# B2: Symmetry and Relativity

Toby Adkins

June 19, 2017

# Contents

# 1. *An Introduction to Flat Spacetime*

This chapter aims to introduce the fundamental concepts used in the treatment of Special Relativity, including:

- Tensor Mathematics
- Minkowski Space
- Time, position, and kinematics
- Symmetries

Some of the material included in this chapter will already be familiar to readers who have had some experience of Special Relativity. The first two chapters focus on setting up the mathematical apparatus and basic transformations that underpin the study of Special Relativity, so if readers are already familiar with this, it is recommended that they skip immediately to the third section.

## 1.1 Tensor Mathematics

The study of Special and General Relativity is achieved through an entirely different language than, say, Quantum Mechanics. Instead of making use of Dirac notation and the associated wavefunctions, we make use of tensor objects, which bring along with themselves associated mathematical apparatus. In this section, we shall introduce tensors and tensor notation, which shall be used - where appropriate - throughout the remainder of this text.

Tensors are geometric objects that describe linear relationships between vectors, scalars and other tensors. More specifically, we can define a tensor $\mathbb{A}$ of valence $(p,q)$ as an assignment of a multidimensional array

$$\mathbb{A}^{i_1,\ldots,i_p}_{j_1,\ldots,j_q}[\mathbf{b}] \tag{1.1}$$

to each basis $\mathbf{b} = \{\mathbf{e}_0,\ldots,\mathbf{e}_{n-1}\}$ of a fixed $n$-dimensional vector space such that, if we apply the change of basis (corresponding to a change of coordinates)

$$\mathbf{b} \mapsto \mathbf{b}' = \mathbf{b}\cdot\Lambda = \{\mathbf{e}_i\Lambda^i_0,\ldots,\mathbf{e}_i\Lambda^i_{n-1}\} \tag{1.2}$$

then the multidimensional array obeys the transformation law

$$\mathbb{A}^{i'_1,\ldots,i'_p}_{j'_1,\ldots,j'_q}[\mathbf{b}'] = \left(\Lambda^{-1}\right)^{i'_1}_{i_1}\ldots\left(\Lambda^{-1}\right)^{i'_p}_{i_p}\mathbb{A}^{i_1,\ldots,i_p}_{j_1,\ldots,j_q}[\mathbf{b}]\Lambda^{j_1}_{j'_1}\ldots\Lambda^{j_q}_{j'_q} \tag{1.3}$$

As mathematical definitions usually are, this may seen quite abstract upon first reading. However, at this stage, we can already tease out some informative points from this definition.

The upper and lower indices $i_1,\ldots,i_p$ and $j_1,\ldots,j_p$ are index notation used to refer to the components of the tensor (as we saw in first year with linear algebra). These are often referred to as the *contravariant* and *covariant* indices respectively; a tensor of only contravariant or covariant indices is said to be itself contravariant or covariant, but if it includes both, it is said to be mixed. The *valence* $(p,q)$ specifies the number of contravariant and covariant indices, and the *rank-$r$* of a tensor is simply the total number of indices required to label a component in the array: $r = p+q$.

The expression on the right-hand side of (1.3) informs us of the way that tensors transform under a change of basis. For a tensor of rank-$r$, we require $r$ applications of the transformation matrix $\Lambda$ and it's inverse $\Lambda^{-1}$. Note that at this stage, we have not made explicit exactly what $\Lambda$ is, apart from the fact that it is some general transformation matrix of rank-2.

### 1.1.1 A Basis and the Metric

Consider the basis $\mathbf{b} = \{\mathbf{e}_0,\ldots,\mathbf{e}_{n-1}\}$. Then, we can expand a rank-1 tensor $A$ in this basis as

$$\mathsf{A} = A^\mu\mathbf{e}_\mu \tag{1.4}$$

$A^\mu$ are the components of the tensor, and the sum is over repeated $\mu$, except in this case $\mu$ indexes the particular basis vector $\mathbf{e}_\mu$ in $\mathbf{b}$. Note that is is not a contraction as outlined in section 1.1.2. Now, we introduce the *dual basis* $\widetilde{\mathbf{b}} = \{\mathbf{e}^0,\ldots,\mathbf{e}^{n-1}\}$. The same tensor can also be expanded in this dual basis as

$$\mathsf{A} = A_\nu\mathbf{e}^\nu \tag{1.5}$$

where once again $A_\nu$ are the components of the rank-1 tensor, and $\mathbf{e}^\nu$ in $\widetilde{\mathbf{b}}$. The basis and the dual basis are related by orthogonality relationship

$$\mathbf{e}_\mu \mathbf{e}^\nu = \delta^\nu_\mu \tag{1.6}$$

The differential line element within this space is then given by

$$ds^2 = (\mathbf{e}_\mu dx^\mu)(\mathbf{e}_\nu dx^\nu) = \mathbf{e}_\mu \mathbf{e}_\nu dx^\mu dx^\nu \tag{1.7}$$

We define the *metric tensor* $g_{\mu\nu}$ as

$$\boxed{g_{\mu\nu} = \mathbf{e}_\mu \mathbf{e}_\nu} \tag{1.8}$$

This is a rank-2 tensor of coefficients that relate the differential spatial elements $dx$ to the distance within the space. Noting that $dx^\mu dx^\nu$ is a symmetric quantity, it is clear that the metric must also be symmetric in order for $ds^2$ to be non-zero:

$$\boxed{g_{\mu\nu} = g_{\nu\mu}} \tag{1.9}$$

Equivalently, we can also write the metric tensor in terms of the dual basis as

$$g^{\mu\nu} = \mathbf{e}^\mu \mathbf{e}^\nu \tag{1.10}$$

Equations (1.8) and (1.10) tell us that $g_{\mu\nu}$ is the coefficient of $\mathbf{e}^\nu$ in the expansion of the vector $\mathbf{e}_\mu$ in the usual basis

$$\mathbf{e}_\mu = g_{\mu\nu}\mathbf{e}^\nu, \tag{1.11}$$

and that $g^{\mu\nu}$ is the coefficient of $\mathbf{e}_\mu$ in the expansion of the vector $\mathbf{e}^\nu$ in the dual basis:

$$\mathbf{e}^\nu = g^{\mu\nu}\mathbf{e}_\mu \tag{1.12}$$

Lastly, it also follows from these equations that

$$\boxed{g_{\mu\nu} g^{\nu\rho} = \delta^\rho_\mu} \tag{1.13}$$

### 1.1.2 Tensor Algebra Rules

In this section, we introduce the commonly used legal operations in tensor algebra. These shall be very similar to the manipulations that readers will be used to when dealing with quantities in linear algebra, except that more care needs to be taken when it comes to the contravariant and covariant nature of indices.

From a notational perspective, tensors of rank-0 shall generally be written as $A$, rank-1 shall be denoted by characters characters of the form $\mathsf{A}$, and rank-2 tensors and above shall be denoted using the more calligraphic characters $\mathbb{A}$. For indices, the Greek alphabet shall usually be adopted, though some Roman Letters shall also be used under certain circumstances.

#### Linear Sum

The sum of two tensors is a linear operation. Let $\mathbb{A}$, $\mathbb{B}$ and $\mathbb{C}$ be tenors of the same valence. Then, we have the well-defined operations

$$\mathbb{C}^{\mu\nu\dots}_{\rho\sigma\dots} = \mathbb{A}^{\mu\nu\dots}_{\rho\sigma\dots} + \mathbb{B}^{\mu\nu\dots}_{\rho\sigma\dots} \tag{1.14}$$

$$(\alpha\mathbb{C})^{\mu\nu\dots}_{\rho\sigma\dots} = \alpha\, \mathbb{C}^{\mu\nu\dots}_{\rho\sigma\dots} \tag{1.15}$$

for some scalar constant $\alpha$. That is, the sum of two tensors of the same valence produces another tensor of the same valence, with the components added entry by entry, and that a scalar can simply be pulled out the front of a tensor if it appears in every entry of said tensor.

### Outer Product

Consider two tensors $\mathbb{A}$ and $\mathbb{B}$. Then, their outer product - often denoted by $\mathbb{A} \otimes \mathbb{B}$ - is defined by

$$\mathbb{A}^{\mu\nu}\mathbb{B}^{\rho}_{\sigma} = (\mathbb{A}\mathbb{B})^{\mu\nu\rho}_{\sigma} = \mathbb{C}^{\mu\nu\rho}_{\sigma} \tag{1.16}$$

The tensor $\mathbb{C}$ is a tensor whose rank is the sum of the tensors that make up the outer product. As an example, let us consider the outer product of two rank-1 tensor objects given by

$$\mathsf{A}^{\mu} = (A^0, A^1, A^2, A^3) \quad \text{and} \quad \mathsf{B}^{\nu} = (B^0, B^1, B^2, B^3) \tag{1.17}$$

Then, their outer product can be written explicitly as

$$\mathbb{C}^{\mu\nu} = \mathsf{A}^{\mu}\mathsf{B}^{\nu} = \begin{pmatrix} A^0B^0 & A^0B^1 & A^0B^2 & A^0B^3 \\ A^1B^0 & A^1B^1 & A^1B^2 & A^1B^3 \\ A^2B^0 & A^2B^1 & A^2B^2 & A^2B^3 \\ A^3B^0 & A^3B^1 & A^3B^2 & A^3B^3 \end{pmatrix} \tag{1.18}$$

It is clear that the outer product is equivalent to a matrix multiplication of $A^T$ and $B$ (if $A$ and $B$ are row vectors, as above). The outer product allows us to form higher-rank tensors out of lower rank ones.

### Contraction

Suppose that a particular index appears in a tensor object as both a contravariant and covariant index. Then, a *contraction* or summation is implied over said index:

$$\mathbb{A}^{\mu\nu\rho}\mathbb{B}_{\sigma} \xrightarrow{\rho=\sigma} \sum_{\rho} \mathbb{A}^{\mu\nu\rho}\mathbb{B}_{\rho} \equiv \mathbb{A}^{\mu\nu\rho}\mathbb{B}_{\rho} = \mathbb{C}^{\mu\nu} \tag{1.19}$$

The second equality follows from what is known as the *Einstein summation convention*; any repeated index contains an implied sum, and we do not have to write out the summation notation at the start of our expression. It is clear that contraction decreases the rank of the original tensor by two; the left hand side was a rank-4 tensor, while the right-hand side is a tensor of rank-2.

From (1.11) and (1.12), it is clear that $g_{\mu\nu}$ can be used to lower indices tensors under contraction, while $g^{\mu\nu}$ can be used to raise the indices of tensors under contraction:

$$g_{\mu\nu}\mathbb{A}^{\mu\rho} = \mathbb{A}^{\rho}_{\nu}, \quad g^{\mu\nu}\mathbb{A}_{\mu\rho} = \mathbb{A}^{\nu}_{\rho} \tag{1.20}$$

This is a very useful algebraic manipulation, as it allows us to do away with, or generate, metric tensors in our expressions when required.

### Differentiation

We notate the partial derivatives on this space by

$$\partial_{\mu} = \frac{\partial}{\partial x^{\mu}}, \quad \partial^{\mu} = \frac{\partial}{\partial x_{\mu}} = g^{\mu\nu}\partial_{\nu} \tag{1.21}$$

Such derivatives follow the normal rules of differentiation (commutativity, product rule). In particular, we shall note here the following identities:

$$\boxed{\begin{aligned} \partial_{\mu}x^{\nu} &= \delta^{\nu}_{\mu} \qquad (1.22) \\ \partial_{\mu}x_{\nu} &= g_{\mu\nu} \qquad (1.23) \end{aligned}}$$

### 1.1.3 Tensors in $\mathbb{R}^4$

We now adopt the basis $\mathbf{b} = \{\mathbf{e}_0, \mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$, such that we are working in $\mathbb{R}^4$. In Special Relativity, we are generally concerned with rank-2 tensors and below, and so we shall now examine the form and transformations of these different ranks of tensor.

**Rank-2 Tensors**

For the purposes of this text, we shall usually only consider tensors that are either fully contravariant or covariant, so either $\mathbb{A}^{\mu\nu}$ or $\mathbb{A}_{\mu\nu}$. Explicitly, the tensor takes the form

$$\mathbb{A}_{\mu\nu} = \left(\begin{array}{c|ccc} A_{00} & A_{01} & A_{02} & A_{03} \\ \hline A_{10} & A_{11} & A_{12} & A_{13} \\ A_{20} & A_{21} & A_{22} & A_{23} \\ A_{30} & A_{31} & A_{32} & A_{33} \end{array}\right) \tag{1.24}$$

The entry in the upper left-hand corner is the *time-time* component, and the entries $A_{0i}$ and $A_{i0}$ for $i = 1, 2, 3$ are the *time-space* and *spacetime* components. The remainder of the tensor is known as the *space-space* part. As we shall see later, these areas of the matrix have particular physical significance.

As anticipated by (1.1), rank-2 tensors transform with two applications of our transformation matrix $\Lambda$

$$\mathbb{A}' = \Lambda^T \mathbb{A} \Lambda \quad \longleftrightarrow \quad \mathbb{A}'^{\rho\sigma} = \Lambda^\rho{}_\mu \Lambda^\sigma{}_\nu \mathbb{A}^{\mu\nu} \tag{1.25}$$

where the prime is the conventional notation used to refer to the new frame/set of coordinates.

**Rank-1 Tensors**

Rank-1 tensors in $\mathbb{R}^4$ are most commonly referred to as *four-vectors*, as they feature four components, and can often be represented as a row/column vector:

$$\mathsf{A}_\mu = (A_0, A_1, A_2, A_3) \tag{1.26}$$

The entry $A_0$ is known as the *time* component, while those corresponding to $A_i$ for $i = 1, 2, 3$ are the *spatial* components. The reason behind this nomenclature shall soon become evident. Generally, the indices $i$ and $j$ shall be reserved to refer only to spatial components of tensors.

Four-vectors transform differently depending on whether they are in contravariant or covariant form. Explicitly, we have that

$$\text{Contravariant}: \quad \mathsf{A}'^\nu = \Lambda^\nu{}_\mu \mathsf{A}^\mu \tag{1.27}$$

$$\text{Covariant}: \quad \mathsf{A}'_\nu = (\Lambda^{-1})_\nu{}^\mu \mathsf{A}_\mu \tag{1.28}$$

**Rank-0 Tensors**

Rank-0 tensors are simply scalar quantities, meaning that they have no free index. In other words, all indices are summed over. We can create such scalar quantities by contracting over indices:

$$\mathsf{A}^\mu \mathsf{A}_\mu \quad \text{or} \quad \mathbb{A}^{\mu\nu}\mathbb{A}_{\mu\nu} \tag{1.29}$$

Having no free index, it is clear from (1.1) that scalars of this form are invariant under the change of basis. We shall show this explicitly when we deal with the Lorentz group in the next section.

## 1.2 Minkowski Space

Special Relativity is the consideration of motion in *flat spacetime*, corresponding to a constant geometry throughout the space. As such, we adopt the Minkowski metric given by

$$g_{\mu\nu} \equiv \eta_{\mu\nu} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{1.30}$$

This metric shall be assumed throughout the remainder of this text. The sign convention $\eta_{\mu\nu} = \mathrm{diag}(+,-,-,-)$ can also be used; all this does is change the sign of components in certain calculations but, thankfully, leaves the physics unchanged. This alternative sign convention is most often used in Subatomic Physics, but is less useful for Special Relativity, General Relativity, and Cosmology.

### 1.2.1 The Lorentz Group

Our differential line element within Minkowski space is then given by

$$ds^2 = \eta_{\mu\nu} dx^\mu dx^\nu = -c^2 dt^2 + d\mathbf{x}^2 \tag{1.31}$$

This is often referred to as *the interval.* Now, suppose that we want to find some non-singular transformation $dx \mapsto dx'$ such that the interval remains invariant. What form do said transformations take? Consider the transformed interval:

$$ds'^2 = \eta_{\rho\sigma} dx'^\rho dx'^\sigma = \eta_{\rho\sigma} \frac{\partial x'^\rho}{\partial x^\mu} \frac{\partial x'^\sigma}{\partial x^\nu} dx^\mu dx^\nu \tag{1.32}$$

For it to be the case that $ds'^2 = ds^2$ for arbitrary $dx^\mu$, then the condition

$$\eta_{\mu\nu} = \eta_{\rho\sigma} \frac{\partial x'^\rho}{\partial x^\mu} \frac{\partial x'^\sigma}{\partial x^\nu} \tag{1.33}$$

must be satisfied. Differentiation with respect to an arbitrary coordinate $x^\lambda$ gives

$$0 = \eta_{\rho\sigma} \frac{\partial x'^\rho}{\partial x^\mu \partial x^\lambda} \frac{\partial x'^\sigma}{\partial x^\nu} + \eta_{\rho\sigma} \frac{\partial x'^\rho}{\partial x^\mu} \frac{\partial x'^\sigma}{\partial x^\nu \partial x^\lambda} \tag{1.34}$$

To solve for second derivatives, we add to this the same equation with $\mu$ and $\lambda$ interchanged, and subtract the same with $\lambda$ and $\nu$ interchanged (this is a common trick with these sort of manipulations):

$$\begin{aligned} 0 = \eta_{\rho\sigma} \Bigg[ & \frac{\partial x'^\rho}{\partial x^\mu \partial x^\lambda} \frac{\partial x'^\sigma}{\partial x^\nu} + \cancel{\frac{\partial x'^\rho}{\partial x^\mu} \frac{\partial x'^\sigma}{\partial x^\nu \partial x^\lambda}} + \frac{\partial x'^\rho}{\partial x^\lambda \partial x^\mu} \frac{\partial x'^\sigma}{\partial x^\nu} + \cancel{\frac{\partial x'^\rho}{\partial x^\lambda} \frac{\partial x'^\sigma}{\partial x^\nu \partial x^\mu}} \\ & - \cancel{\frac{\partial x'^\rho}{\partial x^\mu \partial x^\nu} \frac{\partial x'^\sigma}{\partial x^\lambda}} - \cancel{\frac{\partial x'^\rho}{\partial x^\mu} \frac{\partial x'^\sigma}{\partial x^\lambda \partial x^\nu}} \Bigg] \end{aligned} \tag{1.35}$$

Using the fact that the metric is symmetric, this can be written as

$$0 = \eta_{\rho\sigma} \frac{\partial x'^\rho}{\partial x^\mu \partial x^\lambda} \frac{\partial x'^\sigma}{\partial x^\nu} \quad \longrightarrow \quad 0 = \frac{\partial x'^\rho}{\partial x^\mu \partial x^\lambda} \tag{1.36}$$

where the second expression follows from the fact that both $\eta_{\rho\sigma}$ and $\partial x'^\sigma / \partial x^\nu$ are non-singular matrices. The solution to this is clearly some linear function

$$x'^\rho = \Lambda^\rho{}_\sigma x^\sigma + a^\rho \tag{1.37}$$

where $a^\rho$ is constant. (1.33) makes it clear that our transformation $\Lambda$ must satisfy

$$\boxed{\Lambda^T\eta\Lambda = \eta \quad\longleftrightarrow\quad \eta_{\rho\sigma}\Lambda^\rho{}_\mu\Lambda^\sigma{}_\nu = \eta_{\mu\nu}, \quad \Lambda^\rho{}_\mu = \frac{\partial x'^\rho}{\partial x^\mu}} \tag{1.38}$$

The set of all Lorentz transformations $\Lambda$ of the above from is known as the *inhomogeneous Lorentz group*, where the *homogeneous Lorentz group* are those that satisfy the above with $a^\rho = 0$. The inverse transformation can be obtained easily from this condition. Using (1.20), we have that $\Lambda_{\sigma\mu}\Lambda^\sigma{}_\nu = \eta_{\mu\nu}$. Raising $\mu$ on both sides of this equation gives $\Lambda_\sigma{}^\mu\Lambda^\sigma{}_\nu = \delta^\mu_\nu$. From this, it is clear that the inverse transformation is given by

$$\boxed{(\Lambda^{-1})^\mu{}_\sigma = \Lambda_\sigma{}^\mu} \tag{1.39}$$

This means that one simply has to change the sign of the components for which only one of the indices is zero (namely, $\Lambda^0{}_i$ and $\Lambda^i{}_0$), and take the associated transpose.

The above condition (1.38) has an important consequence that is worth examining here. Suppose that $\mathsf{A}^\rho$ is any contravariant four-vector, and $\mathsf{B}_\rho$ is any covariant four-vector. Then, the scalar $\mathsf{A}^\rho\mathsf{B}_\rho$ must be a frame invariant quantity - that is, independent of our choice of coordinates - as:

$$\mathsf{A}'^\rho\mathsf{B}'_\rho = \mathsf{A}'^\rho\mathsf{B}'^\sigma\eta_{\rho\sigma} = \Lambda^\rho{}_\mu\mathsf{A}^\mu\Lambda^\sigma{}_\nu\mathsf{B}^\nu\eta_{\rho\sigma} = \mathsf{A}^\mu\mathsf{B}^\nu\eta_{\mu\nu} = \mathsf{A}^\mu\mathsf{B}_\mu \tag{1.40}$$

This can easily be extended to show that any scalar quantity that is created through a series of contraction operations (regardless of rank) must also be invariant. Invariants are the saving graces in Relativity, as they allow us to compare the same system in two different frames, which can be very useful for extracting information about the physics of the problem.

### Proper and Improper Lorentz subgroups

Taking the determinant of the first expression in (1.38) gives

$$(\det\Lambda)^2 = 1 \tag{1.41}$$

and letting $\mu = \nu = 0$ in the second expression gives

$$(\Lambda^0{}_0)^2 = 1 + \sum_i(\Lambda^i{}_0)^2 \geq 1 \tag{1.42}$$

where as usual the sum runs over the spatial components $i = 1, 2, 3$. This gives rise to two subgroups:

- Proper Lorentz transforms - These are the transforms that take the solution $\Lambda^0{}_0$ and $\det\Lambda = +1$ of the above two equations. It follows that any $\Lambda^\mu{}_\nu$ that can be converted into the identity $\delta^\mu{}_\nu$ by a continuous variation of its parameters must be a proper Lorentz transformation, because it is impossible by a continuous change of parameters to jump from $\Lambda^0{}_0 \leq -1$ to $\Lambda^0{}_0 \geq +1$, or from $\det\Lambda = -1$ to $\det\Lambda = +1$, and the identity has $\Lambda^0{}_0 = +1$ and $\det\Lambda = +1$

- Improper Lorentz transforms - These involve either a spatial inversion ($\det\Lambda = -1$, $\Lambda^0{}_0 \geq 1$) or time reversal ($\det\Lambda = -1$, $\Lambda^0{}_0 \leq -1$), or a product of the two, all of which are known to be non-exact symmetries of nature

It what follows, we will be dealing exclusively with proper, homogeneous Lorentz transformations, as these behave properly in the infinitesimal transformation limit, reducing to the identity. This is required physically, as two frames should be equivalent if an identity-like transformation is made.

### 1.2.2 The Lorentz Transformations

Suppose that one observer $S$ sees a particle at rest in its own frame, while another observer $S'$ sees it moving a arbitrary velocity $\mathbf{v}$. Then from (1.37), we have

$$dx'^{\mu} = \Lambda^{\mu}{}_{\nu} dx^{\nu} \tag{1.43}$$

As $d\mathbf{x}$ vanishes, this can alternatively be written as

$$dx'^{i} = \Lambda^{i}{}_{0} c dt \quad \text{and} \quad dt' = \Lambda^{0}{}_{0} dt \tag{1.44}$$

Dividing $d\mathbf{x}'$ by $dt'$ gives the relative velocity $\mathbf{v}$, such that

$$c\Lambda^{i}{}_{0} = v_i \Lambda^{0}{}_{0} \tag{1.45}$$

We can obtain another relationship between $\Lambda^{i}{}_{0}$ and $\Lambda^{0}{}_{0}$ by rearranging (1.42)

$$-1 = \sum_i (\Lambda^{i}{}_{0})^2 - (\Lambda^{0}{}_{0})^2 \tag{1.46}$$

The solution to (1.45) and (1.46) is

$$\Lambda^{0}{}_{0} = \gamma, \quad \Lambda^{i}{}_{0} = -\gamma\beta_i, \quad \Lambda^{0}{}_{j} = -\gamma\beta_j \tag{1.47}$$

where we have defined the *Lorentz factor* $\gamma$ as

$$\boxed{\gamma = (1-\beta)^{-1/2}, \quad \boldsymbol{\beta} = \frac{\mathbf{v}}{c}} \tag{1.48}$$

The third expression in (1.47) follows by performing the same calculation as above, but for covariant forms of the differential elements. The other components of $\Lambda^{\mu}{}_{\nu}$ are not uniquely determined, as space-space part may be formed from any arbitrary set of rotations. One convenient choice that satisfies (1.38) is

$$\Lambda^{i}{}_{j} = \delta^{i}_{j} + \beta_i\beta_j \frac{\gamma - 1}{\beta^2} \tag{1.49}$$

It is also easy to see that the components in (1.47) and (1.49) reduce to the identity in the limit that $\mathbf{v} \to 0$, corresponding to the two frames $S$ and $S'$ being equivalent.

