{\epsilon}\,(v/c)\cos\theta} \right)^2 \sin\theta\, d\theta. \tag{7.207}$$

(The integrand is peaked sharply around $\theta = \theta_c$, so to a good approximation we can take the $\sin^2 \theta$ factor outside the integral, calling it $\sin^2 \theta_c$.) Letting $x = \cos\theta$, the remaining integral can be written as

$$\int_{-1}^{1} \left(\frac{\sin[\omega T(1 - \sqrt{\epsilon}\,(v/c)x)]}{\omega T(1 - \sqrt{\epsilon}\,(v/c)x} \right)^2 dx \approx \int_{-\infty}^{\infty} \left(\frac{\sin[\omega T(1 - \sqrt{\epsilon}\,(v/c)x)]}{\omega T(1 - \sqrt{\epsilon}\,(v/c)x} \right)^2 dx. \tag{7.208}$$

(The limits of integration can, to a good approximation, be extended to $\pm\infty$ because the integrand is peaked around $x = \cos\theta_c$.) Letting $\omega T - \omega T \sqrt{\epsilon}\, x/c = -y$, the integral becomes

$$\frac{c}{\omega T \sqrt{\epsilon}\, v} \int_{-\infty}^{\infty} \frac{\sin^2 y}{y^2}\, dy = \frac{\pi c}{\omega T \sqrt{\epsilon}\, v}, \tag{7.209}$$

and so expression (7.207) for the total energy per unit frequency interval becomes

$$\frac{dI}{d\omega} \approx \frac{2e^2 v \omega T \sin^2 \theta_c}{c^2}. \tag{7.210}$$

The distance through the slab is given by $2vT$, and so dividing by this, we obtain an expression for the total energy of Cerenkov radiation per unit frequency interval per unit path length:

$$\frac{d^2 I}{d\omega d\ell} = \frac{e^2 \omega}{c^2} \sin^2 \theta_c = \frac{e^2 \omega}{c^2} \left(1 - \frac{c^2}{v^2 \epsilon} \right). \tag{7.211}$$

This is known as the Frank-Tamm relation. Note that this expression grows linearly with ω, which means that the bulk of the energy is concentrated in the higher frequencies of electromagnetic radiation. Of course there must be some limit, which arises because the

160

dielectric constant will fall off with increasing frequency, and so the Cerenkov effect will cease to operate at high enough frequencies.[29] In practice, the peak of the frequency spectrum for Cerenkov radiation is in the ultra-violet.

The bluish-green glow visible in pictures of nuclear fuel rods immersed in water is a familiar example of Cerenkov radiation. Apart from looking nice, the Cerenkov effect is also of practical use, for measuring the velocity of charged particles moving at relativistic speeds. One can determine the velocity by allowing the particles to move through a slab of suitably-chosen dielectric material, and measuring the Cerenkov angle.

7.11   Thompson scattering

Another application of the Larmor formula is in the phenomenon known as *Thompson scattering*. Consider a plane electromagnetic wave incident on a particle of charge e and mass m. The charge will oscillate back and forth in the electric field of the wave, and so it will therefore emit electromagnetic radiation itself. The net effect is that the electron "scatters" some of the incoming wave.

In most circumstances, we can assume that the induced oscillatory motion of the electron will be non-relativistic. As we saw in (7.75), if Θ is the angle between the acceleration $\vec{a}$ and the unit vector $\vec{n}$ (which lies along the line from the electron to the observation point), then the power radiated per unit solid angle is

$$\frac{dP}{d\Omega} = \frac{e^2 a^2}{4\pi} \sin^2 \Theta .$$
(7.212)

Let us suppose that the plane electromagnetic wave has electric field given by (the real part of)

$$\vec{E} = E_0 \, \vec{\epsilon} \, e^{\mathrm{i}\,(\vec{k}\cdot\vec{r} - \omega t)} ,$$
(7.213)

and that the wave-vector $\vec{k}$ lies along the z axis. The unit polarisation vector $\vec{\epsilon}$, which must therefore lie in the (x,y) plane, may be parameterised as

$$\vec{\epsilon} = (\cos\psi, \sin\psi, 0) .$$
(7.214)

Using standard spherical polar coordinates, the unit vector $\vec{n}$ will be given by

$$\vec{n} = (\sin\theta \, \cos\varphi, \sin\theta \, \sin\varphi, \cos\theta) .$$
(7.215)

---

[29]At sufficiently high frequencies, which implies very small wavelengths, the approximation in which the medium is viewed as a continuum with an effective dielectric constant breaks down, and it looks more and more like empty space with isolated charges present. At such length scales the electron is more or less propagating through a vacuum, and so there is no possibility of its exceeding the local speed of light. Thus the Cerenkov effect tails off at sufficiently high frequencies.

In particular, this means

$$\vec{n} \cdot \vec{\epsilon} = \sin\theta \left(\cos\varphi \cos\psi + \sin\varphi \sin\psi\right) = \sin\theta \, \cos(\varphi - \psi) \,. \tag{7.216}$$

The acceleration of the electron will be given by

$$m\vec{a} = e\vec{E} \,, \qquad \text{so} \qquad \vec{a} = \frac{e}{m} E_0 \, \vec{\epsilon} \, e^{i\,\omega(z-t)} \,. \tag{7.217}$$

Note that this means

$$\vec{n} \cdot \vec{a} = \frac{e}{m} E_0 \, \vec{n} \cdot \vec{\epsilon} e^{i\,\omega(z-t)} = \frac{e}{m} E_0 \, e^{i\,\omega(z-t)} \sin\theta \, \cos(\varphi - \psi) \,. \tag{7.218}$$

Since $\vec{n} \cdot \vec{a} = a \cos\Theta$, it follows that (7.212) becomes

$$\frac{dP}{d\Omega} = \frac{e^2}{4\pi} \left(a^2 - (\vec{n} \cdot \vec{a})^2\right) \,, \tag{7.219}$$

and so the time average will be given by

$$\left\langle \frac{dP}{d\Omega} \right\rangle = \frac{e^4}{8\pi m^2} \, |E_0|^2 [1 - (\vec{n} \cdot \vec{\epsilon})^2] \,. \tag{7.220}$$

Thus we find

$$\left\langle \frac{dP}{d\Omega} \right\rangle = \frac{e^4}{8\pi m^2} \, |E_0|^2 \left[1 - \sin^2\theta \, \cos^2(\varphi - \psi)\right] \,. \tag{7.221}$$

The direction of the polarisation (in the (x, y) plane) of the incoming electromagnetic wave is parameterised by the angle ψ. For unpolarised incoming waves, we should average over all angles ψ. Thus we obtain

$$\begin{aligned}
\left\langle \left\langle \frac{dP}{d\Omega} \right\rangle \right\rangle_\psi &\equiv \frac{1}{2\pi} \int_0^{2\pi} d\psi \left\langle \frac{dP}{d\Omega} \right\rangle = \frac{e^4}{8\pi m^2} \, |E_0|^2 \left(1 - \tfrac{1}{2} \sin^2\theta\right) \,, \\
&= \frac{e^4}{16\pi m^2} \, |E_0|^2 \left(1 + \cos^2\theta\right) \,.
\end{aligned} \tag{7.222}$$

The scattering cross section $d\sigma/d\Omega$ is then defined by

$$\frac{d\sigma}{d\Omega} = \frac{\text{Energy radiated/unit time/unit solid angle}}{\text{Incident energy flux/unit area/unit time}} \,. \tag{7.223}$$

The denominator here will just be $|E_0|^2/(8\pi)$, which is the time average of the Poynting flux for the incoming wave. Thus we arrive at the *Thompson Formula* for the cross section:

$$\frac{d\sigma}{d\Omega} = \frac{e^4(1 + \cos^2\theta)}{2m^2} \,. \tag{7.224}$$

The total scattering cross section is obtained by integrating $d\sigma/d\Omega$ over all solid angles, which gives

$$\begin{aligned}
\sigma &= \int \frac{d\sigma}{d\Omega} d\Omega = 2\pi \int_0^\pi \frac{d\sigma}{d\Omega} \sin\theta d\theta \,, \\
&= \frac{\pi e^4}{m^2} \int_0^\pi \sin^3\theta d\theta = \frac{\pi e^4}{m^2} \int_{-1}^1 (1 + c^2) dc \,,
\end{aligned} \tag{7.225}$$

and so we find

$$\sigma = \frac{8\pi e^4}{3m^2} \,. \tag{7.226}$$

8 Radiating Systems

8.1 Fields due to localised oscillating sources

Consider a localised system of oscillating charge, at a single frequency ω. The charge density and current density can therefore be written as

$$\rho(\vec{r}, t) = \rho(\vec{r})\, e^{-\mathrm{i}\,\omega t}\,, \qquad \vec{J}(\vec{r}, t) = \vec{J}(\vec{r})\, e^{-\mathrm{i}\,\omega t}\,. \tag{8.1}$$

From the expressions (7.22) and (7.23) for the retarded potentials, we shall have

$$\begin{aligned}
\phi(\vec{r}, t) &= \int \frac{\rho(\vec{r}', t - |\vec{r} - \vec{r}'|)}{|\vec{r} - \vec{r}'|}\, d^3\vec{r}'\,, \\
&= e^{-\mathrm{i}\,\omega t} \int \frac{\rho(\vec{r}')}{|\vec{r} - \vec{r}'|}\, e^{\mathrm{i}\,k|\vec{r} - \vec{r}'|}\, d^3\vec{r}'\,.
\end{aligned} \tag{8.2}$$

Note that here k is simply equal to ω, and we have switched to the symbol k in the exponential inside the integral because it looks more conventional. In a similar fashion, we shall have

$$\vec{A}(\vec{r}, t) = e^{-\mathrm{i}\,\omega t} \int \frac{\vec{J}(\vec{r}')}{|\vec{r} - \vec{r}'|}\, e^{\mathrm{i}\,k|\vec{r} - \vec{r}'|}\, d^3\vec{r}'\,. \tag{8.3}$$

From these expressions for ϕ and $\vec{A}$, we can calculate $\vec{E} = -\vec{\nabla}\phi - \partial\vec{A}/\partial t$ and $\vec{B} = \vec{\nabla} \times \vec{A}$. In fact, because of the simple monochromatic nature of the time dependence, we can calculate $\vec{E}$ easily, once we know $\vec{B}$, from the Maxwell equation

$$\vec{\nabla} \times \vec{B} - \frac{\partial \vec{E}}{\partial t} = 4\pi \vec{J}\,. \tag{8.4}$$

Away from the localised source region we have $\vec{J} = 0$. From the time dependence we have $\partial\vec{E}/\partial t = -\mathrm{i}\,\omega\vec{E} = -\mathrm{i}\,k\vec{E}$, and so we shall have

$$\vec{E} = \frac{\mathrm{i}}{k}\, \vec{\nabla} \times \vec{B}\,. \tag{8.5}$$

Let us suppose that the region where the source charges and currents are non-zero is of scale size d. The wavelength of the monochromatic waves that they generate will be given by

$$\lambda = \frac{2\pi}{\omega} = \frac{2\pi}{k}\,. \tag{8.6}$$

We shall assume that $d << \lambda$, i.e. the scale size of the source region is very small compared with the wavelength of the electromagnetic waves that are produced. It will be convenient to choose the origin of the coordinate system to lie within the neighbourhood of the source region, so that we may therefore assume

$$|\vec{r}'| << \lambda \tag{8.7}$$

for all integration points $\vec{r}\,'$ in the expressions (8.2) and (8.3).

The discussion of the electromagnetic fields generated by these sources can then be divided, like all Gaul, into three parts:

$$\text{Near zone, or } \textbf{Static zone}: \quad d << r << \lambda\,,$$
$$\text{Intermediate zone, or } \textbf{Induction zone}: \quad d << r \sim \lambda\,,$$
$$\text{Far zone, or } \textbf{Radiation zone}: \quad d << \lambda << r\,. \tag{8.8}$$

8.1.1   The static zone

First, let us consider the near zone, where $r << \lambda$. Equivalently, we may say that

$$\text{Static zone}: \quad kr << 1\,. \tag{8.9}$$

Since we are also assuming $d << \lambda$, and that the origin of the coordinate system is located in the neighbourhood of the source region, it follows that in the near zone, we can just approximate $e^{\mathrm{i}\,k|\vec{r}-\vec{r}\,'|}$ by 1. Thus we shall have

$$\vec{A}(\vec{r},t) \approx e^{-\mathrm{i}\,\omega t} \int \frac{\vec{J}(\vec{r}\,')}{|\vec{r}-\vec{r}\,'|}\, d^3\vec{r}\,'\,, \tag{8.10}$$

in the near zone. Aside from the time-dependent factor $e^{-\mathrm{i}\,\omega t}$, this is just like the expression for the magnetostatic case. We can make a standard expansion, in terms of spherical harmonics:

$$\frac{1}{|\vec{r}-\vec{r}\,'|} = \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} \frac{4\pi}{2\ell+1} \frac{r_<^\ell}{r_>^{\ell+1}} Y_{\ell m}^*(\theta',\varphi')\, Y_{\ell m}(\theta,\varphi)\,, \tag{8.11}$$

where $r_>$ is the larger of $r = |\vec{r}|$ and $r' = |\vec{r}\,'|$, and $r_<$ is the smaller of r and r', and (θ,φ) and (θ',φ') are the spherical polar angles for $\vec{r}$ and $\vec{r}\,'$ respectively. Thus we shall have, since in our case $r' << r$,

$$\vec{A}(\vec{r},t) = e^{-\mathrm{i}\,\omega t} \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} \frac{4\pi}{2\ell+1} \frac{1}{r^{\ell+1}} Y_{\ell m}(\theta,\varphi) \int \vec{J}(\vec{r}\,')\, r'^\ell Y_{\ell m}^*(\theta',\varphi')\, d^3\vec{r}\,'\,. \tag{8.12}$$

In the near zone, therefore, the electromagnetic fields will just be like static fields, except that they are oscillating in time.