#### General Form

Putting these components together, it follows that we can write the general Lorentz transformations for some relative velocity $\mathbf{v}$ between the two frames as

$$\Lambda = \begin{pmatrix} \gamma & -\gamma\beta_x & -\gamma\beta_y & -\gamma\beta_z \\ \cdot & 1 + \frac{\gamma^2}{1+\gamma}\beta_x^2 & 1 + \frac{\gamma^2}{1+\gamma}\beta_x\beta_y & 1 + \frac{\gamma^2}{1+\gamma}\beta_x\beta_z \\ \cdot & \cdot & 1 + \frac{\gamma^2}{1+\gamma}\beta_y^2 & 1 + \frac{\gamma^2}{1+\gamma}\beta_y\beta_z \\ \cdot & \cdot & \cdot & 1 + \frac{\gamma^2}{1+\gamma}\beta_z^2 \end{pmatrix} \tag{1.50}$$

where we have made use of the identity (1.88). Let us now introduce the four-displacement $\mathsf{X}$ given by

$$\boxed{\mathsf{X} = (ct, \mathbf{x}) = (ct, x^1, x^2, x^3)} \tag{1.51}$$

where $t$ and $\mathbf{x}$ are the time and space coordinates of an event in some frame $S$. It becomes clear that we can also write the general Lorentz transforms component-wise as

$$ct' = \gamma(ct - \boldsymbol{\beta}\cdot\mathbf{x}) \tag{1.52}$$

$$\mathbf{x}' = \mathbf{x} + \left(-\gamma ct + \frac{\gamma - 1}{\beta^2}\boldsymbol{\beta}\cdot\mathbf{x}\right)\boldsymbol{\beta} \tag{1.53}$$

### One-dimensional

Evidently, the one-dimensional form of the transformations is a simple case of the transformations outlined in the previous section. However, we shall be making such extensive use of this particular case that is worth treating it separately here. Two frames $S$ and $S'$ are said to be in *standard configuration* when their relative velocity $\mathbf{v}$ is aligned along the positive $\mathbf{e}_1$ direction (more commonly known as the $x$-axis), such that the origin of both coordinate systems coincide at $t = 0$. Then, they are related by the Lorentz transformation

$$\Lambda = \begin{pmatrix} \gamma & -\beta\gamma & 0 & 0 \\ -\beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{1.54}$$

or in component form

$$\boxed{\begin{aligned} ct' &= \gamma\,(ct - vx/c) && (1.55) \\ x' &= \gamma(x - vt) && (1.56) \\ y' &= y && (1.57) \\ z' &= z && (1.58) \end{aligned}}$$

Using (1.39), we shall also make an explicit note of the inverse transformation

$$\Lambda^{-1} = \begin{pmatrix} \gamma & \beta\gamma & 0 & 0 \\ \beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{1.59}$$

Of course, the forms of (1.54) and (1.59) are reversed in the case that the relative velocity is aligned along the negative $\mathbf{e}_1$ direction.

### Classical Limit

Let us round off our consideration of the Lorentz transformation by considering its classical or Newtonian limit. In Special Relativity, taking the classical limit corresponds to letting $c \to \infty$. In this limit, we have that $\gamma \approx 1$, such that the transformations become

$$t' = t \tag{1.60}$$
$$x' = x - vt \tag{1.61}$$
$$y' = y \tag{1.62}$$
$$z' = z \tag{1.63}$$

which are simply the familiar Galilean transformations. Thus, $\Lambda$ is well behaved in both the unitry and classical limits.

## 1.3 Time, position and kinematics

Almost all of Special Relativity in encapsulated in the Lorentz transformations (1.52) and (1.53), as they allow us to relate any frame in Minkowski space (flat spacetime) to any other. Let us now examine some of the basic consequences of these transformations.

### 1.3.1 Proper Time

In section 1.2.2, we introduced the displacement four-vector $\mathsf{X}^\mu$ that describes the position of a particle or event in spacetime. The *wordline* of a particle is the path of said particle in our four-dimensional space, tracing the history of it's coordinates at each instant in time. This means that if $\mathsf{X}^\mu$ does indeed describe a particle, that we can think of $\mathsf{X}^\mu$ as actually being a function of some parameter $\tau$ that parametrises the motion of the particle along it's worldline.

As such, we define the *proper time $\tau$ as the time measured by a clock following a worldline.* In other words, it is the time measured by a clock that is in the instantaneous rest frame of the observer that we are interested in. We then define the proper time by

$$\boxed{c^2 d\tau^2 = -ds^2 = -\eta_{\mu\nu} dx^\mu dx^\nu} \tag{1.64}$$

with proper time interval moving along some worldline $\ell$ being given by

$$\tau = \int_\ell d\tau = \frac{1}{c}\int_\ell \sqrt{-\eta_{\mu\nu} dx^\mu dx^\nu} \tag{1.65}$$

The Lorentz transformations were derived in the previous section on the basis that they leave the interval invariant; as the proper time is related to the interval by a multiplicative constant, this means that the Lorentz transformations also preserve the proper time.

#### Time-Dilation

Consider two frames $S$ and $S'$ in standard configuration with velocity $\mathbf{v}$. Suppose that we observe a clock at rest in frame $S'$ from frame $S$. By proper time invariance, we have that

$$-c^2 d\tau'^2 = -c^2 d\tau^2 = -c^2 dt^2 \left(1 - \frac{1}{c^2}\left|\frac{d\mathbf{x}}{dt}\right|^2\right) \tag{1.66}$$

as above. Re-arranging this expression allows us to obtain the traditional time-dilation result that

$$\boxed{\frac{dt}{d\tau} = \gamma} \tag{1.67}$$

where the $\gamma$ factor is the one that is associated with the velocity of the particle in the reference frame in which the frame time $t$ is calculated.

Now, consider the propagation of a wavefront of light. Then, $|d\mathbf{x}/dt|$ in frame $S$ is equal to the speed of light, such that

$$-c^2 d\tau^2 = -c^2 dt^2 \left(1 - \frac{1}{c^2}\left|\frac{d\mathbf{x}}{dt}\right|^2\right) = 0 \tag{1.68}$$

This means that the propagation of light is described by the statement that $d\tau = 0$. Performing a Lorentz transformation to frame $S'$ does not change $d\tau$ by definition, meaning that $d\tau'^2 = 0$, and therefore that $|d\mathbf{x}'/dt| = c$. That is, the speed of light is constant across all frames of reference. We have thus obtained one of Einstein's postulates of Special Relativity from our considerations of interval invariance.

**Timelike and Spacelike Vectors**

From (1.40), it is clear that quantities of the form $\mathsf{A}\cdot\mathsf{B} = \eta_{\mu\nu}\mathsf{A}^\mu\mathsf{B}^\nu = \mathsf{A}^\mu\mathsf{B}_\mu$ are Lorentz invariant. For a specific vector of the form

$$\mathsf{A} = (A^0, A^1, A^2, A^3) \tag{1.69}$$

we have that

$$\mathsf{A}^\mu\mathsf{A}_\mu = -(A^0)^2 + (A^1)^2 + (A^2)^2 + (A^3)^2 \tag{1.70}$$

Depending on this sign of this invariant, we can classify a four-vector into three different types:

1. Timelike ($\mathsf{A}^\mu\mathsf{A}_\mu < 0$) - A timelike vector connects two events that are causally connected. The timelike vector can be considered to define a four-velocity direction of an observer, and thus a time axis of said observer

2. Spacelike ($\mathsf{A}^\mu\mathsf{A}_\mu > 0$) - A spacelike vector connects two events that are causally disconnected, meaning that it can be thought of as defining a spatial direction of an observer

3. Null ($\mathsf{A}^\mu\mathsf{A}_\mu = 0$) - Null vectors can be seen as a vector type that separates the timelike and spacelike vectors. More specifically, null vectors represent quantities that are associated with the propagation of light ($d\tau = 0$)

These definitions may seem a little abstract to begin with, but shall become more clear when we cover spacetime diagrams in the next section. From these definitions, we can deduce some properties of these different types of vectors:

- *For any timelike vector, there exists a frame in which its spatial part is zero* - Without loss of generality, we can choose our coordinate system such that

$$\mathsf{A} = (A^0, A^1, 0, 0) \tag{1.71}$$

We want to show that

$$\mathsf{A}'^\nu = \Lambda^\nu{}_\mu\mathsf{A}^\mu = (A'^0, 0, 0, 0) \tag{1.72}$$

is possible for a timelike vector under a Lorentz transformation oriented along $\mathbf{e}_1$. We must then have that

$$A'^1 = -\beta\gamma A^0 + \gamma A^1 = 0 \quad \longrightarrow \quad \beta = \frac{A^1}{A^0} \tag{1.73}$$

Thus, a Lorentz transformation with the above velocity will cause the spatial part to be zero. This is only possible for $A^1/A^0 < 0$, meaning that we require the vector to be timelike

- *Any vector orthogonal to a timelike vector must be spacelike* - Consider the two four-vectors

$$\mathsf{A} = (A^0, A^1, 0, 0), \quad \mathsf{B} = (B^0, B^1, 0, 0) \tag{1.74}$$

If these two vectors are to be orthogonal (or in other words, their product is null), it must be the case that

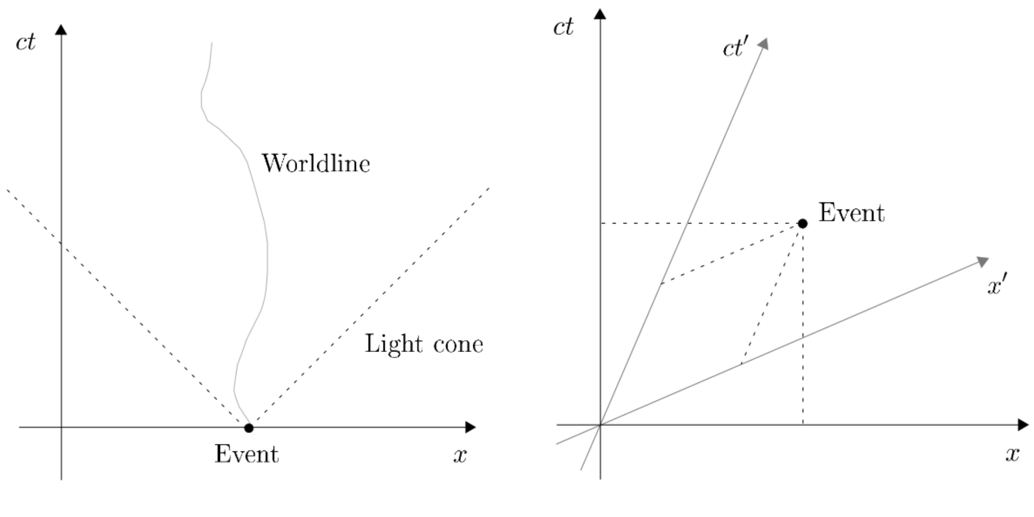
$$\mathsf{A}^\mu\mathsf{B}_\mu = -A^0B^0 + A^1B^1 = 0 \tag{1.75}$$

If $A^\mu$ is timelike, it must be the case that $A^0 > A^1$. From the above, this implies that $B^0 < B^1$, meaning that $B^\mu$ must be spacelike

- *Any vector orthogonal to a null vector is null, or spacelike* - Consider the two four-vectors
$$\mathsf{A} = (A^0, A^0, 0, 0), \quad \mathsf{B} = (B^0, B^1, 0, 0) \tag{1.76}$$
where $\mathsf{A}^\mu$ is clearly null. Then
$$\mathsf{A}^\mu \mathsf{B}_\mu = -A^0 B^0 + A^0 B^1 = -A^0(-B^0 + B^1) = 0 \tag{1.77}$$
meaning that $\mathsf{B}^\mu$ is also null. This will not be the case if $\mathsf{B}^\mu$ has components in $\mathbf{e}_2$ and $\mathbf{e}_3$, meaning that it will then be spacelike

### 1.3.2 Spacetime Diagrams

A useful analytical tool in Special Relativity is that of a spacetime diagram. In essence, they are a graph of spacetime, in which we can represent events, and the worldlines of particles. In two dimensions, the time part is typically represented on the vertical axis, with the spatial part represented on the horizontal axis, as in figure 1.1 (a). From the axes, it is clear that the propagation of light should be represented with a straight line with a gradient of unity.




(a) (b)

Figure 1.1: (a) An example spacetime diagram showing an event, with associated worldline and light cone (b) A spacetime diagram showing an event as viewed from two frames $S$ and $S'$

We can consider any event to correspond to a particular point on a worldline. With each event, we associate a *light cone*, which essentially represents the propagation of light outwards from this event. Events can only effect other events that occur within their light cone, as the information about this event can only propagate outwards at the speed of light. This is the reason for the causality associated with the definitions of timelike, null, and spacelike vectors; it is clear that these lie inside, on, and outside the lightcone respectively.

We can also represent multiple frames on the single diagram. For example, figure 1.1 (b) shows an event in $S$, and another frames $S'$ that is moving with a positive velocity with respect to $S$. We have included lines describing the locus of events that occur at the same time (simultaneous) or same position (coincident) in both frames.

### 1.3.3 Position and Velocity

We have already introduced the four displacement $\mathsf{X}$ that we use to represent the spacetime coordinates of a particular event, or describe an particle's worldline. This clearly has invariant

$$\boxed{\eta_{\mu\nu}\mathsf{X}^\mu\mathsf{X}^\nu = -c^2\tau^2} \tag{1.78}$$

We then define the four-velocity of a particle as the time rate of change $\mathsf{X}^\mu$, such that the ratio of $d\mathsf{X}$ to some small time interval is Lorentz invariant. The natural definition to adopt is

$$\boxed{\mathsf{U} = \frac{d\mathsf{X}}{d\tau} = \gamma(c, \mathbf{u}), \quad \eta_{\mu\nu}\mathsf{U}^\mu\mathsf{U}^\nu = -c^2} \tag{1.79}$$

where we have used the chain rule and (1.67). Note that $\gamma \equiv \gamma_u$ is the Lorentz factor corresponding to the velocity of the particle in the frame in which it is measured. How can we simply write down the invariant as $-c^2$? As invariant quantities are, well, invariant, we can evaluate the invariant in any frame that we choose. Thus, we can evaluate it in the frame where $\mathbf{u} = 0$, such that $\gamma = 1$ and $\mathsf{U} = (c, 0)$, which clearly gives the desired result.

#### Transformation of Velocities

Consider two frames $S$ and $S'$ in standard configuration. From equations (1.55) and (1.56), we have that

$$t' = \gamma_v(t - vx/c^2) = \gamma_v t(1 - u_x v/c^2) \tag{1.80}$$
$$x' = \gamma_v(x - vt) = \gamma_v t(u_x - v) \tag{1.81}$$
$$y' = y \tag{1.82}$$

where we have used the fact that $x = u_x t$ in frame $S$. Then, we clearly have that

$$u_x' = \frac{dx'}{dt'} = \frac{u_x - v}{1 - u_x v/c^2}, \quad u_y = \frac{dy'}{dt'} = \frac{u_y}{\gamma_v(1 - u_y v/c^2)} \tag{1.83}$$

We can perform a similar calculation for $v$ orientated along $\mathbf{e}_2$ and $\mathbf{e}_3$, meaning that these transformations can be generalised to give

$$\boxed{\mathbf{u}_\parallel' = \frac{\mathbf{u}_\parallel - \mathbf{v}}{1 - \mathbf{u}\cdot\mathbf{v}/c^2}, \quad \mathbf{u}_\perp' = \frac{\mathbf{u}_\perp}{\gamma_v(1 - \mathbf{u}\cdot\mathbf{v}/c^2)}} \tag{1.84}$$

where the parallel and perpendicular directions are defined with respect to the transformation velocity $\mathbf{v}$ between the two frames.

Now, consider the case where we have two particles that have velocities $\mathbf{u}$ and $\mathbf{v}$ in frame $S$. Their four-velocities and associated invariant are

$$\mathsf{U} = \gamma_u(c, \mathbf{u}, \quad \mathsf{V} = \gamma_v(c, \mathbf{v}), \quad \eta_{\mu\nu}\mathsf{U}^\mu\mathsf{V}^\nu = \gamma_u\gamma_v(-c^2 + \mathbf{u}\cdot\mathbf{v}) \tag{1.85}$$

Now, move to a frame where the first particle has zero velocity, and meaning that the second moves at the relative velocity $\mathbf{w} = \mathbf{v} - \mathbf{u}$. In this frame, their four-velocities and associated invariant are

$$\mathsf{U} = (c, 0), \quad \mathsf{V} = \gamma_w(c, \mathbf{w}), \quad \eta_{\mu\nu}\mathsf{U}'^\mu\mathsf{V}'^\nu = -\gamma_w c^2 \tag{1.86}$$

Equating invariants gives us a relationship between the Lorentz factors of the individual particles, and that corresponding to their relative velocity:

$$\boxed{\gamma_w = \gamma_u\gamma_v(1 - \mathbf{u}\cdot\mathbf{v}/c^2)} \tag{1.87}$$

this expression holds true for any three velocities $\mathbf{u}$, $\mathbf{v}$ and $\mathbf{w}$.

**Identities Involving $\gamma$**

The Lorentz factor $\gamma$ crops up repeatedly in our consideration of Special Relativity, and so it is worth noting some identities involving this quantity, as follows.

$$\boxed{\begin{aligned} \frac{\gamma-1}{\beta^2} &= \frac{\gamma^2}{1+\gamma} \quad (1.88) \\ \frac{d\gamma}{dv} &= \frac{\gamma^3 v}{c^2} \quad (1.89) \\ \frac{d}{dv}(\gamma v) &= \gamma^3 \quad (1.90) \end{aligned}}$$

In these identities, $\gamma = (1-v^2/c^2)^{-1/2}$. These are all easily derived by either simple algebraic manipulation. Note also

$$\frac{dt}{d\tau} = \gamma, \quad \frac{dt'}{dt} = \gamma_v(1-\mathbf{u}\cdot\mathbf{v}/c^2), \quad \gamma_w = \gamma_u\gamma_v(1-\mathbf{u}\cdot\mathbf{v}/c^2)$$

from previous sections. These have been included here such that the main identities involving $\gamma$ are all in one place for reference.

### 1.3.4 Wave Motion

To finish off our consideration of kinematics in Special Relativity, let us now consider waves. A general wave motion can be described by an expression of the form

$$\mathbf{a} = \mathbf{a}_0 e^{i(\mathbf{k}\cdot\mathbf{x}-\omega t)} \tag{1.91}$$

where $\mathbf{a}_0$ is some constant amplitude, $\mathbf{k}$ is the wave-vector, and $\omega$ the associated frequency. We now need to ask ourselves what is invariant about this sort of motion? The amplitude of the oscillation is clearly not Lorentz invariant, as this is defined in terms of a length, which may change under Lorentz transformation. However, all observers must agree on events where the displacement is maximal. From this, it follows that the 'wave-crest' locations of the oscillation are Lorentz invariant, and so more generally the phase $\phi = \mathbf{k}\cdot\mathbf{x}-\omega t$, as all frames will agree on how far through the oscillation cycle the motion is.

We can thus define the wave four-vector $\mathsf{K}$ by

$$\boxed{\mathsf{K} = (\omega/c, \mathbf{k}), \quad \eta_{\mu\nu}\mathsf{K}^\mu\mathsf{K}^\nu = \omega^2\left(\frac{1}{v_p^2}-\frac{1}{c^2}\right)} \tag{1.92}$$

where we have introduced the phase velocity $v_p = \omega/k$. When $v_p < c$, the wave four-vector is spacelike, and when $v_p > c$, the wave four-vector is timelike. It is null for $v_p = c$, and this further shows that a wave of any kind whose phase velocity is $c$ in some reference frame will have that same velocity in all reference frames. The speed of light is again constant!

We can also relate $\mathsf{K}$ to our scalar phase $\phi$:

$$\boxed{\phi = \eta_{\mu\nu}\mathsf{K}^\mu\mathsf{X}^\nu = \mathbf{k}\cdot\mathbf{x}-\omega t, \quad \mathsf{K}^\mu = \partial^\mu\phi} \tag{1.93}$$

The last identity follows from applying a contravariant partial derivative to both sides of this equation:

$$\partial^\mu\phi = \partial^\mu\mathsf{K}^\nu\mathsf{X}_\nu = \mathsf{K}^\nu(\partial^\mu\mathsf{X}_\nu) = \mathsf{K}^\nu\delta^\mu{}_\nu = \mathsf{K}^\mu \tag{1.94}$$

### Four-gradient

As an aside, we can now ask the interesting question as to the form that the operator $\partial_\mu$ has to be in order for the last two identities in (1.92) to be consistent. We know that

$$\partial^\mu \phi = \eta^{\mu\nu}\partial_\nu(\mathbf{k}\cdot\mathbf{x}-\omega t) = K^\mu = (\omega/c, \mathbf{k}) \tag{1.95}$$

Writing this in explicit component form, we must have that

$$\eta^{\mu 0}\partial_0(\mathbf{k}\cdot\mathbf{x}-\omega t) = \eta^{00}\partial_0(\mathbf{k}\cdot\mathbf{x}-\omega t) = \mathsf{K}^0 = \omega/c \tag{1.96}$$

$$\eta^{\mu i}\partial_i(\mathbf{k}\cdot\mathbf{x}-\omega t) = \eta^{ii}\partial_i(\mathbf{k}\cdot\mathbf{x}-\omega t) = \mathsf{K}^i = k^i \tag{1.97}$$

where we are again using $i = 1, 2, 3$ to refer to the spatial components of a four-vector, and have used the fact that $\eta_{\mu\nu}$ is zero for non-diagonal components. A sensible solution to these equations is

$$\boxed{\partial_\mu = \left(\frac{1}{c}\frac{\partial}{\partial t}, \nabla\right)} \tag{1.98}$$

This is in fact the definition of the four-gradient, which satisfies all the normal tensor rules of differentiation. It is clear that this is simply the normal gradient operator for the spatial components, and a normalised time derivative for the timelike component. It you are getting confused with minus signs in (1.95), remember that the action of the metric on $\partial_\mu$ is to raise the index, which is equivalent to introducing a minus sign in the timelike component.

### The Doppler Effect

Let us now return to wave motion. Suppose that a wave source in frame $S'$ emits a plane wave with angular velocity $\omega_0$. Let $\omega$ be the observed frequency in frame $S$. Then:

$$\text{In } S'\text{:} \quad \mathsf{K}' = (\omega_0/c, \mathbf{k}_0), \quad \mathsf{U}' = (c, 0) \tag{1.99}$$

$$\text{In } S\text{:} \quad \mathsf{K} = (\omega/c, \mathbf{k}), \quad \mathsf{U} = \gamma(c, \mathbf{v}) \tag{1.100}$$

Equating invariants, we have that

$$\eta_{\mu\nu}\mathsf{K}'^\mu \mathsf{U}'^\nu = \eta_{\mu\nu}\mathsf{K}^\mu \mathsf{U}^\nu = -\omega_0 = \gamma(-\omega + \mathbf{k}\cdot\mathbf{v}) \tag{1.101}$$

Taking the observed phase velocity to be that of light $v_p = \omega/k = c$, we arrive at the formula for the *relativistic Doppler effect*

$$\boxed{\frac{\omega}{\omega_0} = \frac{1}{\gamma(1-(v/v_p)\cos\theta)}} \tag{1.102}$$

Let us consider some special cases of this formula:

- $\theta = 0, \pi$ - This corresponds to the source moving directly towards or away from the observer respectively

  $$\frac{\omega}{\omega_0} = \sqrt{\frac{1\pm\beta}{1\mp\beta}} \tag{1.103}$$

  It is clear that this equation will reduce to the normal Doppler effect formula in the non-relativistic limit where $\beta \ll 1$

- $\theta = \pi/2$ - This corresponds to the source moving across the line of sight of the observer. In this case, the formula reduces to a very simple form

$$\frac{\omega}{\omega_0} = \frac{1}{\gamma} \tag{1.104}$$

This can be interpreted as an example of time dilation; the process of oscillation in the source is slowed down by a factor $\gamma$

*A rapidly rotating star is modelled approximately as a sphere of radius $R$ rotating at an angular frequency $\Omega$ about the $z$ axis in some inertial frame $S$. An observer is far away in the $y$ direction, at rest relative to the centre of the star. Consider the material on the surface of the star in frame $S$ that emits radiation at an angular frequency $\omega_0$. Find an expression for the observed angular frequency $\omega$ in terms of $R$, $\Omega$, and spatial coordinates.*

We want to consider material moving with velocity

$$\mathbf{v} = \dot{\mathbf{x}} = \frac{d}{dt}(R\sin\theta\cos\phi, R\sin\theta\sin\phi, R\cos\theta) \tag{1.105}$$

where $\theta$ and $\phi$ are defined as normal for polar coordinates. Let us assume that the material of the star remains planar (disk like), such that $\theta = \pi/2 =$ constant during the rotation. Observing that $\phi = \Omega t$, it follows that

$$\mathbf{v} = (-y\Omega, x\Omega, 0) \tag{1.106}$$

As the observer is a long way away from the source along $\mathbf{e}_y$, we can approximate that the light is propagating purely along $\mathbf{e}_y$ when it reaches the observer, such that $\mathbf{k} = k\mathbf{e}_y$. Now, considering (1.101), we have that $\mathbf{k}\cdot\mathbf{v} = \Omega x$. The velocity of the moving surface area element is $\rho\Omega$, where $\rho = R^2 - z^2$ is the cylindrical radius. This means that the Lorentz factor becomes

$$\gamma = (1 - v^2/c^2)^{-1/2} = (1 - (\rho\Omega)^2/c^2)^{-1/2} = (1 - (\Omega/c)^2(R^2 - z^2))^{-1/2} \tag{1.107}$$

This means that we obtain the final expression of

$$\frac{\omega}{\omega_0} = \frac{\sqrt{1 - (\Omega/c)^2(R^2 - z^2)}}{1 - \Omega x} \tag{1.108}$$

It is clear that the star will not emit a uniform frequency across its entire surface; for material moving towards the observer $\omega > \omega_0$, that moving away from the observer $\omega < \omega_0$, and $\omega \backsim \omega_0$ for that along $x = 0$.