8.1.2   The radiation zone

Next, we shall consider the far zone, or radiation zone. Here, we shall have $kr >> 1$, while at the same time $kr' << 1$. In other words, the source is (as always) small compared with

the wavelength, but that the fields are being observed from a large distance which is much larger than the wavelength. This means that

$$|\vec{r} - \vec{r}'|^2 = r^2 - 2\vec{r} \cdot \vec{r} + r'^2 \approx r^2 - 2\vec{r} \cdot \vec{r}', \tag{8.13}$$

and so

$$|\vec{r} - \vec{r}'| \approx r \left(1 - \frac{2\vec{r} \cdot \vec{r}'}{r^2}\right)^{1/2} \approx r - \frac{\vec{r} \cdot \vec{r}'}{r}, \tag{8.14}$$

and hence

$$|\vec{r} - \vec{r}'| \approx r - \vec{n} \cdot \vec{r}', \tag{8.15}$$

where $\vec{n}$ is the unit vector along $\vec{r}$:

$$\vec{n} = \frac{\vec{r}}{r}. \tag{8.16}$$

For a more complete and comprehensive treatment, see section 8.3

From (8.3) we shall therefore have

$$\vec{A}(\vec{r}, t) \approx e^{i(kr - \omega t)} \int \frac{\vec{J}(\vec{r}')}{r - \vec{n} \cdot \vec{r}'} \, e^{-ik\vec{n} \cdot \vec{r}'} \, d^3\vec{r}', \tag{8.17}$$

in the far zone, and hence, to a very good approximation,

$$\vec{A}(\vec{r}, t) \approx \frac{1}{r} e^{i(kr - \omega t)} \int \vec{J}(\vec{r}') \, e^{-ik\vec{n} \cdot \vec{r}'} \, d^3\vec{r}'. \tag{8.18}$$

(Recall that k and ω are just two names for the same quantity here.)

The magnetic field is given by $\vec{B} = \vec{\nabla} \times \vec{A}$, or $B_i = \epsilon_{ijk} \partial_j A_k$. We are interested in the contribution that dominates at large distance, and this therefore comes from the term where the derivative lands on the e^{ikr} factor rather than the $1/r$ factor. Thus we shall have

$$
\begin{aligned}
B_i &\approx \epsilon_{ijk} \frac{1}{r}(\partial_j e^{ikr}) e^{-i\omega t} \int J_k(\vec{r}') \, e^{-ik\vec{n} \cdot \vec{r}'} \, d^3\vec{r}', \\
&= i\,k \frac{x_j}{r} \epsilon_{ijk} \frac{1}{r} e^{ikr} e^{-i\omega t} \int J_k(\vec{r}') \, e^{-ik\vec{n} \cdot \vec{r}'} \, d^3\vec{r}'
\end{aligned}
\tag{8.19}
$$

and so

$$\vec{B} \approx i\,k\,\vec{n} \times \vec{A}. \tag{8.20}$$

Note that the magnetic field in this leading-order approximation falls off like $1/r$. This is characteristic of electromagnetic radiation, as we have seen previously.

The electric field can be calculated from the magnetic field using (8.5), and again the leading-order behaviour comes from the term where the gradient operator lands on the e^{ikr} factor. The rule again is therefore that $\vec{\nabla} \to i\,k\,\vec{n}$, and so we find

$$\vec{E} \approx -k\,\vec{n} \times (\vec{n} \times \vec{A}). \tag{8.21}$$

165

Note that (8.20) and (8.21) imply

$$\vec{n} \cdot \vec{B} = 0, \qquad \vec{n} \cdot \vec{E} = 0, \qquad \vec{E} \cdot \vec{B} = 0. \tag{8.22}$$

Thus $\vec{E}$ and $\vec{B}$ are transverse and orthogonal. This is characteristic of radiation fields.

Since we are assuming that the characteristic size d of the source is very small compared with the wavelength, $d << \lambda = 2\pi/k$, it follows that $kd << 1$ and so the quantity $k|\vec{n} \cdot \vec{r}'|$ appearing in the exponential in the integrand in (8.18) is much smaller than 1. This means that it is useful to expand the exponential in a Taylor series, giving

$$\vec{A}(\vec{r}, t) \approx \frac{1}{r} e^{i(kr - \omega t)} \sum_{m \geq 0} \frac{(-i\,k)^m}{m!} \int \vec{J}(\vec{r}') \, (\vec{n} \cdot \vec{r}')^m \, d^3\vec{r}', \tag{8.23}$$

where the terms in the sum fall off rapidly with m.

8.1.3 The induction zone

This is the intermediate zone, where $d << \lambda \sim r$, which means that the fields are being observed from a distance that is comparable with the wavelength, and so $kr \sim 1$. In this case, we need to consider the exact expansion of $e^{i\,k|\vec{r} - \vec{r}'|} |\vec{r} - \vec{r}'|^{-1}$. It turns out that this can be written as

$$\frac{e^{i\,k|\vec{r} - \vec{r}'|}}{|\vec{r} - \vec{r}'|} = 4\pi i\,k \sum_{\ell \geq 0} j_\ell(kr') \, h_\ell^{(1)}(kr) \sum_{m=-\ell}^{\ell} Y_{\ell m}^*(\theta', \varphi') \, Y_{\ell m}(\theta, \varphi), \tag{8.24}$$

where $j_\ell(x)$ are spherical Bessel functions and $h_\ell^{(1)}(x)$ are spherical Hankel functions of the first kind. We shall not pursue the investigation of the induction zone further.

8.2 Electric dipole radiation

In the radiation zone, we obtained (8.23)

$$\vec{A}(\vec{r}, t) = \frac{1}{r} e^{i(kr - \omega t)} \sum_{m \geq 0} \frac{(-i\,k)^m}{m!} \int \vec{J}(\vec{r}') \, (\vec{n} \cdot \vec{r}')^m \, d^3\vec{r}', \tag{8.25}$$

The terms in this expansion correspond to the terms in a multipole expansion.

Consider first the $m = 0$ term, for which

$$\vec{A}(\vec{r}, t) = \frac{1}{r} e^{i(kr - \omega t)} \int \vec{J}(\vec{r}') \, d^3\vec{r}', \tag{8.26}$$

This actually corresponds to an electric dipole term. To see this, consider the identity

$$\partial_i(x_j J_i(\vec{r})) = \delta_{ij} J_i(\vec{r}) + x_j \partial_i J_i(\vec{r}) = J_j(\vec{r}) + x_j \, \vec{\nabla} \cdot \vec{J}(\vec{r}). \tag{8.27}$$

166

The integral of the left-hand side over all space gives zero, since it can be turned into a boundary integral over the sphere at infinity (where the localised sources must vanish):

$$\int \partial_i(x_j J_i(\vec{r})) \, d^3\vec{r} = \int_S x_j J_i(\vec{r}) \, dS_i = 0 \,. \tag{8.28}$$

We also have the charge conservation equation

$$\vec{\nabla} \cdot \vec{J}(\vec{r},t) + \frac{\partial \rho(\vec{r},t)}{\partial t} = 0 \,, \tag{8.29}$$

and so with the time dependence $e^{-\mathrm{i}\omega t}$ that we are assuming, this gives

$$\vec{\nabla} \cdot \vec{J}(\vec{r}) = \mathrm{i}\,\omega\,\rho(\vec{r}) = \mathrm{i}\,k\,\rho(\vec{r}) \,. \tag{8.30}$$

Thus we conclude that

$$\int \vec{J}(\vec{r}) \, d^3\vec{r} = -\mathrm{i}\,k \int \vec{r}\,\rho(\vec{r}) \, d^3\vec{r}, \tag{8.31}$$

and so in other words

$$\int \vec{J}(\vec{r}) \, d^3\vec{r} = -\mathrm{i}\,k\,\vec{p}, \tag{8.32}$$

where

$$\vec{p} = \int \vec{r}\,\rho(\vec{r}) \, d^3\vec{r} \tag{8.33}$$

is the electric dipole moment of the charge configuration. Thus we have

$$\vec{A}(\vec{r},t) = -\frac{\mathrm{i}\,k\,\vec{p}}{r}\,e^{\mathrm{i}\,(kr-\omega t)} \,. \tag{8.34}$$

Note that this leading-order term in the expansion of the radiation field corresponds to an electric dipole, and not an electric monopole. The reason for this is that a monopole term would require that the total electric charge in the source region should oscillate in time. This would be impossible, because the total charge in this isolated system must remain constant, by charge conservation.

It is convenient to factor out the time-dependence factor $e^{-\mathrm{i}\omega t}$ that accompanies all the expressions we shall be working with, and to write

$$\vec{A}(\vec{r},t) = \vec{A}(\vec{r})\,e^{-\mathrm{i}\omega t} \,, \quad \vec{B}(\vec{r},t) = \vec{B}(\vec{r})\,e^{-\mathrm{i}\omega t} \,, \quad \vec{E}(\vec{r},t) = \vec{E}(\vec{r})\,e^{-\mathrm{i}\omega t} \,. \tag{8.35}$$

Thus for the electric dipole field we shall have

$$\vec{A}(\vec{r}) = -\frac{\mathrm{i}\,k\vec{p}}{r}\,e^{\mathrm{i}\,kr} \,. \tag{8.36}$$

Then from $\vec{B}(\vec{r}) = \vec{\nabla} \times \vec{A}(\vec{r})$ we find

$$\begin{aligned} B_i &= \epsilon_{ijk}\,\partial_j A_k = -\mathrm{i}\,k\epsilon_{ijk}\,p_k\partial_j\Big(\frac{1}{r}\,e^{\mathrm{i}\,kr}\Big) \,, \\ &= -\mathrm{i}\,k\epsilon_{ijk}p_k\Big(-\frac{x_j}{r^3} + \mathrm{i}\,k\,\frac{x_j}{r^2}\Big)e^{\mathrm{i}\,kr} \,, \end{aligned} \tag{8.37}$$

167

and so

$$\vec{B} = k^2\,(\vec{n} \times \vec{p})\,\frac{e^{i\,kr}}{r}\left(1 + \frac{i}{kr}\right). \tag{8.38}$$

From (8.5) we then have

$$
\begin{aligned}
E_i &= \frac{i}{k}\,\epsilon_{ijk}\,p_m\partial_j\Big[\epsilon_{k\ell m}k^2x_\ell\Big(\frac{1}{r^2} + \frac{i}{kr^3}\Big)e^{i\,kr}\Big], \\
&= i\,k\,(\delta_{i\ell}\delta_{jm} - \delta_{im}\delta_{j\ell})\,p_m\Big[\delta_{j\ell}\Big(\frac{1}{r^2} + \frac{i}{kr^3}\Big) - \frac{2x_jx_\ell}{r^4} - \frac{3i\,x_jx_\ell}{kr^5} + i\,k\frac{x_jx_\ell}{r}\Big(\frac{1}{r^2} + \frac{i}{kr^3}\Big)\Big]e^{i\,kr}, \\
&= \frac{k^2}{r}\,(p_i - n_i\,\vec{n}\cdot\vec{p})\,e^{i\,kr} + \frac{i\,k}{r^2}\,(p_i - 3\vec{n}\cdot\vec{p}\,n_i)\,e^{i\,kr} - \frac{1}{r^3}\,(p_i - 3\vec{n}\cdot\vec{p}\,n_i)\,e^{i\,kr}. \tag{8.39}
\end{aligned}
$$

In 3-vector language, this gives

$$\vec{E} = -k^2\vec{n} \times (\vec{n} \times \vec{p})\,\frac{e^{i\,kr}}{r} + [3(\vec{n}\cdot\vec{p})\,\vec{n} - \vec{p}]\left(\frac{1}{r^3} - \frac{i\,k}{r^2}\right)e^{i\,kr}. \tag{8.40}$$

The reason why we have kept all terms in these expressions for $\vec{B}$ and $\vec{E}$, rather than just the leading-order $1/r$ terms, is that if one makes a multipole expansion, in which the electric dipole contribution we have obtained here is the first term in the series, the expressions are in fact exact to all orders in $1/r$. We shall discuss this in more detail later.

Note that we have $\vec{n}\cdot\vec{B} = 0$ everywhere, but that $\vec{n}\cdot\vec{E} = 0$ only in the radiation zone (i.e. at order $1/r$). In the radiation zone we have

$$\vec{B} = k^2(\vec{n} \times \vec{p})\,\frac{e^{i\,kr}}{r}\,, \qquad \vec{E} = -k^2\vec{n} \times (\vec{n} \times \vec{p})\,\frac{e^{i\,kr}}{r} = -\vec{n} \times \vec{B}\,. \tag{8.41}$$

Note that we have $|\vec{B}| = |\vec{E}|$, as usual for radiation fields.

Since (8.38) and (8.40) are in fact valid everywhere, we can also use these expressions in the static zone (i.e. the near zone). Here, in the regime $kr \ll 1$, we therefore have

$$\vec{B} = i\,k\,(\vec{n} \times \vec{p})\,\frac{1}{r^2}\,, \qquad \vec{E} = [3(\vec{n}\cdot\vec{p})\,\vec{n} - \vec{p}]\,\frac{1}{r^3}\,. \tag{8.42}$$

The electric field here is precisely like that of a static electric dipole, except that it is oscillating in time. Note that in the near zone we have $|\vec{B}| \sim (kr)\,|\vec{E}|$, which means $|\vec{B}| \ll |\vec{E}|$.