### Aberration

The change in direction of travel of waves when the same wave is viewed in two different inertial frames is known as *aberration*. In Special Relativity, this just corresponds to the change in direction of a four-vector under a Lorentz transformation.

Suppose that the source emits radiation at an angle of $\theta_0$ with respect to $\mathbf{e}_{x'}$ in frame $S'$, such that $\mathbf{k}_0 = k_0(\cos\theta_0, \sin\theta_0, 0)$. Writing $\mathbf{k} = k(\cos\theta, \sin\theta, 0)$ in frame $S$, we can find a relationship between the angles using the fact that $\mathsf{K} = \Lambda^{-1}\mathsf{K}'$, yielding:

$$\frac{\omega}{c} = \gamma\left(\frac{\omega_0}{c} + \beta k_0\cos\theta_0\right) \tag{1.109}$$

$$k\cos\theta = \gamma\left(\frac{\beta\omega_0}{c} + k_0\cos\theta_0\right) \tag{1.110}$$

Taking the ratio of these two expressions, and using the dispersion relation for light $\omega = ck$, it follows that

$$\boxed{\cos\theta = \frac{\beta + \cos\theta_0}{1 + \beta\cos\theta_0}} \tag{1.111}$$

This leads to what is often known as the *headlight effect*; we see that as $\beta \to \infty$, $\theta \to 0$. This means that the emitted radiation becomes concentrated into a small angular spread around the direction of motion of the emitter. For example, suppose that $\theta_0 = \pi/2$. Then $\theta = \cos^{-1}\beta$, which clearly describes a cone of apex half-angle $\theta$.

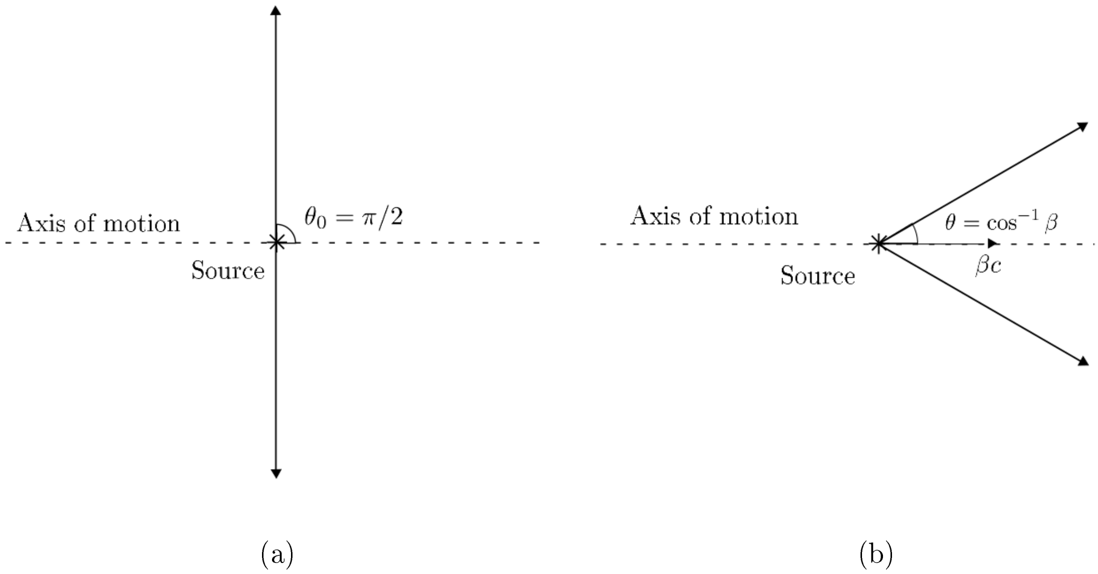



Figure 1.2: The headlight effect (a) The emission of radiation in the rest frame of the source (b) The emission of radiation in the frame of the observer

How does the associated power or intensity transform? For this, we need to consider how the solid angle transforms between both frames.

$$\frac{d\Omega}{d\Omega_0} = \frac{\sin\theta d\theta}{\sin\theta_0 d\theta_0} = \frac{d(\cos\theta)}{d(\cos\theta_0)} = \left(\frac{\omega}{\omega_0}\right)^2 \tag{1.112}$$

where we have made use of the fact that

$$\cos\theta_0 = \frac{\cos\theta - \beta}{1 - \beta\cos\theta} \tag{1.113}$$

Then, if we assume that the area of the wave-front is Lorentz invariant, then the intensity is given by

$$\frac{I}{I_0} = \left(\frac{\omega}{\omega_0}\right)^2 \quad \longrightarrow \quad \frac{dP}{dP_0} = \left(\frac{\omega}{\omega_0}\right)^2 \tag{1.114}$$

Using (1.112), we arrive at the final result that

$$\boxed{\frac{dP}{d\Omega} = \left(\frac{\omega}{\omega_0}\right)^4 \frac{dP_0}{d\Omega_0}} \tag{1.115}$$

There is thus a very strong dependence of the power per unit solid angle on the intensity. In cases where the particle is highly relativistic, this means that the brightness along the direction of motion can be massively magnified, such as in relativistic galactic jets.

## 1.4 Symmetry

A symmetry of a system is a physical or mathematical feature of the system that is preserved or remains changed under some transformation. This transformation may be continuous (such a rotation) or discrete (such a reflection). Symmetries are very important in all branches of physics, as many symmetries gives rise to a cyclic coordinate, and a corresponding conserved quantity. Unfortunately, there is not enough time to properly cover the concept of symmetry, and so we will instead choose to concentrate on a particular symmetry transformation, that of parity inversion.

A *parity inversion* $\mathcal{P}$ is a transformation involving the reversing of the sign of one spatial coordinate. In three dimensions, it describes a simultaneous switch of all the coordinates (a point inversion):

$$\mathcal{P} : \mathbf{x} \mapsto -\mathbf{x} \tag{1.116}$$

We can classify vectors based on the way that they transform under a parity inversion:

- Polar - These reverse their sign under parity inversion. Examples include position $\mathbf{x}$, velocity $\mathbf{v}$, electric field $\mathbf{E}$, electric current density $\mathbf{j}$ and the Del operator $\nabla$
- Axial - These do not reverse their sign under parity inversion, and are often associated with cross-product type quantities. Examples include the magnetic field $\mathbf{B}$, angular momentum $\mathbf{J}$, and magnetic dipole moment $\mathbf{m}$

In most situations, we expect physical theories to be unchanged under parity inversions as, after all, a parity transformation simply corresponds to a change in coordinates. A particularly important example of this that of classical electromagnetism, as encapsulated in Maxwell's equations, and the Lorentz force equation.

$$\begin{aligned}
\nabla \cdot \mathbf{E} &= \frac{\rho}{\epsilon_0} \\
\nabla \cdot \mathbf{B} &= 0 \\
\nabla \times \mathbf{E} &= -\frac{\partial \mathbf{B}}{\partial t} \\
\nabla \times \mathbf{B} &= \mu_0 \mathbf{j} + \mu_0 \epsilon_0 \frac{\partial \mathbf{E}}{\partial t} \\
\mathbf{f} &= q(\mathbf{E} + \mathbf{v} \times \mathbf{B})
\end{aligned} \tag{1.117}$$

For this system of equations, a parity transformation corresponds to $\mathbf{E} \mapsto -\mathbf{E}$, $\nabla \mapsto -\nabla$, $\mathbf{j} \mapsto -\mathbf{j}$, $\mathbf{v} \mapsto -\mathbf{v}$, $\mathbf{f} \mapsto -\mathbf{f}$ and $\mathbf{B} \mapsto \mathbf{B}$. It is clear that the above equations are the same when this transformation is applied. We thus conclude that classical electromagnetism is invariant under parity transformation. This constrains any electromagnetic formalism developed in Special Relativity to obey the same parity invariance, as it must reduce to (1.117). This should be kept in mind when examining the material of chapter 3.

# 2. *Dynamics*

In this chapter, we continue our consideration of relativity in Minkowski Space to include motion under the action of forces, including:

- Energy and Momentum
- Acceleration and Forces
- Angular Momentum

It's a common misconception that Special Relativity cannot handle accelerating objects, or non-inertial reference frames, as it is often understood to only apply to inertial reference frames. This is not true; Special Relativity treats accelerating frames differently from inertial frames, and accelerating objects can actually be dealt with without the use of non-interial frames. We shall thus consider situations in which forces, and the exchange of momentum, is involved.

## 2.1 Energy and Momentum

Let us begin our consideration of dynamics with a treatment of momentum and energy in Special Relativity, as from there we can move to understanding how this changes under the action of forces. Students may already be familiar with much of the material from the following two sections, as this is generally the material that is covered in an introductory course on Special Relativity.

### 2.1.1 Four-momentum

By analogy to classical physics, the four-momentum should be related to the velocity of a particle through an expression of the form

$$\mathsf{P} = m\mathsf{U} = (\gamma_u mc, \gamma_u m\mathbf{u}) \tag{2.1}$$

As this multiplicative constant must be Lorentz invariant, we choose $m$ to represent the rest mass of the particle concerned. Using the familiar expressions for relativistic energy and momentum of

$$E = \gamma mc^2, \quad \mathbf{p} = \gamma m\mathbf{u} \tag{2.2}$$

it is clear that our expression for four-momentum is

$$\boxed{\mathsf{P} = m\mathsf{U} = (E/c, \mathbf{p}), \quad \eta_{\mu\nu}\mathsf{P}^\mu\mathsf{P}^\nu = -m^2c^2} \tag{2.3}$$

We have derived the invariant quantity - as usual - by considering the four-momentum in the rest frame of the particle.

### 2.1.2 The zero-component lemma

Consider some general four-vector $\mathsf{A} = (A^0, A^1, A^2, A^3)$. Suppose that $A^1 = 0$ in all frames. Then, if there are frames in which $A^2$ or $A^3$ are non-zero, then we can perform a rotation of our coordinate system to make $A^1$ non-zero, contrary to the original claim. This means that we are forced to conclude that $A^1 = A^2 = A^3 = 0$. If there is a frame in which $A^0$ is non-zero, then we can apply a Lorentz transformation $\Lambda$ to $\mathsf{A}$ in order to force $A^1$ to be non-zero, which is again contrary to the original claim.

This argument can be repeated for each component of the four-vector, meaning that we have proved the *zero-component lemma*, which states that:

> *If a four-vector has a component which is zero in all frames, then the entire four-vector is zero.*

For four-momentum, we can define the quantity

$$\Delta\mathsf{P} = \mathsf{P}_{\text{final}} - \mathsf{P}_{\text{initial}} \tag{2.4}$$

where 'final' and 'initial' are defined with respect to some spacetime event, and the four-momentum are of the total system. Suppose that we have an isolated system. Then, by definition, energy must be conserved in all frames, meaning that $(\Delta\mathsf{P})^0 = 0$. By the zero-component lemma, this implies that $(\Delta\mathsf{P})^i = 0$ for $i = 1, 2, 3$; that is, the momentum is also conserved. This means that energy conservation implies momentum conservation (and vice-versa), and we can conclude that *the total four-momentum of an isolated system is independent of time (conserved).* In particular, it is not changed by internal interactions among parts of the system.

### 2.1.3 Centre of Momentum Frame

A particularly useful frame when dealing with processes involving the exchange of energy and momentum is the *centre of momentum frame* (CM), which many readers will have already encountered in Classical Mechanics, if not in Special Relativity.

By definition, the CM frame is the frame in which $\mathsf{P}^i = 0$ for $i = 1, 2, 3$. This means that in the CM frame, the total energy of the system is simply the sum of the rest mass energies of the constituents. The CM frame has a few useful properties, which we shall now state and prove.

- The velocity of the CM frame relative to the laboratory frame is

$$\mathbf{v}_{\text{CM}} = \frac{\mathbf{p}_{\text{tot}} c^2}{E_{\text{tot}}} \tag{2.5}$$

  where $\mathbf{p}_{\text{tot}}$ and $E_{\text{tot}}$ are the total momentum and energy in the laboratory frame. Without loss of generality, align $\mathbf{p}_{\text{tot}}$ along $\mathbf{e}_x$ in the laboratory frame. Applying a Lorentz transformation to the four-momentum in the lab frame, we have that

$$p'_{\text{tot}x} = \gamma(-E_{\text{tot}} v/c^2 + p_{\text{tot}x}), \quad p'_{\text{tot}y} = p'_{\text{tot}z} = 0 \tag{2.6}$$

  If the new frame is to be the CM frame, then it is clear that the transformation velocity $\mathbf{v}$ is that given by (2.5)

- If an incoming particle of momentum $\mathbf{p}$ strikes a stationary particle of mass $M$, then the momentum of either particle in the CM frame is given by

$$|\mathbf{p}'| = \frac{Mc^2}{E_{\text{CM}}} |\mathbf{p}| \tag{2.7}$$

  As the particle of mass $M$ is stationary in the lab frame, it has a speed $|\mathbf{v}_{\text{CM}}|$ in the CM frame. This means that

$$|\mathbf{p}'_M| = \gamma M |\mathbf{v}_{\text{CM}}| = \frac{M|\mathbf{p}|c^2}{\sqrt{E_{\text{tot}}^2 - |\mathbf{p}|^2 c^2}} = \frac{Mc^2}{E_{\text{CM}}} |\mathbf{p}| \tag{2.8}$$

  which completes the proof

Note that individual massless particles do not have a CM frame, given that no frame exists in which their momentum is zero. This can actually be used as a powerful tool to argue as to why certain atomic processes are forbidden. For example, an isolated photon cannot transform into an electron-positron pair in free space, as the final state has a well-defined CM frame, while the initial state does not. Similarly, an isolated photon cannot decay into a pair of photos with differing directions of propagation by the same argument.

### 2.1.4 Particle Processes

Armed with the conservation of four-momentum, we can tackle problems with systems involving multiple particles, such as creation, annihilation and scattering processes. For massless particles, we can make use of (1.111) in order to find the way that angles transform when moving between frames.

### Particle Formation

Let us consider the case where a single particle of mass $m$, and momentum $\mathbf{p}$ collides with a stationary particle of mass $M$. The four-momentum in the lab frame is

$$\mathsf{P} = \left(\frac{E}{c} + Mc, \mathbf{p}\right) \tag{2.9}$$

while the four-momentum in the CM frame is given by

$$\mathsf{P}_{\text{CM}} = \left(\frac{E_{CM}}{c}, 0\right) \tag{2.10}$$

$E_{\text{CM}}$ is the energy that is available in the centre of momentum for particle creation, such that $E_{\text{CM}} = \sum_i m_i c^2$ for a series of particles $m_i$. Now, equate the invariants of both frames:

$$\mathsf{P}\cdot\mathsf{P} = \mathsf{P}_{\text{CM}}\cdot\mathsf{P}_{\text{CM}} \quad \longrightarrow \quad E^2 + M^2c^4 + 2MEc^2 - p^2c^2 = E_{\text{CM}}^2 \tag{2.11}$$

Recognising that $E^2 - p^2c^2 = m^2c^4$, this equation can be re-arranged to give

$$\boxed{E_{\text{th}} = \frac{(\sum_i m_i)^2 - m^2 - M^2}{2M}c^2} \tag{2.12}$$

$E_{\text{th}}$ is the threshold energy that the incoming particle must have in order to produce a set of particles $\sum_i m_i$.

### In-flight Decay

Suppose that we have a massive particle with associated four-momentum $\mathsf{P} = (E/c, \mathbf{p})$ that decays into two products of four-momenta

$$\mathsf{P}_1 = \left(\frac{E_1}{c}, \mathbf{p}_1\right), \quad \mathsf{P}_2 = \left(\frac{E_2}{c}, \mathbf{p}_2\right) \tag{2.13}$$

where all quantities are associated with the lab frame. By the conservation and invariance of four-momentum, we have that

$$\mathsf{P}\cdot\mathsf{P} = (\mathsf{P}_1 + \mathsf{P}_2)^2 \quad \longrightarrow \quad M^2 = m_1^2 + m_2^2 - \frac{2}{c^2}\mathsf{P}_1\cdot\mathsf{P}_2 \tag{2.14}$$

Evaluating $\mathsf{P}_1\cdot\mathsf{P}_2$, we arrive at the expression

$$\boxed{M^2 = m_1^2 + m_2^2 + \frac{2}{c^4}(E_1E_2 - |\mathbf{p}_1||\mathbf{p}_2|c^2\cos\theta)} \tag{2.15}$$

where $\theta$ is the angle between the decay products in the lab frame. Given information about $\mathbf{p}_1$ and $\mathbf{p}_2$, one can use invariants to find $\mathbf{E}_1$ and $\mathbf{E}_2$, and thus the mass $M$ of the original particle.

However, suppose that we know the energy $E$ and mass $M$ of the original particle. Then, we can actually find the energies $E_1$ and $E_2$ of the decay products. The CM frame is simply the instantaneous rest frame of the original particle, meaning that the relevant four-momenta are

$$\mathsf{P}_{\text{CM}} = (Mc, 0), \quad \mathsf{P}_1' = \left(\frac{E_1'}{c}, \mathbf{p}_1'\right), \quad \mathsf{P}_2' = \left(\frac{E_2'}{c}, \mathbf{p}_2'\right) \tag{2.16}$$

Let $\mathbf{p}$ be the momentum of either decay product in the CM frame, such that $p = |\mathbf{p}| = |\mathbf{p}_1| = |\mathbf{p}_2|$. By the conservation of energy in the CM, we have that

$$Mc^2 = E_1' + E_2' = \sqrt{m_1^2c^4 + p^2c^2} + \sqrt{m_2^2c^4 + p^2c^2} \tag{2.17}$$

Moving one of the square roots to one side of the equation, squaring both sides, and using the definition of either $E_1'$ or $E_2'$ allows us to obtain

$$E_1' = \frac{M^2 + m_1^2 - m_2^2}{2M}c^2, \quad E_2' = \frac{M^2 + m_2^2 - m_1^2}{2M}c^2 \tag{2.18}$$

Plugging the first of these back into $E_1` = \sqrt{m_1^2c^4 + p^2c^2}$ and re-arranging allows us to obtain the momentum in the CM frame of the decay products as

$$\boxed{p = \frac{c}{2M}\left((m_1^2 + m_2^2 - M^2)^2 - 4m_1^2m_2^2\right)^{1/2}} \tag{2.19}$$

The velocity of the CM is simply the velocity of the original particle, such that $\gamma = E/Mc^2$. Given this, we can perform an inverse Lorentz transform to find the energies $E_1$ and $E_2$ in the lab frame.

*A neutral kaon $K^0$ moving at velocity $\beta_{K^0}$ in the lab frame decays into two pions $\pi^+$ and $\pi^-$ with proper decay times $t_+$ and $t_-$ respectively. If the $\pi^+$ is emitted at an angle $\theta$ to the direction of motion of the $K^0$ in the lab frame, find the angle at which the pions decay simultaneously. Find the lab distance between the decays for this value of $\theta$.*

As the decay products have equal mass, the conservation of momentum requires that they have equal velocities in the instantaneous rest frame of the $K^0$, meaning that they have the same Lorentz factor $\gamma_\pi$ in this frame. In the $K^0$ frame, the decay times are thus $\gamma_\pi t_+$ and $\gamma_\pi t_-$ (time dilated). Consider the decay events in the $K^0$ frame:

$$\mathsf{X}_1' = \begin{pmatrix} c\gamma_\pi t_+ \\ \gamma_\pi\beta_\pi t_+\cos\theta \\ \gamma_\pi\beta_\pi t_+\sin\theta \\ 0 \end{pmatrix}, \quad \mathsf{X}_2' = \begin{pmatrix} c\gamma_\pi t_- \\ -\gamma_\pi\beta_\pi t_-\cos\theta \\ -\gamma_\pi\beta_\pi t_-\sin\theta \\ 0 \end{pmatrix} \tag{2.20}$$

Transforming back to the lab frame yields

$$t_1 = \gamma_{K^0}(\gamma_\pi t_+ + \gamma_\pi\beta_\pi\beta_{K^0}t_+\cos\theta) \tag{2.21}$$

$$t_2 = \gamma_{K^0}(\gamma_\pi t_- - \gamma_\pi\beta_\pi\beta_{K^0}t_-\cos\theta) \tag{2.22}$$

such that

$$t_1 - t_2 = \gamma_{K^0}\gamma_\pi\left[(t_+ - t_-) + \beta_{K^0}\beta_\pi(t_+ + t_-)\cos\theta\right] \tag{2.23}$$

The decays are observed simultaneously in the lab frame for $t_1 = t_2$, meaning that the angle we are looking for is given by

$$\cos\theta = \frac{-(t_+ - t_-)}{\beta_{K^0}\beta_\pi(t_+ + t_-)} \tag{2.24}$$

The distance $d$ between the two decays in the lab frame can be found using invariants. In the lab frame, the time component of $\mathsf{X} = \mathsf{X}_1 - \mathsf{X}_2$ is zero, meaning that its associated invariant is simply the required distance. Defining $\mathsf{X}' = \mathsf{X}_1' - \mathsf{X}_2'$, and equating invariants gives

$$d^2 = \mathsf{X}\cdot\mathsf{X} = \mathsf{X}'\cdot\mathsf{X}' = \gamma_\pi^2\left[c^2(t_+ - t_-)^2 + \beta_\pi^2(t_+ + t_-)^2\right] \tag{2.25}$$

where we have made use of the expression for $\cos\theta$ as above.

### Compton Scattering

Compton scattering is the inelastic scattering of a photon by a charged particle, usually an electron (which is the case we shall now treat). This results in an angle dependent decrease in the energy of the photon. Let the photon have initial energy $E$, and final energy $E'$ when scattering off a stationary electron of mass $m_e$ at an angle $\theta$. The relevant four-momenta are

$$\mathsf{P}_\gamma = \frac{E}{c}(1,1,0), \quad \mathsf{P}_e = m_e c(1,0,0), \quad \mathsf{P}'_\gamma = \frac{E'}{c}(1,\cos\theta,\sin\theta), \quad \mathsf{P}'_e \tag{2.26}$$

We have neglected to specify the form of $P'_e$ as we are not interested in this quantity. The conservation of four-momentum gives us that

$$\mathsf{P}_\gamma + \mathsf{P}_e = \mathsf{P}'_\gamma + \mathsf{P}_e \quad \longrightarrow \quad \mathsf{P}'_e \cdot \mathsf{P}'_e = \left(\mathsf{P}_\gamma + \mathsf{P}_e - \mathsf{P}'_\gamma\right)^2 \tag{2.27}$$

We have chosen to isolate $\mathsf{P}'_e$ as this allows us to simply get rid of is using the invariant $\mathsf{P}'_e \cdot \mathsf{P}'_e = -m_e^2 c^2$. Evaluating the right-hand side of this expression using the fact that photon four-momenta are null gives the final energy of the photon as

$$E' = \frac{E m_e c^2}{m_e c^2 + E(1-\cos\theta)} \tag{2.28}$$

This is usually expressed in terms of wavelengths $\lambda$ and $\lambda'$ of the initial and final states of the photon:

$$\boxed{\lambda' = \lambda + \lambda_C(1-\cos\theta), \quad \lambda_C = \frac{h}{m_e c}} \tag{2.29}$$

The quantity $\lambda_C$ is known as the *Compton wavelength*. Note that if the photon is of low, but sufficient, energy, it can eject an electron from the atom on which it is incident, instead of undergoing Compton scattering. This is why the intensity spectrum (in $\lambda$) will feature two peaks; one around $\lambda'$, and another at the wavelength corresponding to nearly free electrons in the atom.

Inverse Compton scattering is the process whereby a photon scatters off a moving particle, allowing it to actually gain energy. Of course, this is simply normal Compton scattering, but viewed from a different frame. From (2.27), we have that

$$\mathsf{P}_\gamma \cdot \mathsf{P}'_\gamma = \mathsf{P}_e \cdot (\mathsf{P}_\gamma - \mathsf{P}'_\gamma) \tag{2.30}$$

for some new $\mathsf{P}_e$, where we have once again used the null condition for the photons. This result is quite general, but for the sake of an example, consider the case of a head-on collision. Then

$$\mathsf{P}_\gamma = \frac{E}{c}(1,1,0), \quad \mathsf{P}_e = \gamma m_e(c^2,-u/c,0), \quad \mathsf{P}'_\gamma = \frac{E'}{c}(1,-1,0), \quad \mathsf{P}'_e \tag{2.31}$$

such that

$$E'_1 = \frac{\gamma m_e c^2(1+u/c)}{2+\gamma m_e c^2(1-u/c)/E_1} \tag{2.32}$$

For ultra-relativistic electrons, we have that $u/c \approx 1$, meaning that we can write $(1+u/c) \approx 2$, and $(1-u/c) \approx 1/2\gamma^2$, so

$$E'_1 \approx \frac{\gamma m_e c^2}{1 + m_e c^2/4\gamma E_1} \tag{2.33}$$

Inverse Compton scattering of this kind is dominant in astrophysical processes, as it is very rare to find stationary electrons. An example of such a process is the interaction of photons with electrons in the ultra-relativistic jets of active galactic nuclei.

## 2.2 Acceleration and Forces

In previous sections, we have introduced both the four-velocity $\mathsf{U}$ and the four-momentum $\mathsf{P}$. The natural continuation is to consider the rate of change of these quantities, namely the four-acceleration $\dot{\mathsf{U}}$ and four-force $\mathsf{F}$ respectively.

### 2.2.1 Acceleration

The four-acceleration is defined as one would expect, except that the relationship to the three-vector acceleration $\mathbf{a}$ is somewhat more complicated:

$$\dot{\mathsf{U}} = \frac{d\mathsf{U}}{d\tau} = \gamma\frac{d\mathsf{U}}{dt} = \gamma\left(\frac{d\gamma}{dt}c, \frac{d\gamma}{dt}\mathbf{u} + \gamma\mathbf{a}\right) \tag{2.34}$$

We now use the identity (1.89) with the chain rule, such that

$$\frac{d\gamma}{dt} = \frac{d\gamma}{d\mathbf{u}}\cdot\frac{d\mathbf{u}}{dt} = \gamma^3\frac{\mathbf{u}\cdot\mathbf{a}}{c^2} \tag{2.35}$$

We can then write the four-acceleration as

$$\boxed{\dot{\mathsf{U}} = \gamma^2\left(\frac{\mathbf{u}\cdot\mathbf{a}}{c}\gamma^2, \frac{\mathbf{u}\cdot\mathbf{a}}{c^2}\gamma^2\mathbf{u} + \mathbf{a}\right), \quad \eta_{\mu\nu}\dot{\mathsf{U}}^\mu\dot{\mathsf{U}}^\nu = a_0^2} \tag{2.36}$$

where $a_0 = |\mathbf{a}_0|$ is the *proper acceleration*, which is the acceleration as measured in the instantaneous rest frame of the particle. Taking the scalar product with the four-velocity $\mathsf{U}$, it is clear that $\mathsf{U}\cdot\dot{\mathsf{U}} = 0$; that is, the four-acceleration is orthogonal to the four-velocity. This also follows from the fact that $\dot{\mathsf{U}}$ is spacelike, while $\mathsf{U}$ is timelike, as per the results of section 1.3.1.

We can also find a relationship between the acceleration $\mathbf{a}$ observed in any reference frame to the proper acceleration $\mathbf{a}_0$ by equating invariants:

$$\gamma^4\left[-\left(\frac{\mathbf{u}\cdot\mathbf{a}}{c}\right)^2\gamma^4 + \left(\frac{\mathbf{u}\cdot\mathbf{a}}{c^2}\gamma^2\mathbf{u} + \mathbf{a}\right)^2\right] = \gamma^4 a^2 + \gamma^6(\mathbf{u}\cdot\mathbf{a})^2/c^2 = a_0^2 \tag{2.37}$$

Noting the vector identity

$$(\mathbf{u}\cdot\mathbf{a})^2 = u^2a^2 - (\mathbf{u}\times\mathbf{a})^2 \tag{2.38}$$

it follows that we can write

$$\boxed{a_0^2 = \gamma^6\left(a^2 - \frac{(\mathbf{u}\times\mathbf{a})^2}{c^2}\right)} \tag{2.39}$$

We have chosen to write this expression in this form, as the relationships for the special cases of the velocity and acceleration being parallel or perpendicular are immediately obvious.

### 2.2.2 Forces

By analogy to Classical physics, one would be tempted to adopt $\mathsf{F} = m\dot{\mathsf{U}}$ as the definition of four-force. However, this includes the implicit assumption that the rest mass of the system is constant, which is not necessarily the case. We thus adopt the definition

$$\boxed{\mathsf{F} = \frac{d\mathsf{P}}{d\tau} = \gamma\left(\frac{1}{c}\frac{dE}{dt}, \mathbf{f}\right)} \tag{2.40}$$

where $E$ is the total energy and $\mathbf{f} = d\mathbf{p}/dt$ is the three-force exerted on the system.

### Transformations of Forces

Consider two frames $S$ and $S'$ in standard configuration, with relative velocity $\mathbf{v}$. The four-momenta of these two frames are related by

$$\mathsf{P}' = \Lambda \mathsf{P} \tag{2.41}$$

Now, differentiate both sides of this equation with respect to $t'$, the frame time in $S'$. Noting that $\mathsf{F} = d\mathsf{P}/d\tau$ and $\mathsf{F} = d\mathsf{P}'/d\tau$,

$$\frac{dP'}{dt'} = \Lambda \frac{dP}{dt'} \quad \longrightarrow \quad \frac{1}{\gamma_{u'}}\mathsf{F}' = \Lambda \frac{1}{\gamma_u}\mathsf{F}\frac{dt}{dt'} \tag{2.42}$$

We can obtain an expression for $dt/dt'$ by letting $\mathbf{x} = \mathbf{u}t$ in (1.52):

$$ct' = \gamma_v(ct - \boldsymbol{\beta}\cdot\mathbf{x}) \quad \longrightarrow \quad \frac{dt'}{dt} = \gamma_v(1 - \mathbf{u}\cdot\mathbf{v}/c^2) \tag{2.43}$$

Substituting (2.43) into (2.42), we can write explicitly that

$$\left(\frac{1}{c}\frac{dE'}{dt'}, \mathbf{f}'\right) = \frac{\Lambda}{\gamma_v(1-\mathbf{u}\cdot\mathbf{v}/c^2)}\left(\frac{1}{c}\frac{dE}{dt}, \mathbf{f}\right) \tag{2.44}$$

Decomposing the force into parallel and perpendicular components $\mathbf{f} = \mathbf{f}_\parallel + \mathbf{f}_\perp$ with respect to the relative velocity $\mathbf{v}$ between the frames, we arrive at the final transformation equations:

$$\boxed{\mathbf{f}'_\parallel = \frac{\mathbf{f}_\parallel - (\mathbf{v}/c^2)dE/dt}{1-\mathbf{u}\cdot\mathbf{v}/c^2}, \quad \mathbf{f}'_\perp = \frac{\mathbf{f}_\perp}{\gamma_v(1-\mathbf{u}\cdot\mathbf{v}/c^2)}} \tag{2.45}$$

### Pure Forces

Let us now examine the conditions under which the rest mass energy of the system is conserved. Consider the invariant $\mathsf{U}\cdot\mathsf{F}$ :

$$\mathsf{U}\cdot\mathsf{F} = -\gamma^2\left(\frac{dE}{dt} - \mathbf{f}\cdot\mathbf{u}\right) = \underbrace{\mathsf{U}_0\cdot\mathsf{F}_0}_{\text{Rest frame quantites}} = -\frac{dE_0}{dt} \tag{2.46}$$

However, $E_0$ is simply the rest mass energy in the rest frame:

$$\frac{dE_0}{dt} = \frac{d}{dt}(\gamma mc^2) = \gamma c^2 \frac{dm}{dt} \tag{2.47}$$

We have been able to ignore the time derivative of $\gamma$ given that we are in the rest frame of the system, meaning that the Lorentz factor will not change regardless of whether the frame is accelerating or not. This means that the condition for the rest mass of the system to be conserved is

$$\boxed{\frac{dE}{dt} = \mathbf{f}\cdot\mathbf{u} \quad \longleftrightarrow \quad \mathsf{U}\cdot\mathsf{F} = 0} \tag{2.48}$$

Forces $\mathbf{f}$ that satisfy the above condition are known as *pure forces*, such as the Lorentz force. Forces that are not pure forces include the weak and strong forces.

How is the force related to the acceleration, given that it satisfies (2.48)? Using the relationship between $\mathbf{f}$ and $\mathbf{p}$:

$$\mathbf{f} = \frac{d\mathbf{p}}{dt} = \frac{d}{dt}(\gamma m\mathbf{u}) = \gamma m\mathbf{a} + m\mathbf{u}\frac{d\gamma}{dt} \tag{2.49}$$

Then, we note that

$$\frac{dE}{dt} = mc^2\frac{d\gamma}{dt} + \cancel{\gamma c^2\frac{dm}{dt}} = \mathbf{f}\cdot\mathbf{u} \tag{2.50}$$

As we are considering a pure force. Note that $d\gamma/dt \neq 0$ in this case, as we are no longer in the rest frame of the system. Combining (2.49) and (2.50), it follows that

$$\boxed{\mathbf{f} = \gamma m\mathbf{a} + \frac{\mathbf{f}\cdot\mathbf{u}}{c^2}\mathbf{u}} \tag{2.51}$$

This can be split into components parallel and perpendicular to the velocity $\mathbf{u}$ (rather than the transformation velocity $\mathbf{v}$ as before)

$$f_{\parallel} = \gamma^3 ma_{\parallel}, \quad f_{\perp} = \gamma ma_{\perp} \tag{2.52}$$

An important consequence of these equations is that there is a greater inertial resistance to velocity changes (whether an increase or deacrese) along the direction of motion, compared to the inertial resistance to picking up a velocity component transverse to the current motion.

### 2.2.3 Hyperbolic Motion

Hyperbolic motion is the motion of a particle in Special Relativity when it is subject to constant proper acceleration. It is so called because the wordline of such a particle traces a hyperbola in a spacetime diagram, as we shall see. For reference, let us state some hyperbolic trigonometric identities here:

$$\begin{array}{ll}
\cosh^2\rho - \sinh^2\rho = 1 & \frac{d}{d\rho}(\sinh\rho) = \cosh\rho \\
\tanh^2\rho + \operatorname{sech}^2\rho = 1 & \frac{d}{d\rho}(\tanh\rho) = \operatorname{sech}^2\rho \\
\coth^2\rho - \operatorname{csch}^2\rho = 1 & \frac{d}{d\rho}(\coth\rho) = -\operatorname{csch}^2\rho
\end{array}$$

#### Constant Acceleration

Suppose that the (constant) acceleration is directed in the same direction as the motion of the particle, such that by (2.39) the proper acceleration is given by

$$a_0 = \gamma^3 a \tag{2.53}$$

where $a$ is the acceleration in the inertial frame of the observer. As the other frame being considered is the instantaneous rest frame of the particle, we have that $\mathbf{u} = \mathbf{v}$, such that $a = dv/dt$. We can thus write that

$$\int dv\,\gamma^3 = \int dv\,\frac{d}{dv}(\gamma v) = \int dt\,a_0 \quad \longrightarrow \quad v = \frac{a_0 t}{\sqrt{1+(a_0 t/c)^2}} \tag{2.54}$$

We now make the substitution that

$$\rho = \sinh^{-1}\left(\frac{a_0 t}{c}\right) \tag{2.55}$$

where $t$ is the time as measured in the inertial frame of the observer. The parameter $\rho$ is known as the *rapidity*. From this definition, it immediately follows that

$$\beta = \tanh\rho, \quad \gamma = \cosh\rho \tag{2.56}$$

The former of these is usually taken as the definition of rapidity. We can use this to find an explicit form for $\rho$.

$$\frac{d\beta}{d\tau} = \frac{d\rho}{d\tau}\frac{d}{d\rho}(\tanh\rho) = \frac{1}{\cosh^2\rho}\frac{d\rho}{d\tau} \tag{2.57}$$

However, we also have that

$$\frac{d\beta}{d\tau} = \gamma\frac{d\beta}{dt} = \frac{\gamma a}{c} = \frac{a_0}{\gamma^2 c} \tag{2.58}$$

Equating (2.57) and (2.58), it is clear that

$$\boxed{\frac{d\rho}{d\tau} = \frac{a_0}{c} \quad\longleftrightarrow\quad \rho = \frac{a_0\tau}{c}} \tag{2.59}$$

Rapidity can thus be thought of as the proper time (in units of $c/a_0$) measured from the event where $v = 0$.

We can now time the time and position in the observer frame as a function of $\tau$, and hence the position along the worldline. From the metric $\eta_{\mu\nu}$, we have that

$$dt = \gamma d\tau \quad\longrightarrow\quad t = \int d\tau\,\cosh\left(\frac{a_0\tau}{c}\right) = \frac{c}{a_0}\sinh\left(\frac{a_0\tau}{c}\right) \tag{2.60}$$

where we have used the definition of $\gamma$ as per (2.56). Similarly, we have that

$$x = \int dt\,v = \int d\tau\,\gamma v = c\int d\tau\,\sinh\left(\frac{a_0\tau}{c}\right) = \frac{c^2}{a_0}\left[\cosh\left(\frac{a_0\tau}{c}\right) - 1\right] \tag{2.61}$$

From these two equations, it is very clear that the motion of the particle satisfies

$$\left(1 + \frac{a_0 x}{c^2}\right)^2 - \left(\frac{a_0 t}{c}\right)^2 = 1 \tag{2.62}$$

which is clearly the equation of a hyperbola. This can also conveniently be written as

$$\mathsf{X} = \frac{c^2}{a_0}(\sinh\rho, \cosh\rho) \tag{2.63}$$

where the other two coordinates have been suppressed, as they are assumed zero on this problem. From this expression, it becomes very simple to show that the acceleration is constant. Differentiating the invariant in (2.36), we have that

$$\frac{d}{d\tau}(a_0)^2 = \frac{d}{d\tau}(\dot{\mathsf{U}}_\mu\dot{\mathsf{U}}^\mu) = 2\dot{\mathsf{U}}_\mu\frac{d\dot{\mathsf{U}}^\mu}{d\tau} \propto \dot{\mathsf{U}}_\mu\mathsf{U}^\mu = 0 \tag{2.64}$$

**Rapidity**

Let us take a further look at the concept of rapidity. We begin with a re-statement of all of the relations involving rapidity that were derived in the previous section.

$$\rho \equiv \tanh^{-1}\beta = (a_0\tau)/c \tag{2.65}$$
$$\gamma = \cosh\rho \tag{2.66}$$
$$t = (c/a_0)\sinh(a_0\tau/c) \tag{2.67}$$
$$x = (c^2/a_0)\left[\cosh(a_0\tau/c) - 1\right] \tag{2.68}$$

Comparison of these equations with the form of (1.54), it is clear that we can also write the Lorentz transformation as

$$\Lambda = \begin{pmatrix} \cosh\rho & -\sinh\rho & 0 & 0 \\ -\sinh\rho & \cosh\rho & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{2.69}$$

Seen in this form, the Lorentz transform can be though of as a rotation by an imaginary angle $i\rho$, analogous to rotations in three-dimensional space. This is the special case of a more general result associated with the group of proper Lorentz transformations, but more on this in chapter 4.

Consider velocities $v$, $u$ and $w$ all orientated along the same axis, with associated rapidities $\rho_v$, $\rho_u$ and $\rho_w$. Using the definition (2.65), it follows from (1.84) that

$$\tanh\rho_w = \frac{\tanh\rho_v + \tanh\rho_u}{1 + \tanh\rho_v \tanh\rho_u} = \tanh(\rho_v + \rho_u) \tag{2.70}$$

This leads us to conclude that rapidities are additive. This corresponds to $i\rho$ representing a rotation; performing successive rotations (and thus boosting into other frames) will have a cumulative effect. This can also be a useful tool to determine the passage of proper time in different frames.

*The Twin Paradox problem: A travelling twin leaves Earth on board a spaceship undergoing motion at constant proper acceleration $a_0$. After $\tau_0$ years of proper time for the twin in the spaceship, the direction of the rockets are (instantly) reversed so that the spaceship accelerates towards the Earth for $2\tau_0$ proper years. The rockets are then again reversed to allow the spaceship to slow and come to rest on Earth after a further $\tau_0$ years of spaceship proper time. How much do the two twins ages?*

The answer to the question is trivial in the case of the travelling twin; he simply ages according to the passage of time in the rocket frame, meaning he ages by $4\tau_0$ years. The inertial observer time $t$ on Earth accumulated during each acceleration phase of the rocket is given by (2.67):

$$t = \frac{c}{a_0}\sinh\left(\frac{a_0\tau_0}{c}\right) \tag{2.71}$$

As there are four phases of acceleration, the total time measured by the twin on Earth is

$$\Delta t = \frac{4c}{a_0}\sinh\left(\frac{a_0\tau_0}{c}\right) \tag{2.72}$$

Suppose that $a_0 = 9.8\,\mathrm{ms}^{-1} = 1.03\,\mathrm{lightyears}\,(\mathrm{years})^{-2}$ and $\tau_0 = 5\,\mathrm{years}$. We find that the travelling twin ages 20 years, while the twin that remains on Earth ages $\approx 370$ years. Such is Minkowski spacetime.

### Non-constant acceleration

Even in cases where the acceleration is not constant (for which the resultant worldline is not hyperbolic), rephrasing the relevant problem in terms of rapidity can be a useful route to a solution. To illustrate this, let us consider an example.

*A spaceship emits fuel at a relativistic speed $u$ relative to its instantaneous rest frame. Calculate the speed of the spaceship as measured in an inertial frame that is stationary with* *respect to the initial position of the rocket when a fraction $\alpha$ of its mass remains.*

In the instantaneous rest frame, consider momentum and energy conservation when the spaceship loses a small amount of mass $dm$. Let the energy of the ejected fuel be $E_u$, with corresponding momentum $p_u$. Writing the new mass of the rocket as $(m - dm)$, conservation of energy gives

$$mc^2 = E_u + \gamma_{dv}(m - dm)c^2 \tag{2.73}$$

where $\gamma_{dv}$ is the Lorentz factor associated with the small velocity change $dv$. Similarly, momentum conservation gives

$$0 = -p_u + \gamma_{dv}(m - dm)dv \tag{2.74}$$

We can solve these two equations by remarking that $p_u = (u/c^2)E_u$, such that

$$0 = -mu + \gamma_{dv}u(m - dm) + \gamma_{dv}mdv \tag{2.75}$$

where we have ignored the quadratic term $dvdm$. Now, to first order, $\gamma_{dv} \approx 1$ for small velocity changes $dv$, such that we can write that

$$(-dm)u = mdv \tag{2.76}$$

It is now convenient to introduce rapidity into the calculation. We need to calculate the change of variables in the local inertial frame (for which $v$ was initially zero), namely

$$\left.\frac{d\beta}{d\rho}\right|_{\beta=0} = \operatorname{sech}^2\beta\big|_{\beta=0} = 1 \tag{2.77}$$

Using this relationship in (2.76), we find that

$$-\frac{u}{c}\frac{dm}{m} = d\rho \quad \longrightarrow \quad \rho = -\frac{u}{c}\log\alpha \tag{2.78}$$

upon integration. Then, using the definition (2.65), we can express the inverse hyperbolic tangent as

$$\rho = \tanh^{-1}(v/c) = \frac{1}{2}\log\frac{c+v}{c-v} \tag{2.79}$$

which can be equated with (2.78) to give

$$\alpha = \left(\frac{c-v}{c+v}\right)^{c/2u} \tag{2.80}$$

This can then easily be re-arranged to obtain the speed $v$ as a function of $\alpha$.

## 2.3 Angular Momentum

The last dynamical quantity that we will consider is that of angular momentum. This will also be an opportunity for readers to be introduced to their first explicit tensor quantity in these notes, ahead of tackling the tensor quantities that are contained in the study of covariant electromagnetism.

### 2.3.1 Orbital Angular Momentum

Recall the definition of orbital angular momentum $\mathbf{L} = \mathbf{x} \times \mathbf{p}$, where $\mathbf{x}$ and $\mathbf{p}$ are the (three) position and momentum respectively. By analogy, we would expect angular momentum in Special Relativity to be made out of a cross-product like quantity. As such, we define the orbital angular momentum tensor as

$$\boxed{\mathbb{L}^{\mu\nu} = \mathsf{X}^\mu \mathsf{P}^\nu - \mathsf{X}^\nu \mathsf{P}^\mu} \tag{2.81}$$

where $\mathsf{X}$ and $\mathsf{P}$ are the four position and four momenta of the particle in question. In terms of components, this is given by

$$\mathbb{L}^{00} = 0, \quad \mathbb{L}^{0i} = -\mathbb{L}^{i0} = -w^i/c = -x^i E/c + p^i ct, \quad \mathbb{L}^{ij} = \epsilon_{ijk} L_k \tag{2.82}$$

We can write (2.81) explicitly as

$$\mathbb{L}^{\mu\nu} = \left( \begin{array}{c|ccc} 0 & & -\mathbf{w}/c & \\ \hline & 0 & L_z & -L_y \\ \mathbf{w}/c & -L_z & 0 & L_x \\ & L_y & -L_x & 0 \end{array} \right) \tag{2.83}$$

In both Special and General Relativity, tensors representing will often take this form, with the time-time, space-time, and space-space parts of the tensor (separated by the horizontal and vertical lines) representing different physical quantities. In this case, the space-space part clearly represents the three-angular momentum, while the space-time/time-space parts include some 'work-like' quantity $\mathbf{w}$.

#### Conservation of Orbital Angular Momentum

Let us consider the rate of change of orbital angular momentum. Differentiating (2.81) with respect to proper time:

$$\frac{d\mathbb{L}^{\mu\nu}}{d\tau} = \mathsf{U}^\mu \mathsf{P}^\nu - \mathsf{U}^\nu \mathsf{P}^\mu + \mathsf{X}^\mu \mathsf{F}^\nu - \mathsf{X}^\nu \mathsf{F}^\mu = \mathsf{X}^\mu \mathsf{F}^\nu - \mathsf{X}^\nu \mathsf{F}^\mu \tag{2.84}$$

where we have made us of the definition of $\mathsf{U}^\mu$ and recalled that $\mathsf{P}^\mu = m\mathsf{U}^\mu$. It is thus clear that the angular momentum tensor is conserved in the absence of any external torques ($\mathsf{F} = 0$). This is encouraging; we are able to recover classical results from this seemingly unfamiliar tensor. This sort of exercise is often a useful way of checking whether quantities that we encounter in Relativity are indeed well behaved.