Returning now to the radiation zone, we may calculate the radiated power in the usual way, using the Poynting vector. In particular, we saw previously that with the electric and magnetic fields written in the complex notation, the time average of the Poynting flux is given by

$$\langle\vec{S}\rangle = \frac{1}{8\pi}\,\vec{E} \times \vec{B}^*\,. \tag{8.43}$$

168

Then the power radiated into the solid angle $d\Omega$ is given by

$$
\begin{aligned}
dP &= \langle \vec{S} \rangle \cdot \vec{n}\, r^2 d\Omega\,,\\
&= \frac{1}{8\pi}\left[(-\vec{n}\times\vec{B})\times\vec{B}^{\,*}\right]\cdot\vec{n}\,r^2 d\Omega\,,\\
&= \frac{1}{8\pi}\,|\vec{B}|^2\,r^2 d\Omega\,.
\end{aligned}
\tag{8.44}
$$

From (8.41) we therefore have

$$
\frac{dP}{d\Omega} = \frac{k^4}{8\pi}\,|\vec{n}\times\vec{p}|^2 = \frac{k^4}{8\pi}\left(|\vec{p}|^2 - (\vec{n}\cdot\vec{p})^2\right).
\tag{8.45}
$$

If we take θ to be the angle between $\vec{p}$ and $\vec{n}$, so that $\vec{n}\cdot\vec{p} = p\cos\theta$, then this gives

$$
\frac{dP}{d\Omega} = \frac{k^4}{8\pi}\,|\vec{p}|^2\sin^2\theta\,.
\tag{8.46}
$$

Since $d\Omega = \sin\theta\,d\theta\,d\varphi$, the total power radiated by the oscillating dipole is then given by

$$
P = \int \frac{dP}{d\Omega}d\Omega = 2\pi\,\frac{k^4}{8\pi}\,|\vec{p}|^2\int_0^\pi \sin^3\theta\,d\theta = \tfrac{1}{3}k^4\,|\vec{p}|^2\,.
\tag{8.47}
$$

As a concrete example, consider a dipole antenna comprising two thin conducting rods running along the z axis, meeting (but not touching) at the origin, and extending to $z = \pm\tfrac{1}{2}d$ respectively. The antenna is driven at the centre ($z = 0$) by an alternating current source with angular frequency ω. The current will fall off as a function of z, becoming zero at the tips of the antenna at $z = \pm\tfrac{1}{2}d$. A reasonable approximation, in the regime we are considering here where $kd << 1$, is that this fall-off is linear in z.[30] Thus we may assume

$$
I(z,t) = I(z)e^{-\mathrm{i}\omega t} = I_0\left(1 - \frac{2|z|}{d}\right)e^{-\mathrm{i}\omega t}\,.
\tag{8.48}
$$

The equation of charge conservation, $\vec{\nabla}\cdot\vec{J} + \partial\rho/\partial t = 0$ then allows us to solve for the charge density. The current (8.48) is essentially confined to the line $x = y = 0$, since we are assuming the conducting rods that form the antenna are thin. Thus really, we have

$$
\vec{J}(\vec{r},t) = I(z,t)\,\delta(x)\delta(y)\,.
\tag{8.49}
$$

Similarly, the charge density will be given by

$$
\rho(\vec{r},t) = \lambda(z,t)\,\delta(x)\delta(y)\,,
\tag{8.50}
$$

---

[30]It should in principle be possible to *derive* the profile function for the current in the antenna as a function of z. This would involve a complicated time-dependent boundary-value problem, and as far as I know it is not tractable analytically.

where $\lambda(z,t)$ is the charge per unit length in the rods. The charge conservation equation therefore becomes

$$\frac{\partial I(z,t)}{\partial z} + \frac{\partial \lambda(z,t)}{\partial t} = 0 \,, \tag{8.51}$$

and so, in view of the time dependence, which implies also $\lambda(z,t) = \lambda(z)e^{-\mathrm{i}\omega t}$, we have

$$\frac{\partial I(z)}{\partial z} - \mathrm{i}\,\omega\lambda(z) = 0 \,. \tag{8.52}$$

Thus we shall have

$$\lambda(z) = -\frac{\mathrm{i}}{\omega}\frac{\partial I(z)}{\partial z} = -\frac{\mathrm{i}}{\omega}\,I_0\,\frac{\partial}{\partial z}\Big(1 - \frac{2|z|}{d}\Big) \,. \tag{8.53}$$

This implies

$$\begin{aligned}
\lambda(z) &= \frac{2\mathrm{i}\,I_0}{\omega d}\,, & z &> 0\,,\\
\lambda(z) &= -\frac{2\mathrm{i}\,I_0}{\omega d}\,, & z &< 0\,.
\end{aligned} \tag{8.54}$$

The dipole moment $\vec{p}$ is directed along the z axis, $\vec{p} = (0,0,p)$, and is given by

$$p = \int_{-d/2}^{d/2} z\lambda(z)dz = \frac{2\mathrm{i}\,I_0}{\omega d}\int_0^{d/2} zdz - \frac{2\mathrm{i}\,I_0}{\omega d}\int_{-d/2}^0 zdz = \frac{\mathrm{i}\,I_0 d}{2\omega}\,. \tag{8.55}$$

From (8.46), we therefore find that the power per unit solid angle is given by

$$\frac{dP}{d\Omega} = \frac{k^4|\vec{p}|^2}{8\pi}\,\sin^2\theta = \frac{I_0^2(kd)^2}{32\pi}\,\sin^2\theta \,, \tag{8.56}$$

where θ is the angle between $\vec{n} = \vec{r}/r$ and the z axis. The total radiated power is therefore given by

$$P = \tfrac{1}{12}I_0^2\,(kd)^2 \,. \tag{8.57}$$

8.3 Higher multipoles

As mentioned previously, in a multipole expansion we can obtain exact expressions, term by term, for the electric and magnetic fields. To do this, we go back to the general integral expression

$$\vec{A}(\vec{r}) = \int \vec{J}(\vec{r}')\,\frac{e^{\mathrm{i}\,k|\vec{r}-\vec{r}'|}}{|\vec{r}-\vec{r}'|}\,d^3\vec{r}' \,, \tag{8.58}$$

giving $\vec{A}(\vec{r},t) = \vec{A}(\vec{r})\,e^{-\mathrm{i}\omega t}$. Let

$$\frac{1}{r}\,e^{\mathrm{i}\,kr} = f(\vec{r}) = f(r) \,. \tag{8.59}$$

(Note that $f(\vec{r}) = f(r)$, i.e. it depends only on the magnitude of $\vec{r}$.) It follows that

$$\frac{e^{\mathrm{i}\,k|\vec{r}-\vec{r}'|}}{|\vec{r}-\vec{r}'|} = f(\vec{r}-\vec{r}') \,, \tag{8.60}$$

170

which we can therefore express as the Taylor series

$$
\begin{aligned}
f(\vec{r} - \vec{r}\,') &= f(\vec{r}) - x_i' \, \partial_i f(\vec{r}) + \frac{1}{2!} x_i' x_j' \, \partial_i \partial_j f(\vec{r}) + \cdots \\
&= f(r) - x_i' \, \partial_i f(r) + \frac{1}{2!} x_i' x_j' \, \partial_i \partial_j f(r) + \cdots , \\
&= f(r) - x_i' \, (\partial_i r) \, f'(r) + \tfrac{1}{2} x_i' x_j' \left[(\partial_i \partial_j r) f'(r) + (\partial_i r)(\partial_j r) \, f''(r) \right] + \cdots (8.61)
\end{aligned}
$$

Thus we find

$$
\frac{e^{\mathrm{i}\,k|\vec{r} - \vec{r}\,'|}}{|\vec{r} - \vec{r}\,'|} = \frac{1}{r} \, e^{\mathrm{i}\,kr} + \left(\frac{1}{r^2} - \frac{\mathrm{i}\,k}{r} \right) (\vec{n} \cdot \vec{r}\,') \, e^{\mathrm{i}\,kr} + \cdots . \tag{8.62}
$$

The first term in this series gives the electric dipole contribution that we found in the previous section, in (8.26). The second term gives contributions from an electric quadrupole term and a magnetic dipole term. This gives

$$
\vec{A}(\vec{r}) = e^{\mathrm{i}\,kr} \left(\frac{1}{r^2} - \frac{\mathrm{i}\,k}{r} \right) \int (\vec{n} \cdot \vec{r}\,') \, \vec{J}(\vec{r}\,') \, d^3 \vec{r}\,' . \tag{8.63}
$$

In order to interpret this expression, we need to manipulate the integrand a bit. Its i'th component is given by

$$
\begin{aligned}
n_j \, x_j' \, J_i &= \tfrac{1}{2} (J_i \, x_j' - J_j \, x_i') n_j + \tfrac{1}{2} (J_i \, x_j' + J_j \, x_i') n_j , \\
&= \tfrac{1}{2} \epsilon_{ijk} \, \epsilon_{\ell m k} \, J_\ell \, x_m' \, n_j + \tfrac{1}{2} (J_i \, x_j' + J_j \, x_i') n_j , \\
&= -\epsilon_{ijk} \, n_j \mathcal{M}_k + \tfrac{1}{2} (J_i \, x_j' + J_j \, x_i') n_j , \tag{8.64}
\end{aligned}
$$

where

$$
\mathcal{M}_i = \tfrac{1}{2} \epsilon_{ijk} \, x_j' \, J_k , \qquad \text{i.e.} \qquad \vec{\mathcal{M}} = \tfrac{1}{2} \vec{r}\,' \times \vec{J}(\vec{r}\,') \tag{8.65}
$$

is the *magnetisation* resulting from the current density $\vec{J}$.

The remaining term in (8.64), i.e. the symmetric term $\tfrac{1}{2} (J_i \, x_j' + J_j \, x_i') n_j$, can be analysed as follows. Consider

$$
\begin{aligned}
\partial_k (x_i \, x_j \, n_j \, J_k) &= \delta_{ik} \, x_j \, n_j \, J_k + \delta_{jk} \, x_i \, n_j \, J_k + x_i \, x_j \, n_j \, \partial_k J_k , \\
&= (x_i \, J_j + x_j \, J_i) n_j + \mathrm{i}\, x_i \, x_j \, n_j \, \omega \, \rho . \tag{8.66}
\end{aligned}
$$

Integrating this over all space, the left-hand side can be turned into a surface integral over the sphere at infinity, which therefore gives zero. Thus we conclude that

$$
\int (x_i \, J_j + x_j \, J_i) n_j d^3 \vec{r} = -\mathrm{i}\,\omega \int x_i \, x_j \, n_j \, \rho \, d^3 \vec{r} . \tag{8.67}
$$

The upshot is that

$$
\int (\vec{n} \cdot \vec{r}\,') \, \vec{J}(\vec{r}\,') \, d^3 \vec{r}\,' = -\vec{n} \times \int \vec{\mathcal{M}} \, d^3 \vec{r}\,' - \frac{\mathrm{i}\,\omega}{2} \int \vec{r}\,' (\vec{n} \cdot \vec{r}\,') \, \rho(\vec{r}\,') \, d^3 \vec{r}\,' . \tag{8.68}
$$

Defining the magnetic dipole moment $\vec{m}$ by

$$\vec{m} = \int \vec{\mathcal{M}} \, d^3\vec{r}\,' = \tfrac{1}{2} \int \vec{r}\,' \times \vec{J}(\vec{r}\,') \, d^3\vec{r}\,' \,, \tag{8.69}$$

we conclude that

$$\vec{A}(\vec{r}) = e^{\mathrm{i}\,kr} \left(\frac{\mathrm{i}\,k}{r} - \frac{1}{r^2} \right) \vec{n} \times \vec{m} + \frac{\mathrm{i}\,k}{2} e^{\mathrm{i}\,kr} \left(\frac{\mathrm{i}\,k}{r} - \frac{1}{r^2} \right) \int \vec{r}\,' \left(\vec{n} \cdot \vec{r}\,' \right) \rho(\vec{r}\,') \, d^3\vec{r}\,' \,. \tag{8.70}$$

8.3.1   Magnetic dipole term

Consider the magnetic dipole term in (8.70) first:

$$\vec{A}(\vec{r}) = e^{\mathrm{i}\,kr} \left(\frac{\mathrm{i}\,k}{r} - \frac{1}{r^2} \right) \vec{n} \times \vec{m} \,. \tag{8.71}$$

Let

$$f \equiv e^{\mathrm{i}\,kr} \left(\frac{\mathrm{i}\,k}{r^2} - \frac{1}{r^3} \right) , \tag{8.72}$$

so $\vec{A} = rf\vec{n} \times \vec{m} = f\vec{r} \times \vec{m}$. Then from $\vec{B} = \vec{\nabla} \times \vec{A}$ we shall have

$$
\begin{aligned}
B_i &= \epsilon_{ijk}\partial_j A_k = \epsilon_{ijk}\,\epsilon_{k\ell m}\,\partial_j(fx_\ell)m_m \,, \\
&= \epsilon_{ijk}\,\epsilon_{k\ell m}\,(f' x_\ell \frac{x_j}{r} + f\delta_{j\ell}) \,, \\
&= (\delta_{i\ell}\delta_{jm} - \delta_{im}\delta_{j\ell})(f' x_\ell \frac{x_j}{r} + f\delta_{j\ell}) \,, \\
&= rf' n_i \vec{n} \cdot \vec{m} - rf' m_i - 2fm_i \,.
\end{aligned}
\tag{8.73}
$$

From (8.72) we have

$$f' = e^{\mathrm{i}\,kr} \left(-\frac{k^2}{r^2} - \frac{3\mathrm{i}\,k}{r^3} + \frac{3}{r^4} \right) = -\frac{k^2}{r^2} e^{\mathrm{i}\,kr} - \frac{3f}{r} \,, \tag{8.74}$$

and so we find

$$\vec{B} = -k^2 \vec{n} \times (\vec{n} \times \vec{m}) \frac{e^{\mathrm{i}\,kr}}{r} + [3\vec{n}\,(\vec{n} \cdot \vec{m}) - \vec{m}] \left(\frac{1}{r^3} - \frac{\mathrm{i}\,k}{r^2} \right) e^{\mathrm{i}\,kr} \,. \tag{8.75}$$

Note that this is identical to the expression (8.40) for the electric field of an electric dipole source, with the electric dipole $\vec{p}$ replaced by the magnetic dipole, and the electric field replaced by the magnetic field:

$$\vec{p} \longrightarrow \vec{m} \,, \qquad \vec{E} \longrightarrow \vec{B} \,. \tag{8.76}$$

The electric field of the magnetic dipole can be obtained from (8.5). However, a simpler way to find it here is to note that from the Maxwell equation $\vec{\nabla} \times \vec{E} = -\partial\vec{B}/\partial t$ we have

$$\vec{\nabla} \times \vec{E} = \mathrm{i}\,\omega\,\vec{B} = \mathrm{i}\,k\,\vec{B} \,, \tag{8.77}$$

and so

$$\vec{B} = -\frac{\mathrm{i}}{k}\,\vec{\nabla} \times \vec{E}\,. \tag{8.78}$$

Now, we already saw from the calculations for the electric dipole that when the $\vec{B}$ field (8.38) is substituted into (8.5), it gives rise to the $\vec{E}$ field given in (8.40). As we just noted, in the present magnetic dipole case, the expression for the $\vec{B}$ field is just like the expression for the $\vec{E}$ field in the electric dipole case, and we already know that in the electric case, the $\vec{B}$ field is given by (8.38). Therefore, we can conclude that in the present magnetic case, the $\vec{E}$ field that would yield, using (8.78), the result (8.75) for the $\vec{B}$ field will be just the negative of the expression for $\vec{B}$ in the electric case (with $\vec{p}$ replaced by $\vec{m}$). (The reason for the minus sign is that (8.78) has a minus sign, as compared with (8.5), under the exchange of $\vec{E}$ and $\vec{B}$.) Thus the upshot is that the electric field for the magnetic dipole radiation will be given by

$$\vec{E} = -k^2(\vec{n} \times \vec{m})\,\frac{e^{\mathrm{i}\,kr}}{r}\left(1 + \frac{\mathrm{i}}{kr}\right). \tag{8.79}$$

This result can alternatively be verified (after a rather involved calculation) by directlty substituting (8.75) into (8.5).[31]

An observation from the calculations of the electric and magnetic fields for electric dipole radiation and magnetic dipole radiation is that there is a discrete symmetry under which the two situations interchange:

$$
\begin{aligned}
\vec{p} &\longrightarrow \vec{m} \\
\vec{E} &\longrightarrow \vec{B} \\
\vec{B} &\longrightarrow -\vec{E}
\end{aligned}
\tag{8.80}
$$

This is an example of what is known as "electric/magnetic duality" in Maxwell's equations.