#### Motion about a Pivot

(2.81) defines the angular momentum of a single particle. For composite bodies, we define the total orbital angular momentum by

$$\mathbb{L}^{\mu\nu} = \sum_i \mathbb{L}^{\mu\nu}_{(i)}, \quad \mathsf{P}^\mu = \sum_{(i)} \mathsf{P}^\mu_{(i)} \tag{2.85}$$

where the subscript $(i)$ is a summation index, rather than a spatial component index. By linearity, the same argument as in (2.84) holds for the total orbital angular momentum, meaning that

$$\mathbf{w} = \sum_{(i)} \mathbf{x}_{(i)} E_{(i)} - c^2 t \sum_{(i)} \mathbf{p}_{(i)} \tag{2.86}$$

is constant in time. Defining the position of the *centroid* (sometimes known as the 'centre of energy') as

$$\mathbf{x}_C = \frac{\sum_{(i)} \mathbf{x}_{(i)} E_{(i)}}{\sum_{(i)} E_{(i)}} \tag{2.87}$$

it follows quickly from (2.5) that

$$\frac{d\mathbf{x}_C}{dt} = \frac{\mathbf{p}_{\text{tot}} c^2}{E_{\text{tot}}} = \mathbf{v}_{\text{CM}} \tag{2.88}$$

This shows that our treatment of composite objects as single entities will well-defined properties is indeed valid. Now, suppose that we have a pivot located at $\mathsf{R}$. Then:

$$\mathbb{L}^{\mu\nu}(\mathsf{R}) = \sum_{(i)} \left[ \left( \mathsf{X}^{\mu}_{(i)} - \mathsf{R}^{\mu} \right) \mathsf{P}^{\nu}_{(i)} - \left( \mathsf{X}^{\nu}_{(i)} - \mathsf{R}^{\nu} \right) \mathsf{P}^{\mu}_{(i)} \right] = \mathbb{L}^{\mu\nu}(0) - \mathsf{R}^{\mu}\mathsf{P}^{\nu} + \mathsf{R}^{\nu}\mathsf{P}^{\mu} \tag{2.89}$$

In the CM frame, the spatial part of $\mathsf{P}$ is zero, meaning that the spatial part of $\mathbb{L}^{\mu\nu}(\mathsf{R})$ is equal to that of $\mathbb{L}^{\mu\nu}(0)$. That is, the three-angular momentum in the CM frame is independent of the position of the pivot.

### 2.3.2 Spin Angular Momentum

In the case of a point particle at $\mathsf{R}$, we can recognise the second term on the right-hand side of (2.89) as the angular momentum about the origin. Letting $\mathsf{R} = \mathsf{X}_C$, we then define

$$\mathbb{S}^{\mu\nu} = \mathbb{L}^{\mu\nu}(\mathsf{X}_C), \quad \mathbb{L}^{\mu\nu}_C = \mathsf{X}^{\mu}_C \mathsf{P}^{\nu} - \mathsf{X}^{\nu}_C \mathsf{P}^{\mu}, \quad \mathbb{J}^{\mu\nu} = \mathbb{S}^{\mu\nu} + \mathbb{L}^{\mu\nu}_C \tag{2.90}$$

where $\mathbb{J}^{\mu\nu}$ is the total angular momentum about the origin, and $\mathsf{S}^{\mu\nu}$ is the spin angular momentum about the centroid. From this definition, it is clear that $\mathsf{S}^{\mu\nu}$ is antisymmetric, and has zero space-time part. Its spatial part contains the components of the spin angular momentum $\mathbf{s}$.

#### Pauli-Lubanski Spin four-vector

The *Pauli-Lukbanski spin four-vector* is defined as

$$\mathsf{W}_{\rho} = \frac{1}{2}\epsilon_{\sigma\rho\mu\nu}\mathsf{P}^{\sigma}\mathbb{J}^{\mu\nu} = \frac{1}{2}\epsilon_{\sigma\rho\mu\nu}\mathsf{P}^{\sigma}\mathbb{S}^{\mu\nu} \tag{2.91}$$

The second expression follows from the fact that $\epsilon_{\sigma\rho\mu\nu}\mathsf{P}^{\sigma}\mathbb{L}^{\mu\nu}_C = 0$. Component-wise, this is written as

$$\boxed{\mathsf{W} = \left( \mathbf{s}\cdot\mathbf{p}, \frac{E}{c}\mathbf{s} \right), \quad \eta_{\mu\nu} W^{\mu} W^{\nu} = m^2 c^2 s_0^2} \tag{2.92}$$

where $s_0 = |\mathbf{s}_0|$ is the three-spin of the particle in its instantaneous rest frame. It is easy to show that $\mathsf{W}\cdot\mathsf{P} = 0$. Using an appropriate Lorentz transformation, it is also possible to show that

$$\mathbf{s}_{\parallel} = \mathbf{s}_{0\parallel}, \quad \mathbf{s}_{\perp} = \frac{\mathbf{s}_{0\perp}}{\gamma} \tag{2.93}$$

where the parallel and perpendicular directions are defined with respect to the particle velocity. It is clear that as $v \to c$, the spin becomes aligned with the particle velocity, such that $\mathsf{W} \propto \mathsf{P}$. However, we still retain the relationship that $\eta_{\mu\nu}W^\mu P^\nu = 0$, as for $v = c$, the four-momentum of the particle must be null. For massive particles, we define the spin four-vector

$$\mathsf{S} = \mathsf{W}/mc = \left(\frac{\gamma}{c}\mathbf{s}\cdot\mathbf{v}, \gamma\mathbf{s}\right) \tag{2.94}$$

In the absence of external torques, this evolves according to

$$\frac{d\mathsf{S}^\mu}{d\tau} = \frac{\mathsf{S}_\nu \dot{\mathsf{U}}^\nu}{c^2}\mathsf{U}^\mu \tag{2.95}$$

where the dot indicates differentiation with respect to the proper time $\tau$. Making use of (2.35), it follows that

$$\dot{\mathsf{U}} = \left(\frac{\gamma^3}{c}\mathbf{v}\cdot\dot{\mathbf{v}}, \gamma\dot{\mathbf{v}} + \frac{\gamma^3}{c^2}\mathbf{v}(\mathbf{v}\cdot\dot{\mathbf{v}})\right) \quad \longrightarrow \quad \mathsf{S}_\nu\dot{\mathsf{U}}^\nu = \gamma^2\mathbf{s}\cdot\dot{\mathbf{v}} \tag{2.96}$$

such that our evolution equation becomes

$$\frac{d\mathsf{S}^\mu}{d\tau} = \frac{\gamma^2}{c^2}(\mathbf{s}\cdot\dot{\mathbf{v}})\mathsf{U}^\mu \tag{2.97}$$

We can now find the time evolution of the three-spin $\mathbf{s}$. Letting $\mathsf{S}^i = \gamma s^i$ and $\mathsf{U}^i = \gamma v^i$:

$$\frac{d\mathbf{s}}{d\tau} = \frac{\gamma^2}{c^2}\left[(\mathbf{s}\cdot\dot{\mathbf{v}})\mathbf{v} - (\mathbf{v}\cdot\dot{\mathbf{v}})\mathbf{s}\right] \tag{2.98}$$

Now, suppose that the velocity of the particle in question is described by the expression $\mathbf{v}(\tau) = c\hat{\mathbf{v}}[1-\exp(-2\Gamma\tau)]^{1/2}$ for some constant $\Gamma$, such that $v \to c$ as $\tau \to \infty$. Letting $\mathbf{s} = \mathbf{s}_\parallel + \mathbf{s}_\perp$, and remarking that $\gamma^2 e^{-2\Gamma\tau} = 1$, it is simple to show that

$$\frac{d\mathbf{s}_\parallel}{d\tau} = 0, \quad \frac{d\mathbf{s}_\perp}{d\tau} = -\gamma\mathbf{s}_\perp \tag{2.99}$$

meaning that the three-spin evolves with the proper time as

$$\mathbf{s}(\tau) = \mathbf{s}_\parallel(0) + \mathbf{s}_\perp(0)e^{-\Gamma\tau} \tag{2.100}$$

Thus, we recover explicitly the prediction that as $v \to c$, the three-spin becomes aligned with the velocity of the particle.

### Thomas Precession

In the previous section, we stated that the spin four-vector $\mathsf{S}$ evolves according to (2.95) in the absence of external torques. What does this condition actually mean for the three-spin in the instantaneous rest frame of the particle $\mathbf{s}_0$? Dotting (2.95) with $\mathsf{S}$, it follows that

$$\mathsf{S}\cdot\frac{d\mathsf{S}}{d\tau} = \frac{1}{2}\frac{d}{d\tau}(\mathsf{S}\cdot\mathsf{S}) = \frac{1}{2}\frac{d}{d\tau}(s_0^2) = 0 \tag{2.101}$$

where we have made use of the fact that $\mathsf{W}\cdot\mathsf{P} \propto \mathsf{W}\cdot\mathsf{U} = 0$. This means that the proper three-spin is of fixed magnitude $s_0$ during its motion. If the direction of motion of a particle with such a spin is changing with respect to some inertial frame, the sequence of Lorentz transformations that are required to transform the spin vector into the inertial frame causes said spin to precess in the inertial frame. The apparent precession of an accelerated particle that has spin is known as *Thomas Precession.*

Let $\theta_0$ be the angle in the instantaneous rest frame of the particle between $\mathbf{s}_0$ and the particle's velocity $\mathbf{v}$. As this is the relative velocity between the particle's frame, and the inertial lab frame, the angle $\theta_0$ is well-defined in both frames, and they agree on its angle relative to their respective coordinate axes (as the Lorentz transformation relating them is along $\mathbf{v}$). Given the relationships in (2.93), we can write the four-spin $\mathsf{S}$ in the lab frame as

$$\mathsf{S} = \left(\gamma(v/c)s_0\cos\theta_0, \gamma s_0\cos\theta_0\mathbf{e}_\parallel + s_0\sin\theta_0\mathbf{e}_\perp\right) \tag{2.102}$$

where once again $\mathbf{e}_\parallel$ and $\mathbf{e}_\perp$ are defined with respect to the velocity direction $\mathbf{v}$. To find the evolution of $\theta_0$, it is convenient to express $\mathsf{S}$ in terms of the four-vectors $\mathsf{M}$ and $\mathsf{N}$:

$$\mathsf{M} = (\gamma v/c, \gamma\mathbf{e}_\parallel), \quad \mathsf{N} = (0, \mathbf{e}_\perp) \tag{2.103}$$

such that

$$\mathsf{S} = s_0\left[\mathsf{M}\cos\theta_0 + \mathsf{N}\sin\theta_0\right] \tag{2.104}$$

Noticing that $\mathsf{M}$ is a Lorentz boosted version of $(0, \mathbf{e}_\parallel)$, and evaluating in an appropriate frame, it is easy to demonstrate that

$$\mathsf{M}\cdot\mathsf{M} = \mathsf{N}\cdot\mathsf{N} = 1, \quad \mathsf{M}\cdot\dot{\mathsf{M}} = \mathsf{N}\cdot\dot{\mathsf{N}} = 0, \quad \mathsf{M}\cdot\mathsf{U} = \mathsf{N}\cdot\mathsf{U} = \mathsf{M}\cdot\mathsf{N} = 0 \tag{2.105}$$

where again the dots indicate differentiation with respect to the proper time. Then, the time derivative of $\mathbf{S}$ is given by

$$\dot{\mathsf{S}} = s_0\left[\dot{\mathsf{M}}\cos\theta_0 + \dot{\mathsf{N}}\sin\theta_0 + \dot{\theta}_0\left(-\mathsf{M}\sin\theta_0 + \mathsf{N}\cos\theta_0\right)\right] \tag{2.106}$$

Substituting this into (2.95), and dotting both sides with $\mathsf{N}$ gives

$$\mathsf{N}\cdot\dot{\mathsf{M}}\cos\theta_0 + \dot{\theta}_0\cos\theta_0 = 0 \tag{2.107}$$

Thus, we have that

$$\frac{d\theta_0}{d\tau} = -\mathsf{N}\cdot\frac{d\mathsf{M}}{d\tau} = -\mathbf{e}_\perp\cdot\frac{d}{d\tau}(\gamma\mathbf{v}/v) = -\gamma\frac{\mathbf{e}_\perp\cdot\dot{\mathbf{v}}}{v} \tag{2.108}$$

Using (1.67), we can write the observed precession in terms of the time measured in the inertial frame of the lab

$$\frac{d\theta_0}{dt} = -\gamma\frac{\mathbf{e}_\perp\cdot\mathbf{a}}{v} \tag{2.109}$$

Now consider the case of circular motion in the lab frame. The velocity in this case changes direction at a range $\omega_c = a/v$ in the lab frame. This is the rotation of the axis relative to which $\theta_0$ was defined, and so the overall rotation observed will be given by

$$\frac{d\theta}{dt} = \frac{d\theta_0}{dt} + \omega_c = -(\gamma-1)\frac{a}{v} \tag{2.110}$$

Making use of (1.88), we can write this in vector notation as

$$\boxed{\boldsymbol{\omega}_T = -\frac{\mathbf{a}\times\mathbf{v}}{c^2}\frac{\gamma^2}{1+\gamma}} \tag{2.111}$$

This is the typical expression quoted for Thomas Precession. The dependence of the precession on the acceleration, and the non-collinearity of the acceleration and velocity, is clear from the above form. Note that despite this dependence on acceleration, no forces were used or specified in this derivation; the acceleration is already apparent, and the precession is a purely kinematic effect that arises due to the geometrical considerations of the Lorentz transformation.

# 3. *Electromagnetism*

This chapter aims to facilitate a relatively comprehensive study of electromagnetism in the context of Special Relativity, including

- An Introduction to the Covariant Formulation
- General Solutions to Maxwell's Equations
- Oscillations and Radiation
- Lagrangian Mechanics

The study of electromagnetism in the context of Special Relativity is probably one of the most general ways to tackle the subject. The results that we derive in this chapter will be applicable to all electromagnetic systems that occur in flat, Minkowski spacetime. It shall be assumed that readers are already familiar with the classical form of Maxwell's equations.

## 3.1 An Introduction to the Covariant Formulation

The entirety of classical electromagnetism is encapsulated by Maxwell's equations, which we shall state here for reference:

$$\boxed{\begin{aligned} \nabla \cdot \mathbf{E} &= \frac{\rho}{\epsilon_0} && \text{(MI)} \\ \nabla \cdot \mathbf{B} &= 0 && \text{(MII)} \\ \nabla \times \mathbf{E} &= -\frac{\partial \mathbf{B}}{\partial t} && \text{(MIII)} \\ \nabla \times \mathbf{B} &= \mu_0 \mathbf{j} + \mu_0 \epsilon_0 \frac{\partial \mathbf{E}}{\partial t} && \text{(MIV)} \end{aligned}}$$

where $\rho$ and $\mathbf{j} = \rho\mathbf{u}$ are the charge and current densities respectively. The immediate question to ask is as to whether this system of equations holds for all choices of reference frames? It turns out that Maxwell's equations are manifestly *covariant*, meaning that they self-consistently hold regardless of the choice of reference frame. It shall be our aim to demonstrate this fact.

### 3.1.1 Charge Conservation

A good starting point for this consideration is with charge conservation. Consider a volume $V$ bounded by a closed surface $S = \partial V$. In the absence of sources or sinks, the rate of change of the charge contained inside the volume is opposite and equal to the rate at which the charge leaves the volume through the bounding surface. Mathematically, this is stated in the familiar form

$$\frac{\partial}{\partial t}\int_V dV\, \rho = -\int_{\partial V} d\mathbf{S}\cdot\mathbf{j} = -\int_V dV\, \nabla\cdot\mathbf{j} \tag{3.1}$$

where the last expression follows from using the Divergence Theorem. As this equality must hold true for all possible choices of volumes, the integrands must be equal, allowing us to write the charge continuity equation

$$\frac{\partial \rho}{\partial t} + \nabla\cdot\mathbf{j} = 0 \tag{3.2}$$

Noting the form of (1.98), we introduce the four-current density

$$\boxed{\mathsf{J} = \rho_0 \mathsf{U} = (\rho c, \mathbf{j}), \quad \eta_{\mu\nu}\mathsf{J}^\mu \mathsf{J}^\nu = -\rho_0^2 c^2} \tag{3.3}$$

where $\rho_0$ is the charge density observed in some local rest frame. Then, it is clear that we can write (3.2) simply as

$$\boxed{\partial_\mu \mathsf{J}^\mu = 0} \tag{3.4}$$

In other words, that the four-divergence of the four-current is zero. From this, it is immediately obvious that charge conservation will continue to hold regardless of our choice of reference frame.

*A current carrying wire is electrically neutral in its rest frame $S$. The wire has some cross-sectional area, $A$ through which a uniform current $I$ follows. Show that in the rest frame of the current carries, there is a non-zero charge density in the wire. Find the electric field in $S'$.*

Assuming that all the charge carries have the same charge and drift velocity $\mathbf{v}$, the wire will remain electrostatically neutral, while it will have a four-current of

$$\mathsf{J} = (0, I/A, 0, 0) \tag{3.5}$$

where we have chosen (without loss of generality) to orient the wire along $\mathbf{e}_x$. Transforming to a frame moving at relative velocity $v$ to $S$ along the axis of the wire, the new four-current is

$$\mathsf{J}' = \Lambda\mathsf{J} = \left(-\frac{\gamma v I}{Ac^2}c, \frac{\gamma I}{A}, 0, 0\right) \tag{3.6}$$

It is clear that in this frame, there is a non-zero charge density in the wire. This does not imply that charge is not Lorentz invariant; instead, it means that charge density is not Lorentz invariant. There are two charge carries in the wire, one positive and the other negative. The non-zero current means that these charge carries have different velocities, and so their behaviour under a Lorentz transformation will be different. This will modify the charge densities of the two carriers differently, leading to the appearance of a charge.

The charge per unit length in frame $S'$ is

$$\lambda' = \rho' A = \frac{\gamma v I}{c^2} \tag{3.7}$$

The resultant magnitude of the electric field is then

$$E' = \frac{\lambda'}{2\pi\epsilon_0 r} = -\frac{\mu_0 v\gamma I}{2\pi r} = -v\gamma B \tag{3.8}$$

where $B = \mu_0 I/(2\pi r)$ is the magnetic field in frame $S$. An examination of why the electric field and magnetic field are related in this way shall be undertaken after a better understanding of the theory has been developed.

### 3.1.2 Potentials and Gauge

Equation (MII) tells us that the magnetic field $\mathbf{B}$ can always be written as the curl of some vector potential $\mathbf{A}$, namely that we may always define

$$\mathbf{B} = \nabla \times \mathbf{A} \tag{3.9}$$

Substituting this into (MIII), we have that

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} = -\nabla \times \frac{\partial \mathbf{A}}{\partial t} \quad \longrightarrow \quad \nabla \times \left(\mathbf{E} + \frac{\partial \mathbf{A}}{\partial t}\right) = 0 \tag{3.10}$$

This means that the quantity in brackets can be written as the gradient of some scalar potential $\phi$.

$$\mathbf{E} = -\nabla\phi - \frac{\partial \mathbf{A}}{\partial t} \tag{3.11}$$

Two of Maxwell's equations (MII) and (MIII) are thus automatically satisfied if we choose to write our $\mathbf{E}$ and $\mathbf{B}$ fields in terms of the scalar potential $\phi$ and the vector potential $\mathbf{A}$:

$$\boxed{\mathbf{E} = -\nabla\phi - \frac{\partial \mathbf{A}}{\partial t}, \quad \mathbf{B} = \nabla \times \mathbf{A}} \tag{3.12}$$

### Gauge Transformations

However, the choice of $\mathbf{A}$ is not unique. We can add the gradient of any scalar function $\chi = \chi(\mathbf{x}, t)$ to $\mathbf{A}$ without changing the magnetic field $\mathbf{B}$, since $\nabla \times (\nabla\chi) = 0$. Thus, we are at liberty to make the transformation

$$\mathbf{A} \mapsto \mathbf{A} + \nabla\chi \tag{3.13}$$

The electric field $\mathbf{E}$ will remain unchanged under the transformation (3.13) if $\phi$ simultaneously transforms as

$$\phi \mapsto \phi - \frac{\partial\chi}{\partial t} \tag{3.14}$$

The combination of the transformations

$$\boxed{\phi \mapsto \phi - \frac{\partial\chi}{\partial t}, \quad \mathbf{A} \mapsto \mathbf{A} + \nabla\chi} \tag{3.15}$$

is known as a *gauge transformation*, and has no effect on the observed dynamics of the system as the $\mathbf{E}$ and $\mathbf{B}$ fields will remain unchanged. We are thus free to choose our gauge in order to make our equations as simple as possible.

### Electromagnetic Four-Potential

We can encode both of these potentials into a single four-vector called the *electromagnetic four-potential*, defined by

$$\boxed{\mathsf{A} = (\phi/c, \mathbf{A}), \quad \eta_{\mu\nu}\mathsf{A}^\mu\mathsf{A}^\nu = -\phi^2/c^2 + \mathbf{A}^2} \tag{3.16}$$

It is clear that the gauge transformation (3.15) can be written as

$$\mathsf{A}^\mu \mapsto \mathsf{A}^\mu + \partial^\mu\chi \tag{3.17}$$

The fact that we are able to write down this four-potential is particularly significant, as it guarantees that (MII) and (MIII) are satisfied independently of our choice of frame. This follows from the fact that we can relate the $\mathsf{A}$ in different frames by a simple Lorentz transformation, meaning that $\phi$ and $\mathbf{A}$ - and consequently $\mathbf{E}$ and $\mathbf{B}$ - follow suit. What about the other two of Maxwell's equations? To investigate this, substitute (3.12) into (MI) and (MIV), yielding

$$-\nabla^2\phi - \frac{\partial}{\partial t}\nabla\cdot\mathbf{A} = \frac{\rho}{\epsilon_0} \tag{3.18}$$

$$\nabla(\nabla\cdot\mathbf{A}) + \frac{1}{c^2}\frac{\partial}{\partial t}\nabla\phi - \left(-\frac{1}{c^2}\frac{\partial^2}{\partial t^2} + \nabla^2\right)\mathbf{A} = \mu_0\mathbf{j} \tag{3.19}$$

We can write these in the more transparent form of

$$\left(-\frac{1}{c^2}\frac{\partial^2}{\partial t^2} + \nabla^2\right)(\phi/c) + \frac{1}{c}\frac{\partial}{\partial t}\left(\frac{1}{c}\frac{\partial}{\partial t}(\phi/c) + \nabla\cdot\mathbf{A}\right) = -\mu_0(\rho c) \tag{3.20}$$

$$\left(-\frac{1}{c^2}\frac{\partial^2}{\partial t^2} + \nabla^2\right)\mathbf{A} - \nabla\left(\frac{1}{c}\frac{\partial}{\partial t}(\phi/c) + \nabla\cdot\mathbf{A}\right) = \mu_0\mathbf{j} \tag{3.21}$$

These may not initially seem more transparent than the previous two equations; in fact, they appear even more complicated! However, looking at the form of (1.98), and recalling that raising or lowering an index under the Minkowski metric simply changes the sign of the zeroth component of a four-vector, it is clear after a bit of staring that (3.20) and (3.21) can be written as the single equation

$$\boxed{\partial_\mu \partial^\mu \mathsf{A}^\nu - \partial^\nu(\partial_\mu \mathsf{A}^\mu) = -\mu_0 \mathsf{J}^\nu} \tag{3.22}$$

In order to simplify this further, we make the choice of gauge condition that $\partial_\mu \mathsf{A}^\mu = 0$, which is known as the *Lorenz gauge*. A sensible question to ask is to whether we are always able to make this gauge choice. Suppose that we instead have that $\partial_\mu \mathsf{A}^\mu = \zeta$, for some scalar function $\zeta = \zeta(\mathbf{x}, t)$. We can then introduce the gauge transformation

$$\partial_\mu \bar{\mathsf{A}}^\mu = \partial_\mu \mathsf{A}^\mu + \partial_\mu \partial^\mu \chi = \zeta + \partial_\mu \partial^\mu \chi \tag{3.23}$$

However, we are always able to choose a $\chi$ satisfying $\partial_\mu \partial^\mu \chi = \zeta$ such that $\partial_\mu \bar{\mathsf{A}}^\mu = 0$. Thus, the choice of the Lorenz gauge is always available. This means that (3.22) becomes

$$\partial_\mu \partial^\mu \mathsf{A}^\nu = -\mu_0 \mathsf{J}^\nu \tag{3.24}$$

This is quite a neat result. As (MII) and (MIII) are automatically satisfied in the definition of $\mathsf{A}$, and (3.24) explicitly encodes (MI) and (MIV), we have been able to include all of Maxwell's equations in a single expression. Further to this, it is now also clear that all of Maxwell's equations are Lorentz covariant. Both $\mathsf{A}$ and $\mathsf{J}$ transform in the same way under the Lorentz transformation

$$\mathsf{A}'^\mu = \Lambda^\mu{}_\nu \mathsf{A}^\nu, \quad \mathsf{J}'^\mu = \Lambda^\mu{}_\nu \mathsf{J}^\nu \tag{3.25}$$

while $\partial_\nu \partial^\nu$ is invariant, meaning that (3.24) must hold true for all choices of inertial frame.

### Field of Uniformly Moving Charge

Before developing our theory further, we shall pause to take a look at a an illustrative system, that of a uniformly moving charge. The four-potential in the rest frame of the charge is given by

$$\mathsf{A}' = (\phi'/c, 0), \quad \phi' = \frac{q}{4\pi\epsilon_0}\frac{1}{r'} \tag{3.26}$$

where $\phi'$ is the Coulomb potential of a point charge at the origin of our coordinate system in $S'$. Then, the observed four-potential in the lab frame is given by

$$\mathsf{A} = \Lambda^{-1}\mathsf{A}' = \left(\gamma\phi'/c, (\gamma v/c^2)\phi', 0, 0\right) \tag{3.27}$$

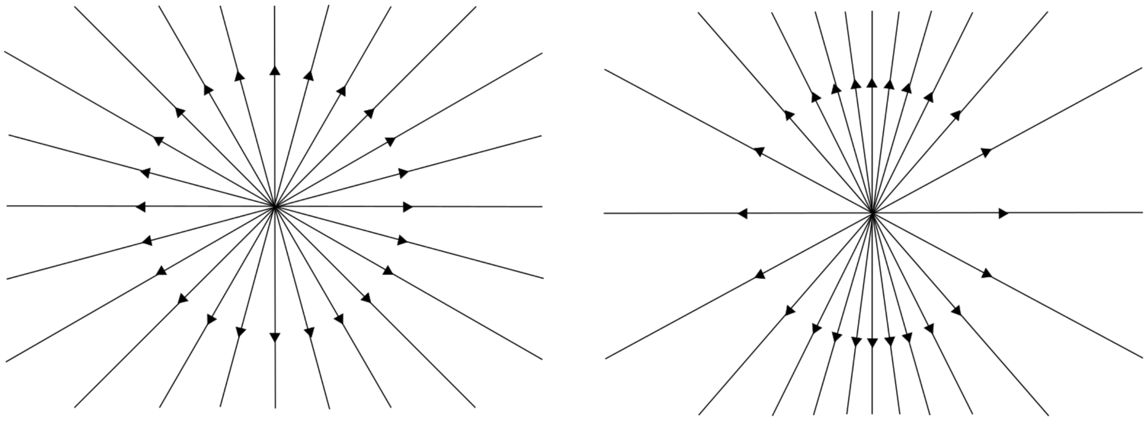
Recognising that the radial coordinate in the rest frame of the charge $r'$ can be related to the Cartesian coordinates in the lab frame by (1.55) and (1.56):

$$r' = \left(x'^2 + y'^2 + z'^2\right)^{1/2} = \left(\gamma^2(x - vt)^2 + y^2 + z^2\right)^{1/2} \tag{3.28}$$

It follows that the resultant electric field in the lab frame can be written as

$$\boxed{\mathbf{E} = \frac{\gamma q}{4\pi\epsilon_0}\frac{(\mathbf{r} - \mathbf{v}t)}{[\gamma^2(x - vt)^2 + y^2 + z^2]^{3/2}}} \tag{3.29}$$

What does the field look like in the lab frame? For large $\gamma$, it is clear that the expression the denominator is dominated by the dependence on $x$. From this, we can infer that the field lines becomes stretched along the direction of motion ($\mathbf{e}_x$), and bunched along the directions perpendicular to the motion. This is shown in figure 3.1. The field is thus no-longer spherically symmetric, and points radially outwards from the *projected position* of the charge at $\mathbf{r} = \mathbf{v}t$. Information about changes in these field lines can only propagate



(a) (b)

Figure 3.1: The field of a point charge (a) At rest (b) In uniform (non-accelerated) motion

outwards from the charge at $c$. Interestingly, we also observe a magnetic field in the lab frame, despite there being no magnetic field in the rest frame of the charge. This often why magnetism is often seen simply as a relativistic correction to electric fields. Explicitly, the magnetic field is given by (3.12)

$$\mathbf{B} = \frac{\gamma q}{4\pi\epsilon_0 c^2}\frac{(\mathbf{v}\times\mathbf{r})}{[\gamma^2(x-vt)^2+y^2+z^2]^{3/2}} \tag{3.30}$$

Let us now make the simplifying assumption that the charge passes through the origin of the lab frame $S$ at time $t$, such that $(ct', \mathbf{r}') = (ct, \mathbf{r}) = 0$, meaning that we can simply replace $x - vt \mapsto x$. Then, we can manipulate the denominator:

$$(\gamma x)^2 + y^2 + z^2 = (\gamma x)^2 + x^2 + y^2 + z^2 - x^2 = r^2 + (\gamma^2 - 1)x^2 = r^2\left(1 + v_r^2\gamma^2/c^2\right) \tag{3.31}$$

where we have used (1.88) and defined $v_r = \mathbf{r}\cdot\mathbf{v}/r$ as the radial component of the velocity. Then, the magnetic field can be written as

$$\mathbf{B} = \frac{\gamma q}{4\pi\epsilon_0 c^2}\frac{(\mathbf{v}\times\mathbf{r})}{r^3\,[1+v_r^2\gamma^2/c^2]^{3/2}} \tag{3.32}$$

In the non-relativistic limit, we can neglect terms in $v_r/c$ and let $\gamma \to 1$, such that

$$\mathbf{B} \to \frac{\gamma q}{4\pi\epsilon_0 c^2}\frac{\mathbf{v}\times\mathbf{r}}{r^3} = \frac{\mu_0 I}{4\pi}\frac{\mathbf{e}_v\times\mathbf{r}}{r^3} \tag{3.33}$$

for $I = q|\mathbf{v}|$, and $\mathbf{e}_v$ is the unit vector along $\mathbf{v}$. This is the familiar expression for the magnetic field that we would obtain by using the Biot-Savart law on a point charge.

### 3.1.3 The Electromagnetic Field Tensor

A consideration of electromagnetic forces - as governed by the Lorentz force equation - has thus far been absent in this chapter. Let us rectify this by looking at the way that we incorporate the considerations of forces into our theory.

The Lorentz force equation for a charge moving at velocity $\mathbf{u}$

$$\mathbf{f} = q(\mathbf{E} + \mathbf{u}\times\mathbf{B}) \tag{3.34}$$

in this form appears as a pure force. Motivated by this, we shall look for an expression for the force $\mathsf{F}$ that satisfies $\mathsf{U}\cdot\mathsf{F}=0$. If the force were independent of $\mathsf{U}$, then this dot product can only vanish for all $\mathsf{U}$ if $\mathsf{F}$ itself was zero. This means that $\mathsf{F}$ must depend on $\mathsf{U}$ in some way. Consider

$$\boxed{\mathsf{F}^\mu = q\mathbb{F}^{\mu\nu}\mathsf{U}_\nu} \tag{3.35}$$

where $q$ is the associated charge, and $\mathbb{F}$ is a rank-2 tensor called the *electromagnetic field tensor*. What are the conditions on $\mathbb{F}$ in order for it to satisfy the pure force condition? Dot through by $\mathsf{U}$ on both sides of this equation:

$$\mathsf{U}_\mu\mathsf{F}^\mu = q\mathbb{F}^{\mu\nu}\mathsf{U}_\mu\mathsf{U}_\nu = q\mathsf{F}^{\mu\nu}\left(\mathsf{U}\mathsf{U}\right)_{\mu\nu} \overset{!}{=} 0 \tag{3.36}$$

Recognising that $(\mathsf{U}\mathsf{U})_{\mu\nu}$ is a symmetric quantity, the only way for the last equality to hold for all choices of $\mathsf{U}$ is if $\mathbb{F}^{\mu\nu}$ is antisymmetric in its indices. As this must be in some way related to the four-potential $\mathsf{A}$, we propose that the electromagnetic field tensor takes the form

$$\boxed{\mathbb{F}^{\mu\nu} = \partial^\mu\mathsf{A}^\nu - \partial^\nu\mathsf{A}^\mu} \tag{3.37}$$

In terms of components, this is given by

$$\mathbb{F}^{00}=0, \quad \mathbb{F}^{0i} = -\mathbb{F}^{i0} = E^i/c, \quad \mathbb{F}^{ij} = \epsilon_{ijk}B_k \tag{3.38}$$

where $\mathbf{E}$ and $\mathbf{B}$ are the electric and magnetic fields. (3.37) can also be written explicitly as

$$\mathbb{F}^{\mu\nu} = \left(\begin{array}{c|ccc} 0 & & \mathbf{E}/c & \\ \hline & 0 & B_z & -B_y \\ -\mathbf{E}/c & -B_z & 0 & B_x \\ & B_y & -B_x & 0 \end{array}\right) \tag{3.39}$$

It is thus clear that all the information about our fields is encoded in this single tensor. Comparison of the form of (3.37) with (3.22) leads us to write that

$$\boxed{\partial_\mu\mathbb{F}^{\mu\nu} = -\mu_0\mathsf{J}^\nu} \tag{3.40}$$

By construction, this equation satisfies all the conditions that we previously placed on $\mathsf{A}$. For example, we can obtain charge conservation immediately from this expression:

$$\partial_\nu\partial_\mu\mathbb{F}^{\mu\nu} = -\partial_\mu\partial_\nu\mathbb{F}^{\nu\mu} = \partial_\nu\partial_\mu\mathbb{F}^{\nu\mu} = 0 \quad\longrightarrow\quad \partial_\nu\mathsf{J}^\nu = 0 \tag{3.41}$$

Together, (3.37) and (3.71) are the *field equations* that encapsulate all of Maxwell's equations in a manifestly covariant form. In most cases, this will have been the first field theory formulation that readers will have encountered.

*Show that the Lorentz force equation is the spatial components of* (3.35). *If* $\mathbb{F}$ *contains only constant electric and magnetic fields, then show that the motion is hyperbolic, and find the magnitude of the proper acceleration in the case that* $\mathbf{E}=(E_0,0,0)$ *and* $\mathbf{B}=0$.

Evaluating the left and right-hand sides of (3.35) explicitly yields:

$$\gamma\left(\frac{1}{c}\frac{dE}{dt},\mathbf{f}\right) = \gamma q\left(\frac{\mathbf{u}\cdot\mathbf{E}}{c},\mathbf{E}+\mathbf{u}\times\mathbf{B}\right) \tag{3.42}$$

The spatial part of this equation is clearly (3.34). In four-vector form, the above equation can also be written as

$$\dot{\mathsf{U}}^\mu = \frac{q}{m}\mathbb{F}^{\mu\nu}\mathsf{U}_\nu \tag{3.43}$$

The rate of change of the proper acceleration is then given by

$$\frac{d}{d\tau}\left(a_0^2\right) = \frac{d}{d\tau}\left(\dot{\mathsf{U}}_\mu \dot{\mathsf{U}}^\mu\right) = 2\dot{\mathsf{U}}_\mu \ddot{\mathsf{U}}^\mu \tag{3.44}$$

However, for constant **E** and **B** fields:

$$\ddot{\mathsf{U}}^\mu = \frac{d}{d\tau}\left(\frac{q}{m}\mathbb{F}^{\mu\nu}\mathsf{U}_\nu\right) = \frac{q}{m}\mathbb{F}^{\mu\nu}\dot{\mathsf{U}}_\nu \tag{3.45}$$

Combining these two equations, it follows that

$$\frac{d}{d\tau}\left(a_0^2\right) = \frac{2q}{m}\mathbb{F}^{\mu\nu}\dot{\mathsf{U}}_\mu \dot{\mathsf{U}}_\nu = 0 \tag{3.46}$$

as we have a symmetric quantity multiplied by an antisymmetric quantity. Thus, the magnitude of the proper acceleration is constant, meaning that the motion is indeed hyperbolic. We can find the magnitude of $a_0$ by solving the equation of motion (3.43).

$$\mathsf{U}^\mu = \frac{q}{m}\mathbb{F}^{\mu\nu}\int d\tau\ \mathsf{U}_\nu = \frac{q}{m}\mathbb{F}^{\mu\nu}\mathsf{X}_\nu \tag{3.47}$$

Using the forms of the **E** and **B** fields in $\mathbb{F}^{\mu\nu}$, we can write this as

$$\frac{d\mathsf{X}}{d\tau} = \frac{qE}{mc}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\mathsf{X} \tag{3.48}$$

where we have suppressed all irrelevant zero components. By making a substitution of the form $\mathsf{X} = \mathsf{X}_0 e^{\Gamma\tau}$ - with $\mathsf{X}_0$ being a constant four-vector - it can be shown that the possible solutions are

$$\Gamma = \frac{qE}{mc}, \quad \mathsf{X}_0 = (1, 1) \tag{3.49}$$

$$\Gamma = -\frac{qE}{mc}, \quad \mathsf{X}_0 = (1, -1) \tag{3.50}$$

such that the full solution can be written as

$$\mathsf{X} = L(\sinh\Gamma\tau, \cosh\Gamma\tau) \tag{3.51}$$

Evaluating the invariant, it is clear that the motion satisfies the equation of a hyperbola, namely that

$$-(ct)^2 + x^2 = -L^2(\sinh^2\Gamma\tau - \cosh^2\Gamma\tau) = -L^2 \tag{3.52}$$

From this solution, we can find $\dot{\mathsf{U}}$, and thus show that $a_0 = L(qE/mc)^2$.

### Invariants of $\mathbb{F}$

Before we can calculate the invariant quantities that are associated with our field tensor $\mathbb{F}$, we first need to introduce the *dual* electromagnetic field tensor, given by

$$\widetilde{\mathbb{F}}_{\mu\nu} = \frac{1}{2}\epsilon_{\mu\nu\rho\sigma}\mathbb{F}^{\rho\sigma} \tag{3.53}$$

This definition might seem quite complicated, but all that it amounts to is that we obtain an expression similar to the field tensor, except that **E** and **B** are reversed in the layout of the tensor (see (3.39) for clarification). We are thus able to form two invariant quantities from our field tensor:

$$D = \frac{1}{2}\mathbb{F}^{\mu\nu}\mathbb{F}_{\mu\nu} = -E^2/c^2 + B^2 \tag{3.54}$$

$$\alpha = \frac{1}{4}\mathbb{F}^{\mu\nu}\widetilde{\mathbb{F}}_{\mu\nu} = \mathbf{E}\cdot\mathbf{B}/c \tag{3.55}$$

The second of these contains the interesting statement that if the **E** and **B** are orthogonal in one frame, they must be orthogonal in all frames definition.

**Field Transformations**

Being a rank-2 tensor, the field tensor transforms according to

$$\mathbb{F}' = \Lambda \mathbb{F} \Lambda^T \quad \longleftrightarrow \quad \mathbb{F}'^{\mu\nu} = \Lambda^{\mu}{}_{\rho} \Lambda^{\nu}{}_{\sigma} \mathbb{F}^{\rho\sigma} \tag{3.56}$$

By either writing out the matrix product explicitly, or evaluating this in component form, we can derive transformations that relating the $\mathbf{E}$ and $\mathbf{B}$ fields in different inertial frames. For example, consider

$$\mathbb{F}'^{01} = \Lambda^0{}_0 \Lambda^1{}_1 \mathbb{F}^{01} + \Lambda^0{}_1 \Lambda^1{}_0 \mathbb{F}^{10} = \gamma^2(1-\beta^2)\mathbb{F}^{01} = \mathbb{F}^{01} \tag{3.57}$$

This means that $\mathbf{E}_{||}$ remains unchanged in the transformation. Carrying this out for each of the components, we obtain the transformation equations

$$\mathbf{E}'_{||} = \mathbf{E}_{||} \tag{3.58}$$

$$\mathbf{B}'_{||} = \mathbf{B}_{||} \tag{3.59}$$

$$\mathbf{E}'_{\perp} = \gamma(\mathbf{E}_{\perp} + \mathbf{v} \times \mathbf{B}) \tag{3.60}$$

$$\mathbf{B}'_{\perp} = \gamma\left(\mathbf{B}_{\perp} - \frac{1}{c^2}\mathbf{v} \times \mathbf{E}\right) \tag{3.61}$$

where $\mathbf{v}$ is the relative velocity between the frames. The last of these equations is particularly interesting. Consider an interial frame in which there is initially no magnetic field, so $\mathbf{B} = 0$. Then, in some other frame moving at a relative velocity $\mathbf{v}$, there will be a magnetic field that is related to the electric field by

$$\mathbf{B}_{\perp} = \frac{1}{c^2}\mathbf{v} \times \mathbf{E} \tag{3.62}$$

This gives further evidence for thinking about magnetic fields as relativistic corrections to electric fields. For example, form of (3.8) follows from (3.60), while we could have found the magnetic field in section 3.1.2 immediately through the use of the above equation (as there is initially no magnetic field in the rest frame of the charge). In the latter case, the magnetic field lines form loops around the direction of motion of the charge.

### 3.1.4 The Stress-Energy Tensor

Thus far, we have only associated momentum with particles, or more macroscopic objects. However, the aim of this section is to show that the electromagnetic field itself has some associated momentum, and that particles are able to exchange energy and momentum with the electromagnetic field.

Let us initially revisit a familiar argument when it comes to energy conservation in electromagnetism, in order to inform our development of a more general formalism. Consider the work done by the electromagnetic on a single charge moving at a velocity $\mathbf{u}$:

$$dW = \mathbf{f} \cdot d\boldsymbol{\ell} = q(\mathbf{E} + \mathbf{u} \times \mathbf{B}) \cdot \mathbf{v} dt = q\mathbf{u} \cdot \mathbf{E} dt \tag{3.63}$$

Assuming that this single charge is part of a larger ensemble of $n$ charges per unit volume that are only weakly interacting, meaning that we can write that

$$\frac{dW}{dt} = \int_V dV\, nq\mathbf{E} \cdot \mathbf{u} = \int_V dV\, \mathbf{j} \cdot \mathbf{E} \tag{3.64}$$

Thus, the rate of work done on charges per unit volume is given by $\mathbf{j}\cdot\mathbf{E}$ (Ohmic heating). Consider (MIV):

$$\begin{aligned}-\mathbf{j}\cdot\mathbf{E} &= -\left(\frac{1}{\mu_0}\nabla\times\mathbf{B}-\epsilon_0\frac{\partial\mathbf{E}}{\partial t}\right)\cdot\mathbf{E}+\underbrace{\left(\nabla\times\mathbf{E}+\frac{\partial\mathbf{B}}{\partial t}\right)\cdot\frac{\mathbf{B}}{\mu_0}}_{=0\text{ by (MIV)}} \\ &= \epsilon_0\mathbf{E}\cdot\frac{\partial\mathbf{E}}{\partial t}+\frac{1}{\mu_0}\mathbf{B}\cdot\frac{\partial\mathbf{B}}{\partial t}+\frac{1}{\mu_0}\left(\mathbf{B}\cdot\nabla\times\mathbf{E}-\mathbf{E}\cdot\nabla\times\mathbf{B}\right) \\ &= \frac{\partial}{\partial t}\left(\frac{1}{2}\epsilon_0E^2+\frac{1}{2\mu_0}B^2\right)+\frac{1}{\mu_0}\nabla\cdot(\mathbf{E}\times\mathbf{B})\end{aligned} \tag{3.65}$$

We now define the *electromagnetic field energy density* $u$ and *Poynting vector* $\mathbf{N}$ by

$$\boxed{u=\frac{1}{2}\epsilon_0E^2+\frac{1}{2\mu_0}B^2,\quad \mathbf{N}=\frac{1}{\mu_0}\mathbf{E}\times\mathbf{B}} \tag{3.66}$$

The Poynting vector gives the energy flux that is associated with the fields $\mathbf{E}$ and $\mathbf{B}$. We can thus write our familiar statement of energy conservation as

$$\frac{\partial u}{\partial t}+\nabla\cdot\mathbf{N}+\mathbf{j}\cdot\mathbf{E}=0 \tag{3.67}$$

It is not immediately obvious (to this author, at least) that this equation will be covariant, which should be a cause for alarm. We cannot have a change in the way that electromagnetic systems conserve energy simply based on our choice of inertial frame for observation. As such, we aim to look for a statement of combined *energy-momentum conservation* that is indeed covariant. We need to take into account both the rate of transfer of energy from the field to the particles, as well as the rate of transfer of momentum from the field to the particles. We thus want to equate the rate of transfer of energy-momentum to the particles to a quantity that must represent the rate of transfer of energy-momentum out of the field.

By analogy to (3.64), we define the rate of transfer of energy-momentum per unit volume from the field to the particles by a moment of the field tensor

$$\boxed{\mathsf{W}^\mu=\mathbb{F}^{\mu\nu}\mathsf{J}_\nu,\quad \mathsf{W}=(\mathbf{j}\cdot\mathbf{E}/c,\rho\mathbf{E}+\mathbf{j}\times\mathbf{B})} \tag{3.68}$$

This is known as the *Lorentz force density*, as it clearly has dimensions of force per unit volume. We expect to see a force in this expression, as a rate of transfer of momentum is simply a force. Particles thus feel a force as a result of the field that is governed by the Lorentz force equation. This is then related to the change in the field quantities by the expression

$$\boxed{\mathsf{W}^\mu=-\partial_\nu\mathbb{T}^{\mu\nu}} \tag{3.69}$$

that encompasses all energy-momentum conservation in a covariant treatment of electromagnetism. $\mathbb{T}$ is known as the *stress-energy* tensor, for reasons which will eventually be made obvious. It has physical dimensions of energy per unit volume, which can be thought of as the energy per unit volume associated with the field.

We shall now find an explicit expression for $\mathbb{T}$, by considering the form of $\mathsf{W}$. Again by analogy to (3.65), we want to express $\mathsf{J}$ in terms of the field tensor. Using (3.71), we write that

$$\mathsf{W}^\mu=-\epsilon_0c^2\mathbb{F}^{\mu\nu}\partial^\rho\mathbb{F}_{\rho\nu}=-\epsilon_0c^2=-\epsilon_0c^2\left[\partial^\rho\left(\mathbb{F}^{\mu\nu}\mathbb{F}_{\rho\nu}\right)-\mathbb{F}_{\rho\nu}\left(\partial^\rho\mathbb{F}^{\mu\nu}\right)\right] \tag{3.70}$$

The first of these expressions is recognisable as being in the form of the divergence of some tensor quantity, and so we concentrate on the second expression. The form of the field tensor (3.37) can be written in the form

$$\partial^\rho \mathbb{F}^{\mu\nu} + \partial^\mu \mathbb{F}^{\nu\rho} + \partial^\nu \mathbb{F}^{\rho\mu} = 0 \tag{3.71}$$

Then,

$$\mathbb{F}_{\rho\nu}\left(\partial^\rho \mathbb{F}^{\mu\nu}\right) = -\mathbb{F}_{\rho\nu}\partial^\mu \mathbb{F}^{\nu\rho} - \mathbb{F}_{\rho\nu}\partial^\nu \mathbb{F}^{\rho\mu} = \mathbb{F}_{\rho\nu}\partial^\mu \mathbb{F}^{\rho\nu} - \mathbb{F}_{\rho\nu}\partial^\nu \mathbb{F}^{\rho\mu} = \partial^\mu D - \mathbb{F}_{\rho\nu}\partial^\nu \mathbb{F}^{\rho\mu} \tag{3.72}$$

for $D$ as defined in (3.54). The last term on the right-hand side appears to be in the same form as the initial term on the left-hand side, except with different dummy indices. However, as the only different indices are those that are contracted over, we can re-label them as required, yielding

$$\mathbb{F}_{\rho\nu}\left(\partial^\rho \mathbb{F}^{\mu\nu}\right) = \frac{1}{2}\partial^\mu D \tag{3.73}$$

We thus have that

$$\mathsf{W}^\mu = -\epsilon_0 c^2\left(\partial^\rho\left(\mathbb{F}^{\mu\nu}\mathbb{F}_{\rho\nu}\right) - \frac{1}{2}\partial^\mu D\right) \quad\longrightarrow\quad \mathsf{W}^\mu = \epsilon_0 c^2 \partial_\nu\left(\mathbb{F}^{\mu\rho}\mathbb{F}_\rho{}^\nu + \frac{1}{2}\eta^{\mu\nu} D\right) \tag{3.74}$$

where the section expression is obtain from a relabelling of indices, appropriate raising and lowering using the metric, as well as using the asymmetric property of $\mathbb{F}$. Comparison of this with (3.69), it is clear that

$$\boxed{\mathbb{T}^{\mu\nu} = -\epsilon_0 c^2\left(\mathbb{F}^{\mu\rho}\mathbb{F}_\rho{}^\nu + \frac{1}{2}\eta^{\mu\nu} D\right), \quad D = \frac{1}{2}\mathbb{F}^{\lambda\sigma}\mathbb{F}_{\lambda\sigma}} \tag{3.75}$$

This is the most general form of the stress-energy tensor in electromagnetic theory in Minkowski space. The presence of $\mathbb{F}$ in this definition means that the rate of energy-momentum transfer to the fields are governed by the fields themselves. Note that $\mathbb{T}^{\mu\nu}$ - unlike $\mathbb{F}^{\mu\nu}$ - is completely symmetric in its indices.

**Interpreting $\mathbb{T}$**

Using the form of (3.39), we find that the stress-energy tensor can be written as

$$\mathbb{T}^{\mu\nu} = \left(\begin{array}{c|c} u & \mathbf{N}/c \\ \hline \mathbf{N}/c & \sigma_{ij}\end{array}\right), \quad \sigma_{ij} = u\delta_{ij} - \epsilon_0(E_iE_j + c^2 B_iB_j) \tag{3.76}$$

where $u$ and $\mathbf{N}$ are defined as per (3.66). The quantity $\sigma_{ij}$ is the three-stress tensor. We shall examine its consequences shortly.