8.3.2 Electric quadrupole term

We now return to the electric quadrupole term in (8.70), namely

$$\vec{A}(\vec{r}) = \frac{\mathrm{i}\,k}{2}\,e^{\mathrm{i}\,kr}\left(\frac{\mathrm{i}\,k}{r} - \frac{1}{r^2}\right)\int \vec{r}\,'\,(\vec{n}\cdot\vec{r}\,')\,\rho(\vec{r}\,')\,d^3\vec{r}\,'\,. \tag{8.81}$$

---

[31]The only "gap" in the simple argument we just presented is that any other vector $\vec{E}\,' = \vec{E} + \vec{\nabla}h$ would also give the same $\vec{B}$ field when plugged into (8.78), where h was an arbitrary function. However, we know that $\vec{\nabla}\cdot\vec{E}$ should vanish (we are in a region away from sources), and it is obvious almost by inspection that the answer given in (8.79) satisfies this condition. Thus if we had arrived at the wrong answer for $\vec{E}$, it could be wrong only by a term $\vec{\nabla}h$ where $\nabla^2 h = 0$. There is no such function with an exponential factor $e^{\mathrm{i}\,kr}$, and so there is no possibility of our answer (8.79) being wrong. If any doubts remain, the reader is invited to substitute (8.75) into (8.5) to verify (8.79) directly.

For simplicity, we shall keep only the leading-order radiation term in this expression,

$$\vec{A}(\vec{r}) = -\tfrac{1}{2}k^2 \frac{e^{ikr}}{r} \int \vec{r}'\, (\vec{n}\cdot\vec{r}')\, \rho(\vec{r}')\, d^3\vec{r}'\,. \tag{8.82}$$

and furthermore when calculating the $\vec{B}$ and $\vec{E}$ fields, we shall keep only the leading-order $1/r$ terms that come from the derivatives hitting e^{ikr}. Thus, from $\vec{B} = \vec{\nabla}\times\vec{A}$ we shall have

$$\begin{aligned}
\vec{B} &= -\tfrac{1}{2}k^2\,(i\,k)\,\vec{n}\times\frac{e^{ikr}}{r}\int \vec{r}'\,(\vec{n}\cdot\vec{r}')\,\rho(\vec{r}')\,d^3\vec{r}'\,,\\
&= -\frac{i\,k^3}{2}\frac{e^{ikr}}{r}\int(\vec{n}\times\vec{r}')(\vec{n}\cdot\vec{r}')\,\rho(\vec{r}')\,d^3\vec{r}'\,.
\end{aligned} \tag{8.83}$$

This radiation field can therefore be written simply as

$$\vec{B} = i\,k\,\vec{n}\times\vec{A}\,. \tag{8.84}$$

In fact, in any expression where we keep only the leading-order term in which the derivative lands on e^{ikr}, we shall have the rule

$$\vec{\nabla} \longrightarrow i\,k\,\vec{n}\,. \tag{8.85}$$

For the electric field, we have, using (8.5) and (8.85),

$$\vec{E} = \frac{i}{k}\vec{\nabla}\times\vec{B} = -\vec{n}\times\vec{B} = -i\,k\vec{n}\times(\vec{n}\times\vec{A})\,. \tag{8.86}$$

The electric quadrupole moment tensor Q_{ij} is defined by

$$Q_{ij} = \int (3x_i x_j - r^2\delta_{ij})\,\rho(\vec{r})\,d^3\vec{r}\,. \tag{8.87}$$

Define the vector $\vec{Q}(\vec{n})$, whose components $Q(\vec{n})_i$ are given by

$$Q(\vec{n})_i \equiv Q_{ij}\,n_j\,. \tag{8.88}$$

Consider the expression $\tfrac{1}{3}\vec{n}\times\vec{Q}(\vec{n})$. We shall have

$$\begin{aligned}
[\tfrac{1}{3}\vec{n}\times\vec{Q}(\vec{n})]_i &= \tfrac{1}{3}\epsilon_{ijk}\,n_j Q_{k\ell}\,n_\ell\,,\\
&= \tfrac{1}{3}\epsilon_{ijk}\,n_j n_\ell \int (3x_k x_\ell - r^2\delta_{k\ell})\,\rho(\vec{r})\,d^3\vec{r}\,,\\
&= \int (\vec{n}\times\vec{r})_i\,(\vec{n}\cdot\vec{r})\,\rho(\vec{r})\,d^3\vec{r}\,,
\end{aligned} \tag{8.89}$$

(since the trace term gives zero). This implies that the expression (8.83) for the electric-quadrupole $\vec{B}$ field can be written as

$$\vec{B} = -\frac{i\,k^3}{6r}\,e^{ikr}\,\vec{n}\times\vec{Q}(\vec{n})\,. \tag{8.90}$$

174

Since we have $\vec{E} = \vec{B} \times \vec{n}$ (see (8.86)), it follows that the time-averaged power per unit solid angle will be given by

$$
\begin{aligned}
\frac{dP}{d\Omega} &= \frac{1}{8\pi} (\vec{E} \times \vec{B}^*) \cdot \vec{n} \, r^2 \,, \\
&= \frac{k^6}{288\pi} |(\vec{n} \times \vec{Q}(\vec{n})) \times \vec{n}|^2 = \frac{k^6}{288\pi} (|\vec{Q}(\vec{n})|^2 - |\vec{n} \cdot \vec{Q}(\vec{n})|^2) \,,
\end{aligned}
\tag{8.91}
$$

and so

$$
\frac{dP}{d\Omega} = \frac{k^6}{288\pi} |\vec{n} \times \vec{Q}(\vec{n})|^2 \,.
\tag{8.92}
$$

Written using indices, this is therefore

$$
\frac{dP}{d\Omega} = \frac{k^6}{288\pi} (Q_{ki} Q^*_{kj} n_i n_j - Q_{ij} Q^*_{k\ell} n_i n_j n_k n_\ell) \,.
\tag{8.93}
$$

As always, having obtained an expression for the power radiated per unit solid angle, it is natural to integrate this up over the sphere, in order to obtain the total radiated power. In this case, we shall need to evaluate

$$
\int n_i n_j \, d\Omega \,, \qquad \text{and} \qquad \int n_i n_j n_k n_\ell \, d\Omega \,.
\tag{8.94}
$$

One way to do this is to parameterise the unit vector $\vec{n}$ in terms of spherical polar angles (θ, φ) in the usual way,

$$
\vec{n} = (\sin\theta \cos\varphi, \sin\theta \sin\varphi, \cos\theta) \,,
\tag{8.95}
$$

and slog out the integrals with $d\Omega = \sin\theta d\theta d\varphi$.

A more elegant way to evaluate the integrals in (8.94) is as follows. For the first integral, we note that the answer, whatever it is, must be a symmetric 2-index tensor. It must also be completely isotropic, since by the time we have integrated over all solid angles it is not possible for the result to be "biased" so that it favours any direction in space over any other. There is only one possibility for the symmetric isotropic tensor; it must be a constant multiple of the Krönecker delta, so we must have

$$
\int n_i n_j \, d\Omega = c \delta_{ij} \,.
\tag{8.96}
$$

The constant c can be determined by taking the trace, and using $n_i n_i = 1$:

$$
4\pi = \int d\Omega = 3c \,,
\tag{8.97}
$$

and so we have

$$
\int n_i n_j \, d\Omega = \frac{4\pi}{3} \delta_{ij} \,.
\tag{8.98}
$$

175

In case one doubted this result, it is not too hard in this case to confirm the result by evaluating all the integrals explicitly using (8.95).

Turning now to the second integral in (8.94), we can use a similar argument. The answer must be a 4-index totally symmetric isotropic tensor. In fact the *only* symmetric isotropic tensors are those that can be made by taking products of Krönecker deltas, and so in this case it must be that

$$\int n_i n_j n_k n_\ell \, d\Omega = b \left(\delta_{ij}\delta_{k\ell} + \delta_{ik}\delta_{j\ell} + \delta_{i\ell}\delta_{jk} \right), \tag{8.99}$$

for some constant b. We can determine the constant by multiplying both sides by $\delta_{ij}\delta_{k\ell}$, giving

$$4\pi = \int d\Omega = (9 + 3 + 3)b = 15b, \tag{8.100}$$

and so

$$\int n_i n_j n_k n_\ell \, d\Omega = \frac{4\pi}{15} \left(\delta_{ij}\delta_{k\ell} + \delta_{ik}\delta_{j\ell} + \delta_{i\ell}\delta_{jk} \right). \tag{8.101}$$

With these results we shall have from (8.93) that

$$\begin{aligned}
P &= \int \frac{dP}{d\Omega} \, d\Omega = \frac{k^6}{288\pi} \left[Q_{ki}Q_{kj}^* \int n_i n_j \, d\Omega - Q_{ij}Q_{k\ell}^* \int n_i n_j n_k n_\ell \, d\Omega \right], \\
&= \frac{k^6}{288\pi} \left[\frac{4\pi}{3} Q_{ki}Q_{kj}^* \delta_{ij} - \frac{4\pi}{15} Q_{ij}Q_{k\ell}^* (\delta_{ij}\delta_{k\ell} + \delta_{ik}\delta_{j\ell} + \delta_{i\ell}\delta_{jk}) \right], \\
&= \frac{k^6}{216} \left[Q_{ij}Q_{ij}^* - \tfrac{2}{5}Q_{ij}Q_{ij}^* - \tfrac{1}{5}Q_{ii}Q_{jj}^* \right], \\
&= \frac{k^6}{360} Q_{ij}Q_{ij}^*. \tag{8.102}
\end{aligned}$$

(Recall that Q_{ij} is symmetric and traceless.)

Since the quadrupole moment tensor Q_{ij} is symmetric, it is always possible to choose an orientation for the Cartesian coordinate system such that Q_{ij} becomes diagonal. (This is because the matrix U that diagonalises Q, $Q \to Q_{\text{diag}} = U^T Q U$ is itself orthogonal, $U^T U = \mathbb{1}$, and therefore the diagonalisation is achieved by an orthogonal transformation of the coordinates.) Thus, having chosen an appropriate orientation for the Cartesian axes, we can assume that

$$Q_{ij} = \begin{pmatrix} Q_1 & 0 & 0 \\ 0 & Q_2 & 0 \\ 0 & 0 & Q_3 \end{pmatrix}, \quad \text{where} \quad Q_1 + Q_2 + Q_3 = 0. \tag{8.103}$$

The expression (8.93) for the angular power distribution will give

$$\frac{dP}{d\Omega} = \frac{k^6}{288\pi} \left(Q_1^2 n_1^2 + Q_2^2 n_2^2 + Q_3^2 n_3^2 - (Q_1 n_1^2 + Q_2 n_2^2 + Q_3 n_3^2)^2 \right). \tag{8.104}$$

176

One can substitute (8.95) into this in order to obtain an explicit expression for the $dP/d\Omega$ in terms of spherical polar angles (θ, φ).

Consider for simplicity the special case where $Q_1 = Q_2$. This means that there is an axial symmetry around the z axis, and also we shall have

$$Q_1 = Q_2 = -\tfrac{1}{2}Q_3\,. \tag{8.105}$$

Substituting (8.95) and (8.105) into (8.104), we obtain

$$\frac{dP}{d\Omega} = \frac{k^6\,Q_3^2}{128\pi}\,\sin^2\theta\,\cos^2\theta = \frac{k^6\,Q_3^2}{512\pi}\,\sin^2 2\theta\,. \tag{8.106}$$

This is indeed, as expected, azimuthally symmetric (it does not depend on φ). It describes a quadrafoil-like power distribution, with four lobes, unlike the figure-of-eight power distribution of the electric dipole radiation. Note also that its frequency dependence is proportional to ω^6 $(= k^6)$, unlike the electric dipole radiation that is proportional to ω^4. A plot of the power distribution for quadrupole radiation is given in Figure 3 below.
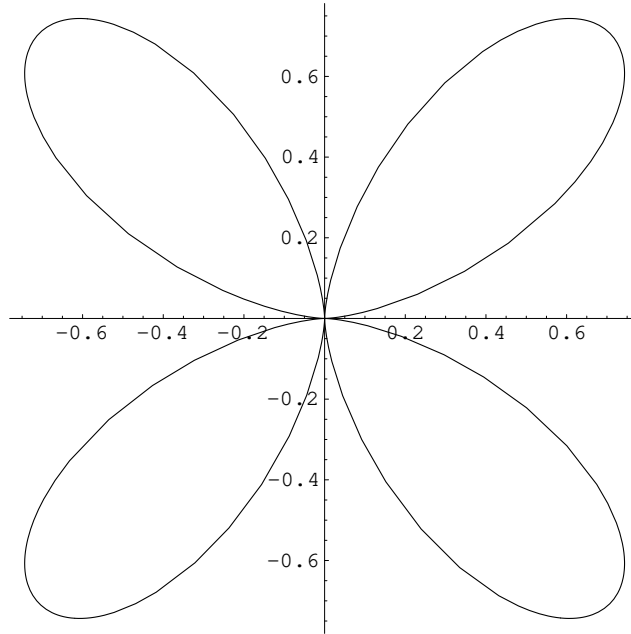

Figure 6: The angular power distribution for electric quadrupole radiation

177

8.4 Linear antenna

In the later part of section 8.2, we considered a centre-fed dipole antenna. In that section we made the assumption that the wavelength of the electromagnetic radiation was very large compared with the length of the dipole, i.e. that $kd << 1$. In that limit, one could assume to a good approximation that the current in each arm of the dipole antenna fell off in a linear fashion as a function of z (the axis along which the dipole is located). Thus, with the dipole arms extanding over the intervals

$$-\tfrac{1}{2}d \le z < 0 \qquad \text{and} \qquad 0 < z \le \tfrac{1}{2}d \,, \tag{8.107}$$

we assumed there that the current in each arm was proportional to $(d/2 - |z|)$.