First, how do we interpret this form of $\mathbb{T}$? Suppose that $\widehat{\mathsf{E}}$ represents some normalised four-vector direction. Then, $\mathbb{T}^{\mu\nu}\widehat{\mathsf{E}}_\nu$ quantifies the flow of energy and momentum in that direction. The first row of $\mathbb{T}$ is used to calculate the exchange of energy, while the elements of the first column are used, together with $\sigma_{ij}$, to calculate the flow of momentum. The fact that the first row is equal to the first column represents the equality of energy flux and momentum density.

The three-stress tensor $\sigma_{ij}$ has units of momentum per unit area, per unit time. This means that each component represents the flux of momentum in the $i^{\text{th}}$ direction, crossing a surface with normal along the $j^{\text{th}}$ direction. In other words, it gives the force per unit area (pressure) in the $i^{\text{th}}$ direction on a surface with normal along the $j^{\text{th}}$ direction. For diagonal $\sigma_{ij}$, it is very easy to read off the forces involved from the components of the tensor. Note that negative pressure is often referred to as a *tension*.

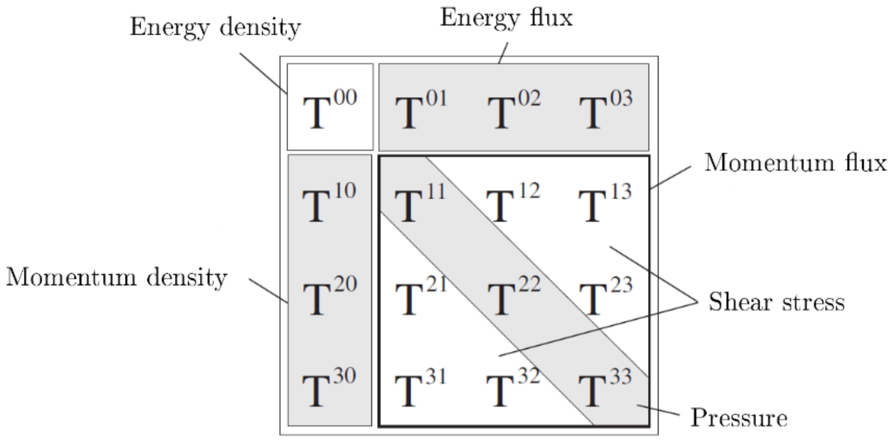



Figure 3.2: A summary of the physics encoded in the stress-energy tensor

Now that we have a physical interpretation of the stress-energy tensor, let us take another look at energy-momentum conservation, as given by (3.69). From the time component of $\mathsf{W}$, we re-obtain the familiar energy conservation result:

$$\mathsf{W}^0 = -\partial_\nu \mathbb{T}^{0\nu} = -\partial_0 \mathbb{T}^{00} - \partial_j \mathbb{T}^{0j} \quad \longrightarrow \quad \mathbf{j}\cdot\mathbf{E} = -\frac{\partial u}{\partial t} - \nabla\cdot\mathbf{N} \tag{3.77}$$

It is often useful to define the quantity $\mathbf{g} = \mathbf{N}/c^2$ as the momentum per unit volume carried by the field. Given this, we can evaluate the components $\mathsf{W}^i$:

$$W^i = -\partial_\nu \mathbb{T}^{i\nu} \quad \longrightarrow \quad \rho(\mathbf{E} + \mathbf{u}\times\mathbf{B})_i = -\frac{\partial g_i}{\partial t} - \partial_j \sigma_{ij} \tag{3.78}$$

Integrating this over some volume $V$ bounded by a closed surface $S = \partial V$, and using the divergence theorem, it follows that

$$q(\mathbf{E} + \mathbf{u}\times\mathbf{B})_i = -\frac{\partial}{\partial t}\int_V dV\, g_i - \int_{\partial V} dS\, e_j \sigma_{ij} \tag{3.79}$$

where $e_j$ is the unit vector in the $j^{\text{th}}$ direction. This equation can be seen as a statement of Newtons Third law for the interaction between the charge and the field. The left-hand side is the rate of momentum transfer from the field to the particles, as encapsulated by the Lorentz force. On the right-hand side, the first term is the rate of change of momentum in the field inside the given volume, while the second is the flux of field momentum into the volume from other regions due to $\sigma_{ij}$. Like with energy above, momentum is also conserved; it is either stored in particles, in the field, or flows from one region to another!

### Some Simple cases of $\mathbb{T}$

Let us consider some simple cases of the stress energy tensor. These are easily calculated from the form (3.76)

- Capacitor plates orientated along $\mathbf{e}_1$:

$$\mathbb{T}^{\mu\nu} = \frac{1}{2}\epsilon_0 E^2 \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{3.80}$$

A tension along $\mathbf{e}_1$ means that the field is pulling inwards on the capacitor plates, an effect that we would usually describe by the attraction between opposite charges on each plate surface. The other components tell us that there is a pressure outwards at right angles to the field direction. In general - in the absence of magnetic fields - there is always a tension along field lines (attraction between charges) and a pressure at right angles to the field lines, pushing them apart (repulsion between charges)

- A semi-infinite solenoid with axis oriented along $\mathbf{e}_1$:

$$\mathbb{T}^{\mu\nu} = \frac{1}{2\mu_0}E^2 \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{3.81}$$

  There is a tension along the axis of the solenoid, while there is an outwards pressure on the walls of the solenoid

- A plane wave of angular frequency $\omega$ propagating along $\mathbf{e}_1$:

$$\mathbb{T}^{\mu\nu} = \epsilon_0 E_0^2 \cos^2(kx - \omega t) \begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \tag{3.82}$$

  We have taken the fields to be $\mathbf{E} = E_0 \cos(kx - \omega t)\mathbf{e}_2$ and $\mathbf{B} = (E_0/c)\cos(kx - \omega t)\mathbf{e}_3$. It is clear that plane electromagnetic waves have an associated pressure along the direction of motion, but no components perpendicular to said direction

The case of the plane electromagnetic wave is worth investigating further. It is clear that the wave four-vector is given by $\mathsf{K} = k(1, 1, 0, 0)$. Recalling (1.93), and using the vacuum dispersion relation $\omega = ck$, we can write the stress energy tensor as

$$\boxed{\mathbb{T}^{\mu\nu} = \epsilon_0 c^2 \left(\frac{E_0}{\omega}\right)^2 \cos^2\left(\mathsf{X}_\rho \mathsf{K}^\rho\right) \mathsf{K}^\mu \mathsf{K}^\nu} \tag{3.83}$$

which is actually the general relationship for a plane, monochromatic electromagnetic wave. Now, both sides of this equation must transform in the same way, as they are both tensors. As $\mathsf{X}_\rho \mathsf{K}^\rho$ is already an invariant quantity, we must conclude that $E_0^2/\omega^2$ is also an invariant scalar. Recalling that $u = \epsilon_0 E_0^2$, $\mathbf{g} = \mathbf{N}/c^2 \propto u/c\,\mathbf{e_k}$ and $\mathbf{p} = \langle \mathbf{N} \rangle / c$ for a plane wave, it follows that

$$\frac{E_0'^2}{E_0^2} = \frac{\omega'^2}{\omega^2} = \frac{I'}{I} = \frac{u'}{u} = \frac{g'}{g} = \frac{p'}{p} \tag{3.84}$$

for two inertial frames $S$ and $S'$ with some relative velocity. The fact that the intensity transforms in this way allows us to conclude that the width of the wavefront remains the same in all frames.

## 3.2 General Solutions to Maxwell's Equations

Field Tensor Equations:

$$\mathbb{F}^{\mu\nu} = \partial^\mu \mathsf{A}^\nu - \partial^\nu \mathsf{A}^\mu \tag{3.85}$$

$$\partial_\mu \mathbb{F}^{\mu\nu} = -\mu_0 \mathsf{J}^\nu \tag{3.86}$$

$$\mathsf{F}^\mu = q\mathbb{F}^{\mu\nu}\mathsf{U}_\nu \tag{3.87}$$

Conservation Equations:

$$\mathsf{W}^\mu = -\partial_\nu \mathbb{T}^{\mu\nu} \tag{3.88}$$

$$\mathsf{W}^\mu = \mathbb{F}^{\mu\nu}\mathsf{J}_\nu \tag{3.89}$$

$$\mathbb{T}^{\mu\nu} = -\epsilon_0 c^2 \left( \mathbb{F}^{\mu\rho}\mathbb{F}_\rho{}^\nu + \frac{1}{2}\eta^{\mu\nu} D \right), \quad D = \frac{1}{2}\mathbb{F}^{\lambda\sigma}\mathbb{F}_{\lambda\sigma} \tag{3.90}$$

In the previous section, we derived the equations that govern our covariant field theory formulation of electromagnetism, the principal equations of which have been included in the box above. We shall now consider some general solutions to these equations that we can be applied to a variety of systems.

### 3.2.1 Retarded Potentials

In order to characterise the electric and magnetic fields present in a system, we want to develop a way to determine the potentials $\phi$ and $\mathbf{A}$ given an arbitrary distribution of charges $\rho$ and currents $\mathbf{j}$. To begin, substitute (3.85) into (3.86), yielding

$$\partial_\mu \partial^\mu \mathsf{A}^\nu - \partial_\mu \left( \partial^\nu \mathsf{A}^\mu \right) = -\mu_0 \mathsf{J}^\nu \tag{3.91}$$

Introducing the Lorenz gauge, this becomes

$$\partial_\mu \partial^\mu \mathsf{A}^\nu = \left( -\frac{1}{c^2}\frac{\partial^2}{\partial t^2} + \nabla^2 \right) \mathsf{A}^\nu = -\mu_0 \mathsf{J}^\nu \tag{3.92}$$

This is clearly a wave-equation for the four-potential $\mathsf{A}$ for some arbitrary source term $\mathsf{J}$. Solving this equation will give us the relationship we need between the potentials and the source terms.

The Green's function for the d'Alembertian $\partial_\mu \partial^\mu$ is defined by

$$\left( -\frac{1}{c^2}\frac{\partial^2}{\partial t^2} + \nabla^2 \right) G(\mathbf{x} - \mathbf{x}', t - t') = \delta^3(\mathbf{x} - \mathbf{x}')\delta(t - t') \tag{3.93}$$

Take the time and spatial Fourier transforms of both sides of this equation:

$$\left( \frac{\omega^2}{c^2} - k^2 \right) \widetilde{G}(\mathbf{k} - \mathbf{k}', \omega - \omega') = e^{-i\mathbf{k}\cdot\mathbf{x}'} e^{i\omega t'} \quad \longrightarrow \quad \widetilde{G}(\mathbf{k}, \omega) = \frac{1}{(\omega/c)^2 - k^2} \tag{3.94}$$

We have chosen to set $\mathbf{k}' = \omega' = 0$ without loss of generality. Transforming back to real space:

$$G(\mathbf{x}, t) = \int \frac{d^3\mathbf{k}}{(2\pi)^3} \int \frac{d\omega}{2\pi} \frac{1}{(\omega/c)^2 - k^2} e^{i(\mathbf{k}\cdot\mathbf{x} - \omega t)} \tag{3.95}$$

Align our choice of coordinates along $\mathbf{x}$, such that we can write $\mathbf{k}\cdot\mathbf{x} = kr\cos\theta$ for $r = |\mathbf{x}|$. Consider the integral over $\mathbf{k}$ space:

$$\begin{aligned}\int \frac{d^3\mathbf{k}}{(2\pi)^3}\frac{1}{(\omega/c)^2 - k^2}e^{i\mathbf{k}\cdot\mathbf{x}} &= \frac{1}{(2\pi)^3}\int_0^{2\pi} d\phi \int_0^{\pi} d\theta \sin\theta \int_0^{\infty} dk\, \frac{k^2}{(\omega/c)^2 - k^2}e^{ikr\cos\theta} \\ &= \frac{1}{(2\pi)^2}\int_0^{\infty} dk\, \frac{k^2}{(\omega/c)^2 - k^2}\int_{-1}^{1} d(\cos\theta)\, e^{ikr\cos\theta} \\ &= \frac{1}{(2\pi)^2}\int_{-\infty}^{\infty} dk\, \frac{k}{ir}\frac{e^{ikr}}{(\omega/c)^2 - k^2}\end{aligned} \tag{3.96}$$

Then, our full integral expression becomes

$$G(\mathbf{x},t) = \frac{1}{(2\pi)^3}\frac{1}{ir}\int_{-\infty}^{\infty} dk\, ke^{ikr}\int_{-\infty}^{\infty} d\omega\, \frac{e^{-i\omega t}}{((\omega/c) - k)\,((\omega/c) + k)} \tag{3.97}$$

We now have to invoke some maths from complex analysis. Namely, *Cauchy's Residue Theorem* states that for a complex valued function $f(z)$ with poles at $z = z_n$, the value its integral around some closed contour $\Gamma$ is given by

$$\oint_\Gamma dz\, f(z) = 2\pi i \sum_n \mathrm{Res}(f(z = z_n)) \tag{3.98}$$

where $\mathrm{Res}(f(z = z_n))$ is a residue contained within the contour $\Gamma$. (3.97) has poles at $\omega = \pm c|k|$, along the real axis. Choosing a semicircular contour containing both poles closed in the lower half plane, we can evaluate this integral as

$$G(\mathbf{x},t) = -\frac{2\pi i}{(2\pi)^3}\frac{c^2}{ir}\int_{-\infty}^{\infty} dk\, ke^{ikr}\left[\frac{e^{-ikct}}{2kc} - \frac{e^{ikct}}{2kc}\right] = \frac{c}{4\pi r}\left[\delta(r+ct) - \delta(r-ct)\right] \tag{3.99}$$

Re-introducing $r = |\mathbf{x}|$, $\mathbf{x}'$ and $\mathbf{t}'$, we can write the Greens function for the wave equation as

$$G_\pm(\mathbf{x}-\mathbf{x}', t-t') = -\frac{c}{4\pi\,|\mathbf{x}-\mathbf{x}'|}\delta\left(|\mathbf{x}-\mathbf{x}'| \pm c\,|t-t'|\right) \tag{3.100}$$

The positive and negative solutions correspond to $t < t'$ and $t > t'$ respectively. We can now reconstruct our final solution:

$$\mathsf{A}^\nu = \int d^3\mathbf{x}' \int dt'\, G_\pm\left(-\mu_0 \mathsf{J}^\nu(\mathbf{x}', \mathbf{t}')\right) = \frac{\mu_0 c}{4\pi}\int d^3\mathbf{x}'\, \frac{\mathsf{J}^\nu\left(\mathbf{x}', t \mp |\mathbf{x}-\mathbf{x}'|/c\right)}{|\mathbf{x}-\mathbf{x}'|} \tag{3.101}$$

Define the *field event* as the space-time event $\mathsf{X}_\mathrm{f} = (ct, \mathbf{r})$ at which the fields are to be measured. Then, the *source event* $\mathsf{X}_s = (ct_s, \mathbf{r}_s)$ is the event on the past light cone of the field event where the source is located. Suppose that the source is located $\mathbf{x}'$, and occurs at $t_s = t - |\mathbf{x}-\mathbf{x}'|/c$. Then, our solution to Maxwell's equations in our covariant formalism is given by

$$\boxed{\mathsf{A}^\nu(\mathbf{r},t) = \frac{1}{4\pi\epsilon_0 c}\int d^3\mathbf{r}_s\, \frac{\mathsf{J}^\nu\left(\mathbf{r}_s, t - r_\mathrm{sf}/c\right)}{r_\mathrm{sf}}, \quad r_\mathrm{sf} = |\mathbf{r}-\mathbf{r}_s|} \tag{3.102}$$

It is clear that the four-potential is simply given by an integral over four-current source term. However, note that the source term - and hence the potentials/fields - depend on the time $t_s = t - r_\mathrm{sf}/c$. This is commonly known as the *retarded time*, which describes how information about changes in the source are only able to propagate outwards at a finite speed $c$. Thus, if we want to know the fields at some time $t$, we have to evaluate the source at this earlier, retarded time to obtain the correct fields. Note that if we have kept the positive solution in (3.101), this would have been a function of the *advanced time*, which describes wave propagation towards the source.

### 3.2.2 Potential of an Arbitrarily Moving Charge

Armed with this solution, let us now examine the particular case of a single point charge in an arbitrary state of motion. In this case, the charge serves as the source event for the fields. Consider the four-potential of the charge its instantaneous rest frame:

$$\mathsf{A}_0 = \left(\frac{q}{4\pi\epsilon_0 c}\frac{1}{r_{\mathrm{sf}}}, \mathbf{0}\right) \tag{3.103}$$

We can find the potential in any other frame by applying a Lorentz transformation with some velocity $\mathbf{v}_s$. Now, in this frame, the only quantities that can realistically contribute to the $r_{\mathrm{sf}}$ in the denominator are the four-displacement from the source event to the field event $\mathsf{X}_{\mathrm{sf}} = (ct_{\mathrm{sf}}, \mathbf{r}_{\mathrm{sf}})$, and the four-velocity of the source $\mathsf{U}_s = \gamma(c, \mathbf{v}_s)$. Their scalar product is

$$\mathsf{X}_{\mathrm{sf}} \cdot \mathsf{U}_s = \gamma(-t_{\mathrm{sf}}c^2 + \mathbf{r}_{\mathrm{sf}} \cdot \mathbf{v}_s) = \gamma(-r_{\mathrm{sf}}c + \mathbf{r}_{\mathrm{sf}} \cdot \mathbf{v}_s) \tag{3.104}$$

as $t_{\mathrm{sf}} = r_{\mathrm{sf}}/c$. This clearly evaluates to $-r_{\mathrm{sf}}c$ in the rest frame of the charge, meaning that we can write

$$\mathsf{A}_0 = \frac{q}{4\pi\epsilon_0}\frac{1}{(-\mathsf{X}_{\mathrm{sf}} \cdot \mathsf{U}_s)}(1, \mathbf{0}) \tag{3.105}$$

By inspection, we can see that the general result is given by

$$\boxed{\mathsf{A} = \frac{q}{4\pi\epsilon_0}\frac{\mathsf{U}_s/c}{(-\mathsf{X}_{\mathrm{sf}} \cdot \mathsf{U}_s)}} \tag{3.106}$$

where $\mathsf{X}_{\mathrm{sf}} = \mathsf{X}_{\mathrm{f}} - \mathsf{X}_s$ is the null four-displacement from the source event to the field event, and $\mathsf{U}_s$ is the four-velocity of the source event, as evaluated in the given frame. Remember that $\mathsf{X}_s$ and $\mathsf{U}_s$ is are both functions of the retarded time $t_s = t - r_{\mathrm{sf}}/c$.

*Consider a moving charge that is performing circular motion in the $x$-$y$ plane, with world-line given by*

$$\mathsf{X}_s = \begin{pmatrix} ct_s \\ \ell\cos\omega t_s \\ \ell\sin\omega t_s \\ 0 \end{pmatrix}$$

*Find the components of the four-potential $\mathsf{A}$ at $\mathsf{X} = (ct, \mathbf{0})$, and from this find the electric $\mathbf{E}(t)$ and magnetic $\mathbf{B}(t)$ fields at the origin.*

In this problem, we have the following four-vectors

$$\mathsf{X}_s = (ct_s, \mathbf{r}_s), \quad \mathsf{U}_s = \gamma(c, \mathbf{v}_s), \quad \mathsf{X}_{\mathrm{sf}} = \mathsf{X}_{\mathrm{f}} - \mathsf{X}_s = (\ell, -\mathbf{r}_s) \tag{3.107}$$

where we have remarked that the charge remains at a constant displacement $ct = \ell$ from the source. Evaluating the denominator of (3.106), we have that

$$-\mathsf{X}_{\mathrm{sf}} \cdot \mathsf{U}_s = -(\ell, -\mathbf{r}_s,) \cdot \gamma(c, \mathbf{v}_s) = \gamma\ell c, \quad \mathbf{r}_s \cdot \mathbf{v}_s = 0 \tag{3.108}$$

It follows that the four-potential is given by

$$\mathsf{A} = \frac{q}{4\pi\epsilon_0 c^2}\frac{1}{\ell}\begin{pmatrix} c \\ \mathbf{v}_s \end{pmatrix} \tag{3.109}$$

Now, we have that

$$\mathbf{v}_s(t_s) = \frac{d\mathbf{x}_s}{dt_s}(t_s) = \begin{pmatrix} -\ell\omega\sin\left[\omega(t-\ell/c)\right] \\ -\ell\omega\cos\left[\omega(t-\ell/c)\right] \\ 0 \end{pmatrix} \tag{3.110}$$

such that our four-potential can be written explicitly as

$$\mathsf{A} = \frac{q}{4\pi\epsilon_0 c^2}\frac{1}{\ell}\begin{pmatrix} c \\ -\ell\omega\sin\left[\omega(t-\ell/c)\right] \\ -\ell\omega\cos\left[\omega(t-\ell/c)\right] \\ 0 \end{pmatrix} \tag{3.111}$$

Now, we are interested in $\nabla_{\mathrm{f}}$ of the above expression, where the subscript indicates differentiation with respect to the field event. Now,

$$\mathsf{X}_{\mathrm{sf}} = \mathsf{X}_{\mathrm{f}} - \mathsf{X}_s \quad\longrightarrow\quad \nabla_{\mathrm{f}}\,\mathsf{X}_{\mathrm{sf}} = -\nabla_s\,\mathsf{X}_{\mathrm{sf}} \tag{3.112}$$

such that

$$\nabla_s\left(-\mathsf{X}_{\mathrm{sf}}\cdot\mathsf{U}_s\right) = \frac{d}{d\mathbf{x}_s}\left(|\mathbf{x}_s|\gamma c + \gamma\mathbf{x}_s\cdot\mathbf{u}_s\right) = \frac{\mathbf{x}_s}{|\mathbf{x}_s|}\gamma c + \gamma\mathbf{u}_s \tag{3.113}$$

where again $\mathbf{x}_s$ and $\mathbf{u}_s$ are functions of the source time $t_s = t-\ell/c$. We can now use (3.12) to find the fields.

$$\nabla\phi = -\frac{q}{4\pi\epsilon_0 c}\frac{\gamma c^2}{\left(-\mathsf{X}_{\mathrm{sf}}\cdot\mathsf{U}_s\right)^2}\nabla_{\mathrm{f}}\left(-\mathsf{X}_{\mathrm{sf}}\cdot\mathsf{U}_s\right) = \frac{q}{4\pi\epsilon_0}\frac{1}{\ell^2}\left(\widehat{\mathbf{x}}_s + \mathbf{u}_s/c\right) \tag{3.114}$$

$$\frac{\partial\mathbf{A}}{\partial t} = \frac{q}{4\pi\epsilon_0 c}\frac{1}{\gamma\ell c}\frac{\partial}{\partial t}(\gamma\mathbf{u}_s) = -\frac{q}{4\pi\epsilon_0}\frac{\Omega^2}{\ell^2}\widehat{\mathbf{x}}_s \tag{3.115}$$

where we have defined $\Omega = \omega\ell/c$. Then, the electric field is given by

$$\mathbf{E} = \frac{q}{4\pi\epsilon_0\ell^2}\left[\left(\Omega^2-1\right)\widehat{\mathbf{x}}_s(t_s) - \mathbf{u}_s(t_s)/c\right] \tag{3.116}$$

Finding the magnetic field requires a little more algebra. As $\mathbf{B} = \nabla\times\mathbf{A}$, we have that $B_i = \epsilon_{ijk}\partial_j A_k$. Now, $A_z = 0$, and $\partial_z(A_{x,y}) = 0$, so we have that $B_x = B_y = 0$. Now:

$$\partial_j A_k = \frac{q}{4\pi\epsilon_0 c}\partial_j\left(\frac{\gamma u_k}{\left(-\mathsf{X}_{\mathrm{sf}}\cdot\mathsf{U}_s\right)}\right) = \frac{q}{4\pi\epsilon_0 c^2}\frac{1}{\ell^2}(u_k\widehat{x}_j - u_k u_j/c) \tag{3.117}$$

As $B_z = \partial_x A_y - \partial y A_x$, the $\mathbf{e}_z$ component of the magnetic field is given by

$$B_z = \frac{q}{4\pi\epsilon_0 c^2}\frac{1}{\ell^2}\left[(\widehat{x}u_y - u_x u_y/c) - (\widehat{y}u_x - u_x u_y/c)\right] = \frac{\mu_0(\omega\ell)q}{4\pi\ell^2} \tag{3.118}$$

This means that we can finally write our electric and magnetic fields for this charge as:

$$\mathbf{E} = \frac{q}{4\pi\epsilon_0\ell^2}\begin{bmatrix} \left(\Omega^2-1\right)\cos\left(\omega t-\Omega\right) - \Omega\sin\left(\omega t-\Omega\right) \\ \left(\Omega^2-1\right)\sin\left(\omega t-\Omega\right) + \Omega\cos\left(\omega t-\Omega\right) \\ 0 \end{bmatrix} \tag{3.119}$$

$$\mathbf{B} = \frac{\mu_0 q}{4\pi\ell^2}\begin{bmatrix} 0 \\ 0 \\ \omega\ell \end{bmatrix} \tag{3.120}$$

The magnetic field is simply that which we would expect from using the Biot-Savart law on a point charge moving about the origin. However, the electric field is significantly more complicated, due to the finite propagation time between the source event and the field event.