In this section, we shall consider the case where the dipole arms of not assumed to be short campared to the wavelength. Under these circumstances, it can be shown that the current distribution in the dipole arms takes the form

$$\vec{J}(\vec{r}, t) = I_0 \, \sin k(d/2 - |z|) \, e^{-\mathrm{i}\omega t} \, \delta(x)\delta(y) \, \vec{Z} \,, \qquad |z| \le \tfrac{1}{2}d \,, \tag{8.108}$$

where $\vec{Z} = (0, 0, 1)$ is the unit vector along the z-axis, which is the axis along which the dipole is located.

We then have $\vec{A}(\vec{r}, t) = \vec{A}(\vec{r}) \, e^{-\mathrm{i}\omega t}$, where

$$\vec{A}(\vec{r}) = \int \frac{\vec{J}(\vec{r}', t - |\vec{r} - \vec{r}'|) \, d^3\vec{r}'}{|\vec{r} - \vec{r}'|} \,. \tag{8.109}$$

Thus in the radiation zone, with $|\vec{r} - \vec{r}'| \approx r - \vec{n} \cdot \vec{r}'$ as usual, we therefore have

$$
\begin{aligned}
\vec{A}(\vec{r}) &\approx \vec{Z} \frac{I_0 \, e^{\mathrm{i}kr}}{r} \int_{-d/2}^{d/2} \sin k(d/2 - |z|) \, e^{-\mathrm{i}kz\cos\theta} \, dz \,, \\
&= \vec{Z} \frac{2I_0 \, e^{\mathrm{i}kr}}{r} \frac{\cos(\tfrac{1}{2}kd\cos\theta) - \cos(\tfrac{1}{2}kd)}{\sin^2\theta}
\end{aligned}
\tag{8.110}
$$

As we saw earlier, the magnetic field is given by $\mathrm{i}\,k\,\vec{n} \times \vec{A}$ in the radiation zone, and $\vec{E} = -\vec{n} \times \vec{B}$. Therefore the radiated power per unit solid angle is given by

$$\frac{dP}{d\Omega} = \frac{r^2}{8\pi} \, |\vec{E} \times \vec{B}^*|^2 = \frac{r^2}{8\pi} \, |(\vec{B} \cdot \vec{B}^*) \, \vec{n}|^2 = \frac{r^2}{8\pi} \, |\vec{B}|^2 \,. \tag{8.111}$$

Here we have

$$|\vec{B}|^2 = |\mathrm{i}\,k\vec{n} \times \vec{A}|^2 = k^2(|\vec{A}|^2 - (\vec{n} \cdot \vec{A})^2) = k^2|\vec{A}|^2 \sin^2\theta \,, \tag{8.112}$$

since $\vec{n} \cdot \vec{Z} = \cos\theta$, and so the radiated power per unit solid angle is given by

$$\frac{dP}{d\Omega} = \frac{I_0^2}{2\pi} \left[\frac{\cos(\tfrac{1}{2}kd\cos\theta) - \cos(\tfrac{1}{2}kd)}{\sin\theta} \right]^2 \,. \tag{8.113}$$

We can now consider various special cases:

8.4.1 $kd << 1$:

In this case, we can make Taylor expansions of the trigonometric functions in the numerator in (8.113), leading to

$$
\begin{aligned}
\frac{dP}{d\Omega} &\approx \frac{I_0^2}{2\pi} \left[\frac{1 - \frac{1}{2}(\frac{1}{2}kd)^2 \cos^2\theta - 1 - \frac{1}{2}(\frac{1}{2}kd)^2}{\sin\theta} \right]^2, \\
&= \frac{I_0^2}{2\pi} \left(\frac{\frac{1}{2}(\frac{1}{2}kd)^2 \sin^2\theta}{\sin\theta} \right)^2, \\
&= \frac{I_0^2 (kd)^2 \sin^2\theta}{128\pi}.
\end{aligned}
\tag{8.114}
$$

This agrees with the result (8.56), after making allowance for the fact that the current in the calculation leading to (8.56) was twice as large as the current in the present calculation.

8.4.2 $kd = \pi$:

In this case, each arm of the dipole has a length equal to $\frac{1}{4}$ of the wavelength, and so $I(z) = I_0 \sin\frac{1}{2}\pi(1 - 2|z|/d)$. In this case, (8.113) becomes

$$
\frac{dP}{d\Omega} = \frac{I_0^2}{2\pi} \frac{\cos^2(\frac{1}{2}\pi\cos\theta)}{\sin^2\theta}.
\tag{8.115}
$$

8.4.3 $kd = 2\pi$:

In this case, each dipole arm has a length equal to $\frac{1}{2}$ of the wavelength, and $I(z) = I_0 \sin\pi(1 - 2|z|/d)$. In this case (8.113) becomes

$$
\frac{dP}{d\Omega} = \frac{I_0^2}{2\pi} \frac{\cos^4(\frac{1}{2}\pi\cos\theta)}{\sin^2\theta}.
\tag{8.116}
$$

9 Electromagnetism and Quantum Mechanics

9.1 The Schrödinger equation and gauge transformations

We saw at the end of chapter 2, in equation (2.108), that in the non-relativistic limit the Hamiltonian describing a particle of mass m and charge e in the presence of electromagnetic fields given by potentials ϕ and $\vec{A}$ is

$$
H = \frac{1}{2m} (\pi_i - eA_i)^2 + e\phi,
\tag{9.1}
$$

where π_i is the canonical 3-momentum. In quantum mechanics, we the standard prescription for writing down the Schrödinger equation for the wavefunction ψ describing the particle is to interpret π_i as an operator, and to write

$$
H\psi = i\hbar \frac{\partial\psi}{\partial t}.
\tag{9.2}
$$

In the position representation we shall have

$$\pi_i = -\mathrm{i}\,\hbar\,\partial_i\,, \qquad \text{or} \qquad \vec{\pi} = -\mathrm{i}\,\hbar\,\vec{\nabla}\,. \tag{9.3}$$

Thus the Schrödinger equation for a particle of mass m and charge e in an electromagnetic field is

$$-\frac{\hbar^2}{2m}\left(\vec{\nabla} - \frac{\mathrm{i}\,e}{\hbar}\,\vec{A}\right)^2 \psi + e\phi\,\psi = \mathrm{i}\,\hbar\,\frac{\partial\psi}{\partial t}\,. \tag{9.4}$$

The Schrödinger equation (9.4) is written in terms of the scalar and vector potentials ϕ and $\vec{A}$ that describe the electromagnetic field. Thus, if we merely perform a gauge transformation

$$\vec{A} \longrightarrow \vec{A}' = \vec{A} + \vec{\nabla}\lambda\,, \qquad \phi \longrightarrow \phi' = \phi - \frac{\partial\lambda}{\partial t}\,, \tag{9.5}$$

then the Schrödinger equation will change its form. On the other hand, we expect that the physics should be unaltered by a mere gauge transformation, since this leaves the physically-observable electric and magnetic fields unchanged. It turns out that we should simultaneously perform the very specific spacetime-dependent phase transformation on the wavefunction ψ,

$$\psi \longrightarrow \psi' = e^{\mathrm{i}\,e\lambda/\hbar}\,\psi\,. \tag{9.6}$$

Then the Schrödinger equation expressed entirely in terms of the primed quantities (i.e. wavefunction ψ' and electromagnetic potentials ϕ' and $\vec{A}'$) will take the identical form to the original unprimed equation (9.4). Thus, we may say that the Schrödinger equation transforms *covariantly* under gauge transformations.

To see the details of how this works, it is useful first to define what are called *covariant derivatives*. We this both for the three spatial derivatives, and also for the time derivative. Thus we define

$$D_i \equiv \partial_i - \frac{\mathrm{i}\,e}{\hbar}\,A_i\,, \qquad D_0 \equiv \frac{\partial}{\partial t} + \frac{\mathrm{i}\,e}{\hbar}\,\phi\,. \tag{9.7}$$

Note that the original Schrödinger equation (9.4) is now written simply as

$$-\frac{\hbar^2}{2m}\,D_i D_i\psi - \mathrm{i}\,\hbar\,D_0\psi = 0\,. \tag{9.8}$$

Next, perform the transformations

$$\begin{aligned}
\vec{A} &\longrightarrow \vec{A}' = \vec{A} + \vec{\nabla}\lambda\,, \qquad \phi \longrightarrow \phi' = \phi - \frac{\partial\lambda}{\partial t}\,, \\
\psi &\longrightarrow \psi' = e^{\mathrm{i}\,e\lambda/\hbar}\,\psi
\end{aligned} \tag{9.9}$$

The crucial point about this is that we have the following:

$$
\begin{aligned}
D_i'\,\psi' &\equiv \left(\partial_i - \frac{\mathrm{i}\,e}{\hbar}\,A_i'\right)\psi' = \left(\partial_i - \frac{\mathrm{i}\,e}{\hbar}\,A_i - \frac{\mathrm{i}\,e}{\hbar}\,(\partial_i\lambda)\right)\left(e^{\mathrm{i}\,e\lambda/\hbar}\,\psi\right), \\
&= e^{\mathrm{i}\,e\lambda/\hbar}\left(\partial_i - \frac{\mathrm{i}\,e}{\hbar}\,A_i - \frac{\mathrm{i}\,e}{\hbar}\,(\partial_i\lambda) + \frac{\mathrm{i}\,e}{\hbar}\,(\partial_i\lambda)\right)\psi, \\
&= e^{\mathrm{i}\,e\lambda/\hbar}\left(\partial_i - \frac{\mathrm{i}\,e}{\hbar}\,A_i\right)\psi,
\end{aligned}
\tag{9.10}
$$

and

$$
\begin{aligned}
D_0'\,\psi' &\equiv \left(\frac{\partial}{\partial t} + \frac{\mathrm{i}\,e}{\hbar}\,\phi'\right)\psi' = \left(\frac{\partial}{\partial t} + \frac{\mathrm{i}\,e}{\hbar}\,\phi - \frac{\mathrm{i}\,e}{\hbar}\,\frac{\partial\lambda}{\partial t}\right)\left(e^{\mathrm{i}\,e\lambda/\hbar}\,\psi\right), \\
&= e^{\mathrm{i}\,e\lambda/\hbar}\left(\frac{\partial}{\partial t} + \frac{\mathrm{i}\,e}{\hbar}\,\phi - \frac{\mathrm{i}\,e}{\hbar}\,\frac{\partial\lambda}{\partial t} + \frac{\mathrm{i}\,e}{\hbar}\,\frac{\partial\lambda}{\partial t}\right)\psi, \\
&= e^{\mathrm{i}\,e\lambda/\hbar}\left(\frac{\partial}{\partial t} + \frac{\mathrm{i}\,e}{\hbar}\,\phi\right)\psi.
\end{aligned}
\tag{9.11}
$$

In other words, we have

$$
D_i'\psi' = e^{\mathrm{i}\,e\lambda/\hbar}\,D_i\psi, \qquad D_0'\psi' = e^{\mathrm{i}\,e\lambda/\hbar}\,D_0\psi.
\tag{9.12}
$$

This means that $D_i\psi$ and $D_0\psi$ transform the same way as ψ itself under the gauge transformations (9.9), namely just by acquiring the phase factor $e^{\mathrm{i}\,e\lambda/\hbar}$. This is a non-trivial statement, since the gauge parameter λ is an arbitrary function of space and time. Had we been considering standard partial derivatives ∂_i and $\partial/\partial t$ rather than the covariant derivatives defined in (9.7), it would most certainly not have been true. For example,

$$
\partial_i\psi' = \partial_i\left(e^{\mathrm{i}\,e\lambda/\hbar}\,\psi\right) = e^{\mathrm{i}\,e\lambda/\hbar}\,\partial_i\psi + e^{\mathrm{i}\,e\lambda/\hbar}\,\frac{\mathrm{i}\,e}{\hbar}\,(\partial_i\lambda)\,\psi \neq e^{\mathrm{i}\,e\lambda/\hbar}\,\partial_i\psi,
\tag{9.13}
$$

precisely because the derivative can land on the space-time dependent gauge-transformation parameter λ and thus give the second term, which spoils the covariance of the transformation. The point about the covariant derivatives is that the contributions from the gauge transformation of the gauge potentials precisely cancels the "unwanted" second term in (9.13).

By iterating the calculation, it also follows that $D_i'D_i'\psi' = e^{\mathrm{i}\,e\lambda/\hbar}\,D_iD_i\psi$, and so we see that the Schrödinger equation (9.8) written in terms of the primed fields, i.e.

$$
-\frac{\hbar^2}{2m}\,D_i'D_i'\psi' - \mathrm{i}\,\hbar\,D_0'\psi' = 0,
\tag{9.14}
$$

just implies the Schrödinger equation in terms of unprimed fields, since

$$
\begin{aligned}
0 &= -\frac{\hbar^2}{2m}\,D_i'D_i'\psi' - \mathrm{i}\,\hbar\,D_0'\psi', \\
&= e^{\mathrm{i}\,e\lambda/\hbar}\left(-\frac{\hbar^2}{2m}\,D_iD_i\psi - \mathrm{i}\,\hbar\,D_0\psi\right).
\end{aligned}
\tag{9.15}
$$

What we have proved above is that the Schrödinger equation transforms covariantly under electromagnetic gauge transformations, provided that at the same time the wave function is scaled by a space-time dependent phase factor, as in (9.9). Note that we use the term "covariant transformation" here in the same sense as we used it earlier in the course when discussing the behaviour of the Maxwell equations under Lorentz transformations. The actual transformation is totally different in the two contexts; here we are discussing the behaviour of the Schrödinger equation under *gauge transformations* rather than Lorentz transformations, but in each case the essential point, which is characteristic of a covariance of any equation under a symmetry transformation, is that the equation expressed in terms of the symmetry-transformed (primed) variables is identical in form to the original equation for the unprimed variables, but with a prime placed on every field.

Note that the two definitions of the spatial and time covariant derivatives in (9.7) can be unified into the single 4-dimensional definition

$$D_\mu = \partial_\mu - \frac{\mathrm{i}\,e}{\hbar}\,A_\mu \qquad (9.16)$$

since we have $A^\mu = (\phi, \vec{A})$, and hence $A_\mu = (-\phi, \vec{A})$.

The Schrödinger equation itself provides only an approximate description of the quantum theory of matter. This is obvious from the fact that it is obtained by starting from a *non-relativistic* classical Hamiltonian, and then applying the quantisation procedure. And indeed, clearly the Schrödinger equation (9.8) does not transform covariantly under Lorentz transformations. (The fact that time is treated on a completely different footing from space makes this obvious.) The non-relativistic Schrödinger equation is therefore inconsistent with causality (essentially, the notion that nothing can travel faster than light). At the very least, one should therefore be taking a *relativistic* classical theory as the the starting point for applying the quantisation procedure. It turns out that this is not sufficient. If one constructs a relativistic generalisation of the Schrödinger equation, one then encounters difficulties in giving a probabilistic interpretation of the wave function, related to the fact that the probability density current is not positive definite. The resolution of this problem requires a further process of quantisation, known as *second quantisation*, in which the wave-function itself becomes an operator that creates and annihilates particles. Theories of this type are known as *quantum field theories*, and they lie at the heart of all the fundamental theories of matter and forces.

An example is *quantum electrodynamics*, which is a quantum field theory describing electromagnetism together with electrons (and necessarily positrons, which are the antiparticles of electrons). In this theory the 4-vector gauge potential A_μ becomes a quantum field,

which creates and annihilates photons, and the electron is described by a quantum field that satisfies a relativistic equation known as the Dirac equation. It is one of the most spectacularly successful theories known, leading to predictions that have been experimentally verified to quite remarkable accuracy.

9.2 Magnetic monopoles

The Maxwell equations

$$\partial_\mu F^{\mu\nu} = -4\pi J^\nu \,,$$
$$\partial_\mu F_{\nu\rho} + \partial_\nu F_{\rho\mu} + \partial_\rho F_{\mu\nu} = 0 \tag{9.17}$$

take on a more symmetrical-looking form if we introduce the dual of the field-strength tensor, defined by

$$\widetilde{F}_{\mu\nu} = \tfrac{1}{2}\epsilon_{\mu\nu\rho\sigma}\,F^{\rho\sigma}\,. \tag{9.18}$$

In terms of $\widetilde{F}_{\mu\nu}$, the second equation in (9.17) (i.e. the Bianchi identity) becomes

$$\partial_\mu \widetilde{F}^{\mu\nu} = 0\,. \tag{9.19}$$

From $F_{0i} = -E_i$ and $F_{ij} = \epsilon_{ijk}B_k$, it is easy to see that

$$\widetilde{F}_{0i} = B_i\,, \qquad \widetilde{F}_{ij} = \epsilon_{ijk}\,E_k\,. \tag{9.20}$$

It follows that $\widetilde{F}_{\mu\nu}$ is obtained from $F_{\mu\nu}$ by making the replacements

$$\vec{E} \longrightarrow -\vec{B}\,, \qquad \vec{B} \longrightarrow \vec{E}\,. \tag{9.21}$$

The symmetry between the two Maxwell equations would become even more striking if there were a current on the right-hand side of (9.19), analogous to the electric 4-current density on the right-hand-side of the first Maxwell equation in (9.17). Since the rôles of $\vec{E}$ and $\vec{B}$ are exchanged when passing from $F_{\mu\nu}$ to $\widetilde{F}_{\mu\nu}$, it is evident that the 4-current needed on the right-hand side of (9.19) must be a *magnetic* 4-current density, J_M^μ. Let us now attach a subscript E to the standard electric 4-current density, in order to emphasise which is which in the following. The generalised Maxwell equations will now be written as

$$\partial_\mu F^{\mu\nu} = -4\pi J_E^\nu\,, \qquad \partial_\mu \widetilde{F}^{\mu\nu} = -4\pi J_M^\nu\,. \tag{9.22}$$

Particles with magnetic charge, known as magnetic monopoles, have never been seen in nature. However, there seems to be no reason in principle why they should not exist, and

it is of interest to explore their properties in a little more detail. A point electric charge e has an electric field given by

$$\vec{E} = \frac{e\,\vec{r}}{r^3}\,. \tag{9.23}$$

Thus by analogy, a point magnetic monopole, with magnetic charge g, will have a magnetic field given by

$$\vec{B} = \frac{g\,\vec{r}}{r^3}\,. \tag{9.24}$$

This satisfies

$$\vec{\nabla} \cdot \vec{B} = 4\pi\,\rho_M\,, \qquad \rho_M = g\,\delta^3(\vec{r})\,, \tag{9.25}$$

where $\rho_M = J_M^0$ is the magnetic charge density.

We shall be interested in studying the quantum mechanics of electrically-charged particles in the background of a magnetic monopole. Since the Schrödinger equation is written in terms of the potentials ϕ and $\vec{A}$, we shall therefore need to write down the 3-vector potential $\vec{A}$ for the magnetic monopole. To do this, we introduce Cartesian coordinates (x, y, z), related to spherical polar coordinates (r, θ, φ) in the standard way,

$$x = r\sin\theta\,\cos\varphi\,, \quad y = r\sin\theta\,\sin\varphi\,, \quad x = r\cos\theta\,, \tag{9.26}$$

and we also define

$$\rho^2 = x^2 + y^2\,. \tag{9.27}$$

Consider the 3-vector potential

$$\vec{A} = g\left(\frac{zy}{r\rho^2}, -\frac{zx}{r\rho^2}, 0\right). \tag{9.28}$$

Using

$$\begin{aligned}
\frac{\partial r}{\partial x} &= \frac{x}{r}\,, & \frac{\partial r}{\partial y} &= \frac{y}{r}\,, & \frac{\partial r}{\partial z} &= \frac{z}{r}\,, \\
\frac{\partial \rho}{\partial x} &= \frac{x}{\rho}\,, & \frac{\partial \rho}{\partial y} &= \frac{y}{\rho}\,, & \frac{\partial \rho}{\partial z} &= 0\,,
\end{aligned} \tag{9.29}$$

it is easily seen that

$$B_x = \partial_y A_z - \partial_z A_y = g\partial_z\left(\frac{zx}{r\rho^2}\right) = \frac{gx}{r\rho^2} - \frac{gxz^2}{r^3\rho^2} = \frac{gx}{r^3}\,, \tag{9.30}$$

and similarly

$$B_y = \frac{gy}{r^3}\,, \qquad B_z = \frac{gz}{r^3}\,. \tag{9.31}$$

Thus indeed we find that

$$\vec{\nabla} \times \vec{A} = \frac{g\,\vec{r}}{r^3}\,, \tag{9.32}$$

and so the 3-vector potential (9.28) describes the magnetic monopole field (9.24).

In terms of spherical polar coordinates we have $\rho^2 = x^2 + y^2 = r^2 \sin^2 \theta$, and so (9.28) can be written as

$$\vec{A} = \frac{g \cot \theta}{r} \left(\sin \varphi, -\cos \varphi, 0 \right). \tag{9.33}$$

Not surprisingly, this potential is singular at $r = 0$, since we are describing an idealised point magnetic charge. In exactly the same way, the potential $\phi = e/r$ describing a point electric charge diverges at $r = 0$ also. However, the potential (9.33) also diverges everywhere along the z axis, i.e. at $\theta = 0$ and $\theta = \pi$. It turns out that these latter singularities are "unphysical," in the sense that they can be removed by making gauge transformations. This is not too surprising, when we note that the magnetic field itself, given by (9.24) has no singularity along the z axis. It is, of course, genuinely divergent at $r = 0$, so that is a real physical singularity.

To see the unphysical nature of the singularities in (9.33) along $\theta = 0$ and $\theta = \pi$, we need to make gauge transformations, under which

$$\vec{A} \longrightarrow \vec{A} + \vec{\nabla}\lambda. \tag{9.34}$$

Consider first taking

$$\lambda = g\,\varphi = g \arctan \frac{y}{x}. \tag{9.35}$$

From this, we find

$$\vec{\nabla}\lambda = -\frac{g}{r} \operatorname{cosec}\theta \left(\sin \varphi, -\cos \varphi, 0 \right). \tag{9.36}$$

Letting the gauge-transformed potential be $\vec{A}'$, we therefore find

$$\vec{A}' = \vec{A} + \vec{\nabla}\lambda = \frac{g}{r} \frac{\cos \theta - 1}{\sin \theta} \left(\sin \varphi, -\cos \varphi, 0 \right) = -\frac{g}{r} \tan \tfrac{1}{2}\theta \left(\sin \varphi, -\cos \varphi, 0 \right). \tag{9.37}$$

It can be seen that $\vec{A}$ is completely non-singular along $\theta = 0$ (i.e. along the positive z axis). It is, however, singular along $\theta = \pi$ (i.e. along the negative z axis).

We could, on the other hand, perform a gauge transformation with λ given by

$$\lambda = -g\,\varphi = -g \arctan \frac{y}{x} \tag{9.38}$$

instead of (9.35). Defining the gauge-transformed potential as $\vec{A}''$ in this case, we find

$$\vec{A}'' = \frac{g}{r} \cot \tfrac{1}{2}\theta \left(\sin \varphi, -\cos \varphi, 0 \right). \tag{9.39}$$

This time, we have obtained a gauge potential that is non-singular along $\theta = \pi$ (i.e. the negative z axis), but it is singular along $\theta = 0$ (the positive z axis).

There is no single choice of gauge in which the 3-vector potential for the magnetic monopole is completely free of singularities away from the origin $r = 0$. We have obtained two expressions for the vector potential, one of which, $\vec{A}'$, is non-singular along the positive z axis, and the other, $\vec{A}''$, is non-singular along the negative z axis. The singularity that each has is known as a "string singularity," since it lies along a line, or string. By making gauge transformations the location of the string can be moved around, but it can never be removed altogether.

In the discussion above, the z axis appears to have played a preferred rôle, but this is, of course, just an artefact of our gauge choices. We could equally well have chosen a different expression for $\vec{A}$, related by a gauge transformation, for which the string singularity ran along any desired line, or curve, emanating from the origin.

9.3 Dirac quantisation condition

We have seen that gauge potentials for the magnetic monopole, free of singularities on the positive and negative z axes resepctively, are given by

$$
\begin{aligned}
\vec{A}' &= -\frac{g}{r}\tan\tfrac{1}{2}\theta\,(\sin\varphi, -\cos\varphi, 0)\,,\\
\vec{A}'' &= \frac{g}{r}\cot\tfrac{1}{2}\theta\,(\sin\varphi, -\cos\varphi, 0)\,.
\end{aligned}
\tag{9.40}
$$

The two are themselves related by a gauge transformation, namely

$$
\vec{A}'' = \vec{A}' + \vec{\nabla}(-2g\varphi)\,.
\tag{9.41}
$$

Now let us consider the quantum mechanics of an electron in the background of the magnetic monopole. As we discussed in section 9.1, the Schrödinger equation for the electron is given by (9.4), where e is its charge, and m is its mass. We shall consider the Schrödinger equation in two different gauges, related as in (9.41). Denoting the corresponding electron wave-functions by ψ' and ψ'', we see from (9.9) (9.41) that we shall have

$$
\psi'' = e^{-2\mathrm{i}\,eg\varphi/\hbar}\,\psi'\,.
\tag{9.42}
$$

However, we have seen that the gauge transformation is not physical, but merely corresponds to shifting the string singularity of the magnetic monopole from the negative z axis to the positive z axis. Quantum mechanically, the physics will only be unchanged if the electron wave-function remains single valued under a complete 2π rotation around the z axis. This means that the phase factor in the relation (9.42) must be equal to unity, and so it must be that

$$
\frac{2eg}{\hbar}\,2\pi = 2\pi\,n\,,
\tag{9.43}
$$

186

where n is an integer. Thus it must be that the product of the electric charge e on the electron, and the magnetic charge g on the magnetic monopole, must satisfy the so-called *Dirac quantisation condition*,

$$2e\,g = n\,\hbar\,. \tag{9.44}$$

It is interesting to note that although a magnetic monopole has never been observed, it would only take the existence of a single monopole, maybe somewhere in another galaxy, to imply that electric charges everywhere in the universe must quantised in units of

$$\frac{\hbar}{2g}\,, \tag{9.45}$$

where g is the magnetic charge of the lonely magnetic monopole. In fact all observed electric charges are indeed quantised; in integer multiples of the charge e on the electron, in everyday life, and in units of $\frac{1}{3}e$ in the quarks of the theory of strong interactions. It is tempting to speculate that the reason for this may be the existence of a magnetic monopole somewhere out in the vastness of space, in a galaxy far far away.

10 Local Gauge Invariance and Yang-Mills Theory

10.1 Relativistic quantum mechanics

We saw in the previous section that the ordinary non-relativistic quantum mechanics of a charged particle in an electromagnetic field has the feature that it is *covariant* under electromagnetic gauge transformations, provided that the usual gauge trnasformation of the 4-vector potential is combined with a phase transformation of the wavefunction for the charged particle:

$$A_\mu \longrightarrow A_\mu + \partial_\mu \lambda\,, \qquad \psi \longrightarrow e^{\mathrm{i}\,e\lambda/\hbar}\,\psi\,. \tag{10.1}$$

The essential point here is that the gauge transformation parameter λ can be an arbitrary function of the spacetime coordinates, and so the phase transformation of the wavefunction is a spacetime-dependent one. Such spacetime-dependent transformations are known as *local* transformations.

One could turn this around, and view the introduction of the electromagnetic field as the necessary addition to quantum mechanics in order to allow the theory to be covariant under local phase transformations of the wavefunction. If we started with quantum mechanics in the absence of electromagnetism, so that for a free particle of mass m we have

$$-\frac{\hbar^2}{2m}\,\vec{\nabla}^2\,\psi = \mathrm{i}\,\hbar\,\frac{\partial\psi}{\partial t}\,, \tag{10.2}$$

then the Schrödinger equation is obviously covariant under *constant* phase transformations of the wavefunction,

$$\psi \longrightarrow e^{\mathrm{i}c}\,\psi\,, \tag{10.3}$$

where c is an arbitrary constant. And, indeed, the physics described by the wavefunction is invariant under this phase transformation, since all physical quantities are constructed using a product of ψ and its complex conjugate $\bar{\psi}$ (for example, the probability density $|\psi|^2$), for which the phase factors cancel out. Also, clearly, the Schrödinger equation (10.2) does *not* transform nicely under local phase transformations, since the derivatives will now land on the (spacetime dependent) phase factor in (10.3) and give a lot of messy terms.

As we now know, the way to achieve a nice covariant transformation of the Schrödinger equation under local phase transformations is to replace the partial derivatives ∂_i and ∂_0 in (10.2) by covariant derivatives

$$D_i = \partial_i - \frac{\mathrm{i}\,e}{\hbar}\,A_i\,, \qquad D_0 = \partial_0 + \frac{\mathrm{i}\,e}{\hbar}\,\phi \tag{10.4}$$

where A_i and ϕ transform in the standard way for the electromagnetic gauge potentials at the same time as the local phase transformation for ψ is performed, as in (10.1). From this point of view, it could be said that we have *derived* electromagnetism as the field needed in order to allow the Schrödinger equation to transform covariantly under local phase transformations of the wavefunction.

The idea now is to extend this idea to more general situations. By again demanding local "phase" transformations of some quantum-mechanical equation, we will now be able to derive a generalisation of electromagnetism known as Yang-Mills theory.

Working with a non-relativistic equation like the Schrödinger equation is rather clumsy, because of the way in which space and time arise on such different footings. It is more elegant (and simpler) to switch at this point to the consideration of relativistic quantum mechanical equation. There are various possible equations one could consider, but they all lead to equivalent conclusions about the generalisation of electromagnetism. Examples one could consider include the Dirac equation, which provides a relativistic description of the electron, or any other fermionic particle with spin $\frac{1}{2}$. A simpler opetion is to consider the Klein-Gordon equation for a relativistic particle of spin 0 (otherwise known as a scalar field). The Klein-Gordon equation for a free scalar field φ with mass m is very simple, namely

$$\Box\,\varphi - m^2\varphi = 0\,, \tag{10.5}$$

where $\Box = \partial^\mu \partial_\mu$ is the usual d'Alembertian operator, which, as we know, is Lorentz invariant.[32] Note that from now on, we shall use units where Planck's constant $\hbar$ is set equal to 1.

In what follows, we shall make the simplifying assumption that the scalar field is massless, and so its Klein-Gordon equation is simly

$$\Box \varphi = 0 \,. \qquad (10.6)$$

We shall do this because no essential feature that we wish to explore will be lost, and it will slightly shorten the equations. It is completely straightforward to add it back in if desired.[33]

The Klein-Gordon equation (10.6) can be derived from the Lagrangian density

$$\mathcal{L} = -\tfrac{1}{2} \partial^\mu \varphi \, \partial_\mu \varphi \,. \qquad (10.7)$$

Varying the action $I = \int d^4x \mathcal{L}$, we find

$$\delta I = -\int d^4x \partial^\mu \varphi \, \partial_\mu \delta\varphi = \int d^4x \, (\partial_\mu \partial^\mu \varphi) \, \delta\varphi \,, \qquad (10.8)$$

(dropping the boundary term at infinity as usual), and so demanding that the action be stationary under arbitrary variations $\delta\varphi$ implies the Klein-Gordon equation (10.6).

Before moving on to the generalisation to Yang-Mills theory, we shall first review, again, the derivation of electromagnetism as the field needed in order to turn a global phase invariance into a local invariance, this time from the viewpoint of the relativistic Klein-Gordon equation. To do this, we first need to enlarge the system of wavefunctions from one real scalar to two. Suppose, then, we have two real scalars called φ_1 and φ_2, each satisfying

---

[32]The non-relativistic Schrödinger equation can be derived from the Klein-Gordon equation (10.5) in an appropriate limit. The leading-order time dependence of a field with mass (i.e. energy) m will be $e^{-\mathrm{i}\,mt}$ (in units where we set $\hbar = 1$). Thus the appropriate non-relativistic approximation is where where the wavefunction φ is assumed to be of the form $\varphi \sim e^{-\mathrm{i}\,mt}\,\psi$, with ψ only slowly varying in time, which means that the term $\partial^2 \psi / \partial t^2$ can be neglected in comparison to the others. Substituting into (10.5) and dropping the $\partial^2 \psi / \partial t^2$ term gives precisely the Schrödinger equation for the free massive particle, namely $-(1/2m)\nabla^2 \psi = \mathrm{i}\,\partial \psi / \partial t$.

[33]Note that we can only discuss a non-relativistic limit for the *massive* Klein-Gordon equation. This is because the non-relativisitic approximation (discussed in the previous footnote) involved assuming that each time derivative of ψ with respect to t was small compared with m times ψ. Clearly this would no longer be true if m were zero. Put another way, a massless particle is inherently relativistic, since it must travel at the speed of light (like the photon). We shall not be concerned with taking the non-relativistic limit in what follows, and so working with a massless field will not be a problem.

a Klein-Gordon equation. These equations can therefore be derived from the Lagrangian density

$$\mathcal{L} = -\tfrac{1}{2}\partial^\mu\varphi_1\,\partial_\mu\varphi_1 - \tfrac{1}{2}\partial^\mu\varphi_2\,\partial_\mu\varphi_2\,. \tag{10.9}$$

We can conveniently combine the two real fields into a complex scalar field ϕ, defined by

$$\phi = \frac{1}{\sqrt{2}}\,(\varphi_1 + \mathrm{i}\,\varphi_2)\,. \tag{10.10}$$

The Lagrangian density can then be written as

$$\mathcal{L} = -\tfrac{1}{2}\partial^\mu\bar{\phi}\,\partial_\mu\phi\,. \tag{10.11}$$

The complex field ϕ therefore satisfies the Klein-Gordon equation

$$\Box\,\phi = 0\,. \tag{10.12}$$

It is clear that the complex field ϕ has a global phase invariance, under

$$\phi \longrightarrow e^{\mathrm{i}\,\alpha}\,\phi\,, \tag{10.13}$$

where α is a constant. (The term "global" is used to describe such phase transformations, which are identical at every point in spacetime.) This can be seen at the level of the Klein-Gordon equation (10.12), since the constant phase factor simply passes straight through the d'Alembertian operator. It can also be seen at the level of the Lagrangian density, since again the derivatives do not land on the phase factor, and furthermore, the $e^{\mathrm{i}\,\alpha}$ phase factor from transforming ϕ is cancelled by the $e^{-\mathrm{i}\,\alpha}$ phase factor from tramsforming $\bar{\phi}$.

It is also clear that the Lagrangian density is *not* invariant under *local* phase transformations, where α is assumed now to be spacetime dependent. This is because we now have

$$\partial_\mu\phi \longrightarrow \partial_\mu(e^{\mathrm{i}\,\alpha}\,\phi) \longrightarrow e^{\mathrm{i}\,\alpha}\,\partial_\mu\phi + \mathrm{i}\,(\partial_\mu\alpha)\,\phi\,. \tag{10.14}$$

It is the second term, where the derivatives land on α, that spoils the invariance.

The remedy, not surprisingly, is to introduce a gauge potential A_μ, and replace the partial derivatives by covariant derivatives

$$D_\mu = \partial_\mu - \mathrm{i}\,e\,A_\mu\,, \tag{10.15}$$

where now ϕ will be interpreted as describing a complex scalar field with electric charge e. As we saw before when discussing the Schrödinger equation, the covariant derivative

acting on ϕ has a nice transformation property under the local phase tramsformations of ϕ, provided at the same time we transform A_μ:

$$\phi \longrightarrow e^{i\alpha}\phi, \qquad A_\mu \longrightarrow A_\mu + \frac{1}{e}\partial_\mu\alpha. \qquad (10.16)$$

This implies that $D_\mu\phi$ transforms nicely as

$$D_\mu\phi \longrightarrow e^{i\alpha}D_\mu\phi, \qquad (10.17)$$

and so the new Lagrangian density

$$\mathcal{L} = -\tfrac{1}{2}(\overline{D^\mu\phi})(D_\mu\phi) \qquad (10.18)$$

is indeed invariant. This is the "derivation" of ordinary electromagnetism.

In this viewpoint, where we are deriving electromagnetism by requiring the local phase invariance of the theory under (10.13), has not yet given any dynamics to the gauge field A_μ. Indeed, one cannot *derive* a dynamical existence for A_μ because in fact there is no unique answer. What one can do, however, is to introduce a dynamical term "by hand," which has all the natural properties one would like. First of all, we want a dynamical term that respects the gauge invariance we have already achieved in the rest of the Lagrangian. Secondly, we expect that it should give rise to a second-order dynamical equation for A_μ. We are back to the discussion of section 4.2, where we derived Maxwell's equations from an action principal. The steps leading to the answer are as follows.

First, to make a gauge-invariant term we need to use the gauge-invariant field strength

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu \qquad (10.19)$$

as the basic "building block." Then, to make a Lorentz-invariant term, the lowest-order possibility is to form the quadratic invariant $F^{\mu\nu}F_{\mu\nu}$. Taking the standard normalisation as discussed in section 4.2, we are therefore lead to propose the total Lagrangian density

$$\mathcal{L} = -\tfrac{1}{2}(\overline{D^\mu\phi})(D_\mu\phi) - \frac{1}{16\pi}F^{\mu\nu}F_{\mu\nu}, \qquad (10.20)$$

where $D_\mu = \partial_\mu - i\,e\,A_\mu$. It is easily verified that the Euler-Lagrange equations resulting from this Lagrangian density are as follows. Requiring the stationarity of the action under variations of the wavefunction ϕ implies

$$D^\mu D_\mu\phi = 0, \qquad \text{i.e.} \qquad (\partial^\mu - i\,eA^\mu)(\partial_\mu - i\,eA_\mu)\phi = 0, \qquad (10.21)$$

This is the gauge-covariant generalisation of the original uncharged Klein-Gordon equation.

191

Requiring stationarity under variations of A_μ implies

$$\partial_\mu F^{\mu\nu} = -4\pi J^\nu \,, \qquad (10.22)$$

where

$$J_\mu = -\mathrm{i}\,e\left(\bar\phi D_\mu \phi - (\overline{D_\mu\phi})\,\phi\right). \qquad (10.23)$$

Thus A_μ satisfies the Maxwell field equation, with a source current density given by (10.23). This is exactly what one would hope for; the complex field ϕ carries electric charge e, and so it is to be expected that it should act as a source for the electromagnetic field. In the process of giving dynamics to the electromagnetic field we have, as a bonus, derived the current density for the scalar field.

10.2 Yang-Mills theory

At the end of the previous subsection we rederived electromagnetism as the field needed in order to turn the global phase invariance of a complex scalar field that satisfies the Kelin-Gordon equation into a local phase invariance. The phase factor $e^{\mathrm{i}\alpha}$ is a unit-modulus complex number. The set of all unit-modulus complex numbers form the group $U(1)$; i.e. 1×1 complex matrices U satisfying $U^\dagger U = 1$. (For 1×1 matrices, which are just numbers, there is of course no distinction between Hermitean conjugation, denoted by a dagger, and complex conjugation.)

In order to derive the generalisation of electromagnetism to Yang-Mills theory we need to start with an extended system of scalar fields, each satisfying the Klein-Gordon equation, whose Lagrangian is invariant under a larger, non-abelian, group.[34] We shall take the example of the group $SU(2)$ in order to illustrate the basic ideas. One can in fact construct a Yang-Mills theory based on any Lie group.

The group $SU(2)$ should be familiar from quantum mechanics, where it arises when one discusses systems with intrinstic spin $\frac{1}{2}$. The group can be defined as the set of 2×2 complex matrices U subject to the conditions

$$U^\dagger U = 1\,, \qquad \det U = 1\,. \qquad (10.24)$$

---

[34]An abelian group is one where the order of combination of group elements makes no difference. By contrast, for a non-ableian group, if two elements U and V are combined in the two different orderings, the results are, in general, different. Thus, for a group realised by matrices under multiplication, for example, one has in general that $UV \neq VU$.

It can therefore be parameterised in the form

$$U = \begin{pmatrix} a & b \\ -\bar{b} & \bar{a} \end{pmatrix} , \tag{10.25}$$

where a and b are complex numbers subject to the constraint

$$|a|^2 + |b|^2 = 1 . \tag{10.26}$$

If we write $a = x_1 + \mathrm{i}\, x_2$, $b = x_3 + \mathrm{i}\, x_4$, the constraint is described by the surface

$$x_1^2 + x_2^2 + x_3^2 + x_4^2 = 1 \tag{10.27}$$

in Euclidean 4-space, and so the elements of the group $SU(2)$ are in one-to-one correspondence with the points on a unit 3-dimensional sphere. Clearly $SU(2)$ has three independent parameters.[35]

The group $SU(2)$ can be generated by exponentiating the three Pauli matrices τ_a, where

$$\tau_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} , \qquad \tau_2 = \begin{pmatrix} 0 & -\mathrm{i} \\ \mathrm{i} & 0 \end{pmatrix} , \qquad \tau_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} . \tag{10.28}$$

They satisfy the commutation relations

$$[\tau_a, \tau_b] = 2\mathrm{i}\, \epsilon_{abc}\, \tau_c , \tag{10.29}$$

i.e. $[\tau_1, \tau = 2] = 2\mathrm{i}\, \tau_3$, and cyclic permutations.

Let

$$T_a = \frac{1}{2\mathrm{i}}\, \tau_a . \tag{10.30}$$

We shall therefore have

$$[T_a, T_b] = \epsilon_{abc}\, T_c . \tag{10.31}$$

Note that the T_a, which are called the generators of the Lie algebra of $SU(2)$, are anti-Hermitean,

$$T_a^\dagger = -T_a . \tag{10.32}$$

They are also, of course, traceless.

---

[35]For comparison, the group $U(1)$, whose elements U can be parameterised as $U = e^{\mathrm{i}\,\alpha}$ with $0 \le \alpha < 2\pi$, has one parameter. Since $e^{\mathrm{i}\,\alpha}$ is periodic in α the elements of $U(1)$ are in one-to-one correspondence with the points on a unit circle, or 1-dimensional sphere. In fact the circle, S^1, and the 3-sphere, S^3, are the only spheres that are isomorphic to groups.

The $SU(2)$ group elements can be written as

$$U = e^{\alpha_a T_a} \,, \tag{10.33}$$

where α_a are three real parameters. (This is the analogue of writing the $U(1)$ elements U as $U = e^{\mathrm{i}\alpha}$.) It is easy to check that the unitarity of U, i.e. $U^\dagger U = 1$, follows from the anti-Hermiticity of the generators T_a:

$$U^\dagger U = \left(e^{\alpha_a T_a}\right)^\dagger e^{\alpha_b T_b} = e^{\alpha_a T_a^\dagger} e^{\alpha_b T_b} = e^{-\alpha_a T_a} e^{\alpha_b T_b} = 1 \,. \tag{10.34}$$

The unit-determinant property follows from the tracelessness of the T_a, bearing in mind that for any matrix X we have $\det X = \exp(\mathrm{tr}\log X)$:

$$\det U = \det(e^{\alpha_a T_a}) = \exp[\mathrm{tr}\log(e^{\alpha_a T_a})] = \exp[\mathrm{tr}(\alpha_a T_a)] = \exp[0] = 1 \,. \tag{10.35}$$

Suppose now that we take a pair of *complex* scalar fields, called ϕ_1 and ϕ_2, each of which satisfies the massless Klein-Gordon equation. We may assemble them into a complex 2-vector, which we shall call ϕ:

$$\phi = \begin{pmatrix} \phi_1 \\ \phi_2 \end{pmatrix} \,. \tag{10.36}$$

This vector-valued field therefore satisfies the Klein-Gordon equation

$$\Box\phi = 0 \,, \tag{10.37}$$

which can be derived from the Lagrangian density

$$\mathcal{L} = -(\partial^\mu \phi^\dagger)(\partial_\mu \phi) \,. \tag{10.38}$$

It is obvious that the Lagrangian density (10.38) is invariant under global $SU(2)$ transformations

$$\phi \longrightarrow U\phi \,, \tag{10.39}$$

where U is a constant $SU(2)$ matrix. Thus, we have

$$\mathcal{L} \longrightarrow -\partial^\mu(\phi^\dagger U^\dagger)\,\partial_\mu(U\phi) = -(\partial^\mu\phi^\dagger)\,U^\dagger U\,(\partial_\mu\phi) = -(\partial^\mu\phi^\dagger)\,(\partial_\mu\phi) = \mathcal{L} \,. \tag{10.40}$$

Obviously, $\mathcal{L}$ would not be invariant if we allowed U to be space-time dependent, for the usual reason that we would get extra terms where the derivatives landed on the U transformation matrix. Based on our experience with the local $U(1)$ phase invariance of the theory coupled to electromagnetism, we can expect that again we could achieve a local

$SU(2)$ invariance by introducing appropriate analogues of the electromagnetic field. In this case, since the $SU(2)$ group is characterised by 3 parameters rather than the 1 parameter characterising $U(1)$, we can expect that we will need 3 gauge fields rather than 1. We shall called these A_μ^a, where $1 \leq a \leq 3$. In fact it is convenient to assemble the three gauge fields into a 2×2 matrix, by defining

$$A_\mu = A_\mu^a T_a\,, \tag{10.41}$$

where T_a are the generators of the $SU(2)$ algebra that we introduced earlier.

We next define the covariant derivative D_μ, whose action on the complex 2-vector of scalar fields ϕ is defined by

$$D_\mu \phi = \partial_\mu \phi + A_\mu\,\phi\,. \tag{10.42}$$

Since we don't, *a priori*, know how A_μ should transform we shall work backwards and demand that its transformation rule should be such that D_μ satisfies the nice property we should expect of a covariant derivative in this case, namely that if we transform ϕ under a *local $SU(2)$* transformation

$$\phi \longrightarrow \phi' = U\,\phi\,, \tag{10.43}$$

then we should also have that

$$(D_\mu \phi) \longrightarrow (D_\mu \phi)' = U(D_\mu \phi)\,. \tag{10.44}$$

Working this out, we shall have

$$\begin{aligned}
(D_\mu \phi)' &= D_\mu'\,\phi' = (\partial_\mu + A_\mu')(U\,\phi)\,, \\
&= (\partial_\mu U)\,\phi + U\,\partial_\mu \phi + A_\mu'\,U\,\phi\,, \\
&= U\,D_\mu \phi = U\,\partial_\mu \phi + U\,A_\mu\,\phi\,.
\end{aligned} \tag{10.45}$$

Equating the last two lines, and noting that we want this to be true for all possible ϕ, we conclude that

$$\partial_\mu U + A_\mu'\,U = U\,A_\mu\,. \tag{10.46}$$

Multiplying on the right with $U^\dagger$ then gives the result that

$$A_\mu' = U\,A_\mu\,U^\dagger - (\partial_\mu U)\,U^\dagger\,. \tag{10.47}$$

This, then, will be the gauge transformation rule for the Yang-Mills potentials A_μ. In other

words, the full set of local $SU(2)$ transformations comprise[36]

$$A_\mu \longrightarrow A'_\mu = U A_\mu U^\dagger - (\partial_\mu U) U^\dagger ,$$
$$\phi \longrightarrow \phi' = U \phi . \tag{10.48}$$

What we have established is that if we replace the Lagrangian density (10.38) by

$$\mathcal{L} = -(D^\mu \phi)^\dagger (D_\mu \phi) , \tag{10.49}$$

then it will be invariant under the *local $SU(2)$* transformations given by (10.48). The proof is now identical to the previous proof of the invariance of (10.38) under global $SU(2)$ transformations. The essential point is that the local transformation matrix U "passes through" the covariant derivative, in the sense that $(D_\mu \phi)' = D'_\mu(U\phi) = U D_\mu \phi$.

So far, we have suceeded in constructing a theory with a local $SU(2)$ symmetry, but as yet, the Yang-Mills potentials A^a_μ that we introduced do not have any dynamics of their own. Following the strategy we applied to the case of electromagnetism and local $U(1)$ invariance, we should now look for a suitable term to add to the Lagrangian density (10.49) that will do the job. Guided by the example of electromagnetism, we should first find a a field strength tensor for the Yang-Mills fields, which will be the analogue of the electromagnetic field strength

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu . \tag{10.50}$$

It is clear that the expression (10.50) is not suitable in the Yang-Mills case. If one were to try adopting (10.50) as a definition for the field strength, then a simple calculation shows that under the $SU(2)$ gauge transformation for A_μ given in (10.48), the field strength would transform into a complete mess. It turns out that the appropriate generalistion that is needed for Yang-Mills is to define

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu + [A_\mu, A_\nu] . \tag{10.51}$$

Of course this would reduce to the standard electromagnetic field strength in the abelian $U(1)$ case, since the commutator $[A_\mu, A_\nu] \equiv A_\mu A_\nu - A_\nu A_\mu$ would then vanish.

---

[36]Note that this non-abelian result, which takes essentially the same form for any group, reduces to the previous case of electromagnetic theory is we specialise to the abelian group $U(1)$. Essentially, we would just write $U = e^{i\alpha}$ and plug into the transformations (10.48). Since left and right multiplication are the same in the abelian case, the previous results for electromagnetic gauge invariance can be recovered.

The first task is to check how $F_{\mu\nu}$ defined by (10.51) transforms under the $SU(2)$ gauge transformation of A_μ given in (10.48). We shall have

$$
\begin{aligned}
F_{\mu\nu} \quad \longrightarrow \quad F'_{\mu\nu} &= \partial_\mu A'_\nu + A'_\mu A'_\nu - (\mu \leftrightarrow \nu)\,, \\
&= \partial_\mu (U A_\nu U^\dagger - (\partial_\nu U)U^\dagger) + (U A_\mu U^\dagger - (\partial_\mu U)U^\dagger)(U A_\nu U^\dagger - (\partial_\nu U)U^\dagger) - (\mu \leftrightarrow \nu)\,, \\
&= (\partial_\mu U) A_\nu U^\dagger + U(\partial_\mu A_\nu) - U A_\nu U^\dagger (\partial_\mu U)U^\dagger - (\partial_\mu \partial_\nu U)U^\dagger + (\partial_\nu U)U^\dagger (\partial_\mu U)U^\dagger \\
&\quad + U A_\mu U^\dagger U A_\nu U^\dagger - U A_\mu U^\dagger (\partial_\nu U)U^\dagger - (\partial_\mu U)U^\dagger U A_\nu U^\dagger \\
&\quad + (\partial_\mu U)U^\dagger (\partial_\nu U)U^\dagger - (\mu \leftrightarrow \nu)\,, \\
&= U(\partial_\mu A_\nu - \partial_\nu A_\mu + A_\mu A_\nu - A_\nu A_\mu)U^\dagger\,, \tag{10.52}
\end{aligned}
$$

where the notation $-(\mu \leftrightarrow \nu)$ means that one subtracts off from the terms written explicitly the same set of terms with the indices μ and ν exchanged. Comparing with (10.51) we see that the upshot is that under the $SU(2)$ gauge transformation for A_μ given in (10.48), the field strength $F_{\mu\nu}$ defined in (10.51) transforms as

$$
F_{\mu\nu} \longrightarrow F'_{\mu\nu} = U F_{\mu\nu} U^\dagger\,. \tag{10.53}
$$

This means that $F_{\mu\nu}$ transforms *covariantly* under $SU(2)$ gauge transformations. It would of course, reduce to the *invariance* of the elctromagnetic field strength transformation ($F'_{\mu\nu} = F_{\mu\nu}$) in the abelian case.

It is now a straightforward matter to write down a suitable term to add to the Lagrangian density (10.49). As for electromagnetism, we want a gauge-invariant and Lorentz-invariant quantity that is quadratic in fields. Thus we shall take

$$
\mathcal{L} = -(D^\mu \phi)^\dagger (D_\mu \phi) + \frac{1}{8\pi} \mathrm{tr}(F^{\mu\nu} F_{\mu\nu})\,, \tag{10.54}
$$

The proof that $\mathrm{tr}(F^{\mu\nu} F_{\mu\nu})$ is gauge invariant is very simple; under the $SU(2)$ gauge transformation we shall have

$$
\begin{aligned}
\mathrm{tr}(F^{\mu\nu} F_{\mu\nu}) \quad \longrightarrow \quad \mathrm{tr}(F^{\mu\nu\prime} F'_{\mu\nu}) &= \mathrm{tr}(U F^{\mu\nu} U^\dagger U F_{\mu\nu} U^\dagger) = \mathrm{tr}(U F^{\mu\nu} F_{\mu\nu} U^\dagger) \\
&= \mathrm{tr}(F^{\mu\nu} F_{\mu\nu} U^\dagger U) = \mathrm{tr}(F^{\mu\nu} F_{\mu\nu})\,. \tag{10.55}
\end{aligned}
$$

The equations of motion for the ϕ and A_μ fields can be derived from (10.54) in the standard way, as the Euler-Lagrange equations that follow from requiring that the action $I = \int d^4 x \mathcal{L}$ be stationary under variations of ϕ and A_μ respectively. First, let us just consider the *source-free* Yang-Mills equations that will result if we just consider the Lagrangian density for the Yang-Mills fields alone,

$$
\mathcal{L}_{YM} = \frac{1}{8\pi} \mathrm{tr}(F^{\mu\nu} F_{\mu\nu})\,. \tag{10.56}
$$

197

Wrting $I_{YM} = \int d^4x \mathcal{L}_{YM}$, we shall have

$$
\begin{aligned}
\delta I_{YM} &= \frac{1}{8\pi} 2 \operatorname{tr} \int d^4x \delta F_{\mu\nu} \, F^{\mu\nu} \,, \\
&= \frac{1}{4\pi} \operatorname{tr} \int d^4x (\partial_\mu \delta A_\nu - \partial_\nu \delta A_\mu + [\delta A_\mu, A_\nu] + [A_\mu, \delta A_\nu]) F^{\mu\nu} \,, \\
&= \frac{1}{2\pi} \operatorname{tr} \int d^4x (\partial_\mu \delta A_\nu + [A_\mu, \delta A_\nu]) F^{\mu\nu} \,, \\
&= \frac{1}{2\pi} \operatorname{tr} \int d^4x (-\delta A_\nu \, \partial_\mu F^{\mu\nu} + A_\mu \delta A_\nu F^{\mu\nu} - \delta A_\nu A_\mu F^{\mu\nu}) \,, \\
&= -\frac{1}{2\pi} \operatorname{tr} \int d^4x \delta A_\nu \, (\partial_\mu F^{\mu\nu} + [A_\mu, F^{\mu\nu}]) \,,
\end{aligned}
\tag{10.57}
$$

and so requiring that the action be stationary implies

$$
\partial_\mu F^{\mu\nu} + [A_\mu, F^{\mu\nu}] = 0 \,.
\tag{10.58}
$$

These are the source-free Yang-Mills equations, which are the generalisation of the source-free Maxwell equations $\partial_\mu F^{\mu\nu} = 0$. Obviously the Yang-Mills equations reduce to the Maxwell equations in the abelian $U(1)$ case.

If we now include the $-(D^\mu \phi)^\dagger (D_\mu \phi)$ term in the above calculation, we shall find that

$$
\delta I = \int d^4x (\phi^\dagger \delta A_\nu D^\nu \phi - (D^\nu \phi)^\dagger \delta A_\nu \phi) - \frac{1}{2\pi} \operatorname{tr} \int d^4x \delta A_\nu \, (\partial_\mu F^{\mu\nu} + [A_\mu, F^{\mu\nu}]) \,.
\tag{10.59}
$$

Requiring stationarity under the variations δA_ν now gives the Yang-Mills equations with sources,

$$
\partial_\mu F^{\mu\nu} + [A_\mu, F^{\mu\nu}] = 2\pi \, J^\nu \,,
\tag{10.60}
$$

where

$$
J_\mu = (D_\mu \phi) \, \phi^\dagger - \phi \, (D_\mu \phi)^\dagger \,.
\tag{10.61}
$$

Note that is is a 2×2 matrix current, as it should be, since the Hermitean-conjugated 2-vector sits to the right of the unconjugated 2-vector.[37]

This completes this brief introduction to Yang-Mills theory. As far as applications are concerned, it is fair to say that Yang-Mills theory lies at the heart of modern fundamental physics. The weak nuclear force is described by the Weinberg-Salam model, based on the Yang-Mills gauge group $SU(2)$. The W and Z bosons, which have been seen in particle accelerators such as the one at CERN, are the $SU(2)$ gauge fields. The strong nuclear force is described by a Yang-Mills theory with $SU(3)$ gauge group, and the 8 gauge fields associated with this theory are the gluons that mediate the strong interactions. Thus one may say that almost all of modern particle physics relies upon Yang-Mills theory.

---

[37]It is helpful to introduce an index notation to label the rows and columns of the 2×2 matrices, in order to verify that (10.61) is the correct expression.