### 3.2.3 Fields of an Arbitrarily Moving Charge

The scalar and vector potentials that result from (3.106) are known as the *Liénard-Wiechert potentials*, given by

$$\phi(\mathbf{r},t) = \frac{q}{4\pi\epsilon_0}\frac{1}{(r_{\text{sf}} - \mathbf{v}\cdot\mathbf{r}_{\text{sf}}/c)} \tag{3.121}$$

$$\mathbf{A}(\mathbf{r},t) = \frac{q}{4\pi\epsilon_0 c^2}\frac{\mathbf{v}}{(r_{\text{sf}} - \mathbf{v}\cdot\mathbf{v}_{\text{sf}}/c)} \tag{3.122}$$

where both $\mathbf{r}_{\text{sf}}$ and $\mathbf{v}$ are evaluated at the retarded time $t_s = t - r_{\text{sf}}/c$. To find the fields, one could use (3.12), but the algebra involved is quite laborious, and unenlightening. Instead, we shall derive the equations for the fields from tensor methods, as this provides a slightly shorter route. As we are interested in the fields, we want to find an expression for $\mathbb{F}$, meaning that we consider

$$\partial_\mu \mathsf{A}^\nu = -\frac{q}{4\pi\epsilon_0 c}\frac{\xi\partial_\mu \mathsf{U}^\nu - \mathsf{U}^\nu\partial_\mu\xi}{\xi^2} \tag{3.123}$$

where we have dropped the subscripts in (3.106), and defined $\xi = \mathsf{X}_\rho\mathsf{U}^\rho$. Let $\dot{\mathsf{U}}^\nu$ be the four-acceleration of the source event, and observe that for any quantity at the source event

$$\partial_\mu = (\partial_\mu\tau)\frac{d}{d\tau} \tag{3.124}$$

Since $\mathsf{X} = \mathsf{X}_{\text{f}} - \mathsf{X}_s$, we have

$$\partial_\mu \mathsf{X}^\nu = \delta_\mu{}^\nu - \frac{d\mathsf{X}_s{}^\nu}{d\tau}\partial_\mu\tau = \delta_\mu{}^\nu - \mathsf{U}^\nu\partial_\mu\tau \tag{3.125}$$

However, since $\mathsf{X}$ is always null, it is orthogonal to its gradient, such that

$$\mathsf{X}_\nu\partial_\mu\mathsf{X}^\nu = \mathsf{X}_\mu - \mathsf{X}_\nu\mathsf{U}^\nu\partial_\mu\tau = 0 \quad\longrightarrow\quad \partial_\mu\tau = \mathsf{X}_\mu/\xi \tag{3.126}$$

This means that we can write

$$\xi\partial_\mu\mathsf{U}^\nu = \dot{\mathsf{U}}^\nu\mathsf{X}_\mu \tag{3.127}$$

We now consider the second term in the numerator of (3.123). Using (3.125) and (3.127), we have that

$$\begin{aligned}\partial_\mu(\mathsf{X}_\nu\mathsf{U}^\nu) &= \left(\mathsf{U}_\mu - \mathsf{U}^\nu\mathsf{U}_\nu\frac{\mathsf{X}_\mu}{\xi}\right) + \mathsf{X}_\nu\dot{\mathsf{U}}^\nu\frac{\mathsf{X}_\mu}{\xi}\\ \mathsf{U}^\nu\partial_\mu\xi &= \mathsf{U}^\nu\mathsf{U}_\mu + \frac{\mathsf{X}_\mu}{\xi}\left(c^2 + \mathsf{X}_\rho\dot{\mathsf{U}}^\rho\right)\mathsf{U}^\nu\end{aligned} \tag{3.128}$$

Putting this all together, we have that

$$\partial_\mu\mathsf{A}^\nu = \frac{q}{4\pi\epsilon_0 c}\left(\frac{\mathsf{U}^\nu\mathsf{U}_\mu}{\xi^2} + \frac{c^2\mathsf{X}_\mu}{\xi^3}\bar{\mathsf{U}}^\nu\right), \quad \bar{\mathsf{U}}^\nu = \mathsf{U}^\nu - \frac{1}{c^2}(\dot{\mathsf{U}}^\nu\xi - \mathsf{U}^\nu\mathsf{X}_\rho\dot{\mathsf{U}}^\rho) \tag{3.129}$$

Substituting this into (3.85), we find that

$$\mathbb{F}^{\mu\nu} = \frac{qc}{4\pi\epsilon_0}\frac{(\mathsf{X}^\mu\bar{\mathsf{U}}^\nu - \mathsf{X}^\nu\bar{\mathsf{U}}^\mu)}{(\mathsf{X}_\rho\mathsf{U}^\rho)^3} \tag{3.130}$$

with $\bar{\mathsf{U}}$ defined as above. We can find the electric field using $E^i = c\mathbb{F}^{0i}$, yielding

$$\boxed{\mathbf{E} = \frac{q}{4\pi\epsilon_0\kappa^3}\left(\frac{\widehat{\mathbf{r}} - \mathbf{v}/c}{\gamma^3 r^2} + \frac{\widehat{\mathbf{r}}\times[(\widehat{\mathbf{r}} - \mathbf{v}/c)\times\mathbf{a}]}{c^2 r}\right), \quad \kappa = 1 - \widehat{\mathbf{r}}\cdot\mathbf{v}/c} \tag{3.131}$$

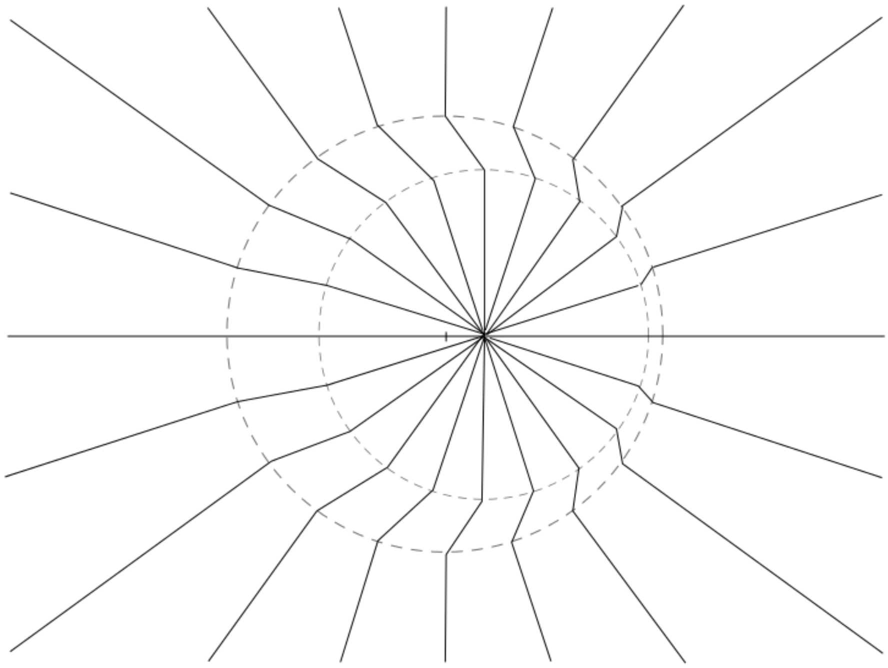


Figure 3.3: An example field of an accelerated charge

where $\mathbf{r} = \mathbf{r}_{\mathrm{f}} - \mathbf{r}_s$, and $\mathbf{v}$ and $\mathbf{a}$ are the velocity and acceleration of the charge at the source event repsectively. The magnetic field can then be found from $\mathbf{B} = \mathbf{v} \times \mathbf{E}/c^2$.

The first term is the bound field (what the field would be for a moving charge if the acceleration was zero), while the second term is the radiative field, having finite energy flux through a sphere at $r \to \infty$ (note the $r^{-1}$ dependence). The total energy of the radiative field is constant when the particle is not accelerating, and it is completely contained inside a finite region of space.

What does this field for an accelerated charge look like? Let us consider the case of a point charge that is initially in uniform motion, but then is quickly accelerated and decelerated, before moving uniformly again at the original velocity. Figure 3.3 shows the lines of the electric field in the plane containing the acceleration vector, in the initial rest frame of the charge. The two dotted circles show the current position of two light spheres that have propagated outwards from the source at the beginning and end of the acceleration respectively. Outside the outermost circle, the field lines do not have any knowledge about the acceleration, and so point towards the projected position of the charge, marked by a small cross. Inside the innermost circle, the field is simply that of a charge in uniform motion. Between the two, it is clear that the field has both a bound (along $\widehat{\mathbf{r}}$) and radiative (along $\widehat{\boldsymbol{\theta}}$) part. The faster the acceleration, the more perpendicular the field lines in this region become, and thus the greater the energy of the radiation.

## 3.3 Oscillations and Radiation

Now that we have derived general solutions to Maxwell's equations, we shall apply them to some simple, but interesting, systems involving the motion of charges. In general, we will only really need to (3.106) and (3.131), as we can build up more complicated charge distributions by considering an ensemble of point charges satisfying these equations.

### 3.3.1 Dipole Oscillations

When considering dipole oscillations, it is often convenient to consider (3.106) in the following form

$$\mathsf{A} = \frac{q}{4\pi\epsilon_0 c^2}\frac{(c, \mathbf{v})}{(r_{\text{sf}} - \mathbf{r}_{\text{sf}}\cdot\mathbf{v}/c)} \tag{3.132}$$

We approximate that the wavelength of the emitted radiation is large in comparison to the size of the dipole itself, meaning that the velocity of the charge satisfied $|\mathbf{v}| \ll c$. Under this approximation, we have that

$$\mathsf{A} \approx \frac{1}{4\pi\epsilon_0 c^2}\frac{(qc, \dot{\mathbf{d}}[t - r_{\text{sf}}/c])}{r_{\text{sf}}} \tag{3.133}$$

where we have defined $\dot{\mathbf{d}} = q\mathbf{v}$ as a function of the retarded time $t_s = t - r_{\text{sf}}/c$. Suppose that the source is sinusoidally oscillating, such that

$$\mathbf{d} = q\mathbf{x}_0 \sin(\omega t - kr_{\text{sf}}) \quad\longrightarrow\quad \dot{\mathbf{d}} = \omega q\mathbf{x}_0\cos(\omega t - kr_{\text{sf}}) \tag{3.134}$$

We make the far-field approximation that $r_{\text{sf}} \approx r$, as the movement of the source is very small in comparison to the distance of the source from where the radiation will be measured. We can thus write our magnetic vector potential as

$$\mathbf{A} \approx \frac{q\omega\mathbf{x}_0\cos(\omega t - kr)}{4\pi\epsilon_0 c^2 r} \tag{3.135}$$

The resultant magnetic field is then

$$\mathbf{B} = \frac{1}{4\pi\epsilon_0 c^2}\nabla\times\left(\frac{\dot{\mathbf{d}}[t - r/c]}{r}\right) \approx \frac{1}{4\pi\epsilon_0 c^3}\frac{\ddot{\mathbf{d}}[t - r/c]\times\mathbf{r}}{r^2} = \frac{\omega^2 q x_0}{4\pi\epsilon_0 c^3}\frac{\sin\theta}{r}\sin(kr - \omega t)\,\widehat{\boldsymbol{\phi}} \tag{3.136}$$

where we have orientated the dipole along $\mathbf{e}_z$. The second expression above follows from keeping only the far-field terms when performing the differentiation. This implies that the fields are light-like, such that $E = cB$, and $\mathbf{E}$, $\mathbf{B}$ and $\mathbf{r}$ form a right-handed set. Then, the electric field is given by

$$\mathbf{E} = c\mathbf{B}\times\mathbf{r}/r = \frac{\omega^2 q x_0}{4\pi\epsilon_0 c^2}\frac{\sin\theta}{r}\sin(kr - \omega t)\,\widehat{\boldsymbol{\theta}} \tag{3.137}$$

#### Antennae

The combination $\omega q x_0$ can be recognised as $Ix_0$, where $I$ is the size of the magnitude of the current oscillations in a short segment of wire of length $x_0$. We can thus write $d(q\omega x_0) = I(z)dz$, such that

$$\boxed{dE = \frac{I}{2\epsilon_0 c}\frac{dz}{\lambda}\frac{\sin\theta}{r}\cos(kr - \omega t)} \tag{3.138}$$

Let us model an antenna as a section of wire of length $x_0 \equiv \ell$ with a current source halfway along its length. By definition, we must have that $I(\pm\ell/2) = 0$, as there cannot be any flow through the ends of the wire. Consider the following two cases:

- Short antenna - The current is maximal in the centre, and decreases linearly towards the ends of the antenna, such that the current distribution is given by

$$I(z) = I_0 \left(1 - \frac{2|z|}{\ell}\right) \tag{3.139}$$

The relevant integral is

$$\int_{-\ell/2}^{\ell/2} dz\, I(z) = \frac{I_0 \ell}{2} \tag{3.140}$$

meaning that the resultant electric field is

$$E = \frac{I_0 \ell}{4\epsilon_0 \lambda c} \frac{\sin\theta}{r} \cos(kr - \omega t) \tag{3.141}$$

- Half-wave antenna - This is another centre fed antenna, except that it is of length $\ell = \lambda/2$, with the current distribution being

$$I(z) = I_0 \cos(kz) \tag{3.142}$$

This clearly satisfies the condition that $I(\pm\ell/2) = I(\pm\lambda/4) = 0$. As before, the relevant integral is

$$\int_{-\lambda/4}^{\lambda/4} dz\, I(z) = \frac{I_0 \lambda}{\pi} \tag{3.143}$$

meaning that the resultant electric field is given by

$$E = \frac{I_0}{4\pi\epsilon_0 c} \frac{\sin\theta}{r} \cos(kr - \omega t) \tag{3.144}$$

It is clear from these two expressions that the short antenna depends on $\lambda^{-1}$, while the half-wave antenna has no wavelength dependence. From (3.141), it appears that the radiated power is proportional to $\ell^2$. However, after $\ell = \lambda/2$, there is not a significant increase in power. Instead, the direction distribution of the field is altered due to phase lag between oscillations in the current and the far-field radiation. This means that bigger antennas are not necessarily better, above a critical size.

### 3.3.2 Radiated Power

To calculate the power radiated by moving charges, let us initially choose to evaluate quantities in the instantaneous rest frame of the charge, such that $\mathbf{v} = 0$. Then, the radiative part of (3.131) becomes

$$\mathbf{E}_{\text{rad}} = \frac{q}{4\pi\epsilon_0 c^2} \frac{\widehat{\mathbf{r}} \times (\widehat{\mathbf{r}} \times \mathbf{a})}{r} \tag{3.145}$$

We can then calculate the Poynting vector associated with this field

$$\mathbf{N} = \frac{1}{\mu_0}\mathbf{E} \times \mathbf{B} = \epsilon_0 c \mathbf{E}_{\text{rad}} \times (\widehat{\mathbf{r}} \times \mathbf{E}_{\text{rad}}) = \epsilon_0 c E_{\text{rad}}^2\, \widehat{\mathbf{r}} \tag{3.146}$$

Suppose that the radiated power is $\mathcal{P}_L$. Then, $d\mathcal{P}_L = NdS = Nr^2 d\Omega$. Performing the integration, we have that

$$\mathcal{P}_L = \frac{q^2}{4\pi\epsilon_0} \frac{a_0^2}{4\pi c^3} \int_0^{2\pi} d\phi \int_0^{\pi} d\theta\, \sin^3\theta = \frac{q^2 a_0^2}{6\pi\epsilon_0 c^3} \tag{3.147}$$

where we have used the fact that $|\dot{\mathbf{v}}| = a_0$ in the instantaneous rest frame of the particle. However, we argue that the power is in fact a Lorentz invariant quantity, and so we should be able to write it in a manifestly covariant form. In the instantaneous rest frame of the particle, no momentum is transferred into the total field. If we have that $\mathcal{P}_L = d\mathcal{E}_0/d\tau$ in this frame, we can write the four-momentum of the radiation as

$$d\mathsf{P} = (d\mathcal{E}_0/c, 0) \tag{3.148}$$

This means that the $\mathcal{E}$ transforms as $d\mathcal{E} = \gamma d\mathcal{E}_0$. Recalling (1.67), it is clear that

$$\frac{d\mathcal{E}}{dt} = \frac{d\mathcal{E}_0}{d\tau} \tag{3.149}$$

meaning that the total radiated power is the same in all frames. This is made more clear if we write (3.147) in the following form

$$\boxed{\mathcal{P}_L = \frac{q^2}{6\pi\epsilon_0 c^3}\dot{\mathsf{U}}_\mu \dot{\mathsf{U}}^\mu} \tag{3.150}$$

where the dot indicates a derivative with respect to the proper time, as per the definition of four-acceleration. This equation is known as *Larmors formula.* It is also useful to express this in terms of the velocity and acceleration of the source event. Using (2.39)

$$\mathcal{P}_L = \frac{q^2}{6\pi\epsilon_0 c^3}\gamma^6\left(a^2 - \frac{(\mathbf{v}\times\mathbf{a})^2}{c^2}\right) \tag{3.151}$$

Lastly, we note that the rate at which four-momentum is carried away by the radiation is given by

$$\frac{d\mathsf{P}}{d\tau} = \frac{\mathcal{P}_L \mathsf{U}}{c^2} = \frac{q^2}{6\pi\epsilon_0 c^5}\left(\dot{\mathsf{U}}_\mu \dot{\mathsf{U}}^\mu\right)\mathsf{U} \tag{3.152}$$

**Linear Motion**

Consider a particle that is accelerated linearly under a pure force. In this case, $\mathbf{v} \parallel \mathbf{a}$, such that the radiated power can be written as

$$\mathcal{P}_L = \frac{q^2}{6\pi\epsilon_0 c^3}\gamma^6 a^2 = \frac{q^2}{6\pi\epsilon_0 m^2 c^3}\left(\frac{dp}{dt}\right)^2 \tag{3.153}$$

where we have used (2.52). As we are assuming a pure force, we can write that

$$\mathbf{f}\cdot\mathbf{v} = \frac{dE}{dt} \quad \longrightarrow \quad \frac{dp}{dt} = \frac{1}{v}\frac{dE}{dt} = \frac{dE}{dx} \tag{3.154}$$

This allows us to write the radiated power is a useful form

$$\frac{\mathcal{P}_L}{dE/dt} = \frac{q^2}{6\pi\epsilon_0 m^2 c^3}\frac{1}{v}\frac{dE}{dx} \xrightarrow{v\backsim c} \frac{2}{3}\left(2.8\times 10^{-15}\,\mathrm{m}\right)\frac{1}{mc^2}\frac{dE}{dx} \tag{3.155}$$

Typically, we are accelerating electrons, such that $mc^2 \backsim 0.511$ MeV and $dE/dx \backsim 10^9$ Vm$^{-1}$. This means that the relative radiative loss given by the above equation is of order $\backsim 10^{-12}$, such that radiated power in linear accelerators is negligible.

**Circular Motion**

In the case of circular motion under a pure force, $\mathbf{v} \perp \mathbf{a}$, so

$$\mathcal{P}_L = \frac{q^2}{6\pi\epsilon_0 c^3}\gamma^6\left(a^2 - \frac{a^2v^2}{c^2}\right) = \frac{q^2}{6\pi\epsilon_0 c^3}\frac{\gamma^4 v^4}{r^2} \tag{3.156}$$

The radiative energy loss per revolution is then

$$\Delta E = \frac{2\pi r}{v}\mathcal{P}_L = \frac{q^2}{3\epsilon_0 c^3}\frac{\gamma^4 v^3}{r} \tag{3.157}$$

Suppose that an electron moves in a magnetic field of 1 T in an orbit of radius 10 m. From (2.52), we have that

$$f_\perp = \gamma m a = qBv \longrightarrow v = \frac{qBr}{\gamma m} = \frac{v_0}{\gamma} \longrightarrow v = \frac{v_0}{\sqrt{1+(v_0/c)^2}} \approx 2.9999{\times}10^8\mathrm{ms}^{-1} \tag{3.158}$$

This means that the energy is of order $E \approx 3\,\mathrm{GeV}$, while $\Delta E \approx 0.717\,\mathrm{MeV}$. This means that radiative losses cannot be ignored in circular accelerators; a very large radius is needed if losses are to be small. The losses are so much higher in this case in comparison to the linear case due to the difference in $\gamma$'s between the two expressions in (2.52); there is a lower inertial resistance to acceleration perpendicular to the motion.

## 3.4 Lagrangian Mechanics

Field theories are often expressed through the language of Lagrangian mechanics. In this section, we will upgrade our classical Lagrangian theory to be able to handle relativistic fields, and examine the specific case of the motion of a charge particle.

### 3.4.1 Classical Lagrangian

Suppose that we have a system consisting of $N$ particles indexed by $i$, each with generalised coordinates $\mathbf{q}_i(t)$ and $\dot{\mathbf{q}}_i(t)$. Then, the Lagrangian of the system is a function defined by

$$\boxed{\mathcal{L} = \mathcal{L}\left(\{\mathbf{q}_i\}, \{\dot{\mathbf{q}}_i\}, t\right) = T - V} \tag{3.159}$$

where $T$ and $V$ are the total kinetic and potential energies of the system respectively. Note that from here onwards we shall use $\mathbf{q}_i$ and $\dot{\mathbf{q}}_i$ as shorthand for $\{\mathbf{q}_i\}$ and $\{\dot{\mathbf{q}}_i\}$ respectively. The *action* is the time integral of the Lagrangian:

$$\mathcal{S} = \int_{t_1}^{t_2} dt\, \mathcal{L}(\mathbf{q}_i, \dot{\mathbf{q}}_i, t) \tag{3.160}$$

*Hamilton's principle* states that the phase space trajectory taken by a set of particles from $\{\mathbf{q}_i(t_1), \dot{\mathbf{q}}_i(t_1)\}$ to $\{\mathbf{q}_i(t_2), \dot{\mathbf{q}}_i(t_2)\}$ is the one for which $\mathcal{S}$ is stationary with respect to variations in $\mathcal{L}$. We write that

$$\delta\mathcal{L} = \mathcal{L}(\mathbf{q}_i + \delta\mathbf{q}_i, \dot{\mathbf{q}}_i + \delta\dot{\mathbf{q}}_i, t) - \mathcal{L}(\mathbf{q}_i, \dot{\mathbf{q}}_i, t) = \frac{\partial\mathcal{L}}{\partial\mathbf{q}_i}\cdot\delta\mathbf{q}_i + \frac{\partial\mathcal{L}}{\partial\dot{\mathbf{q}}_i}\cdot\delta\dot{\mathbf{q}}_i \tag{3.161}$$

Then, the variation in the action is given by

$$\delta S = \int_{t_1}^{t_2} dt\, \left[\frac{\partial\mathcal{L}}{\partial\mathbf{q}_i}\cdot\delta\mathbf{q}_i + \frac{\partial\mathcal{L}}{\partial\dot{\mathbf{q}}_i}\cdot\delta\dot{\mathbf{q}}_i\right] = \int_{t_1}^{t_2} dt\, \delta\mathbf{q}_i \left[\frac{\partial\mathcal{L}}{\partial\mathbf{q}_i} - \frac{d}{dt}\left(\frac{\partial\mathcal{L}}{\partial\dot{\mathbf{q}}_i}\right)\right] = 0 \tag{3.162}$$

where the second expression follows from integrating by parts, and recognising a total derivative. As this equation must hold for all coordinates $\mathbf{q}_i$, we arrive at the *Euler-Lagrange equations*:

$$\boxed{\frac{d}{dt}\left(\frac{\partial\mathcal{L}}{\partial\dot{\mathbf{q}}_i}\right) - \frac{\partial\mathcal{L}}{\partial\mathbf{q}_i} = 0} \tag{3.163}$$

This relationship between $\mathbf{q}_i$ and $\dot{\mathbf{q}}_i$ can be seen as a more general version of Newton's Second Law; we have the rate of change of some generalised momenta $\mathbf{p}_i = \partial\mathcal{L}/\partial\dot{\mathbf{q}}_i$ equal to some generalised force $\partial\mathcal{L}/\partial\mathbf{q}_i$. If the Lagrangian is independent (explicitly) of $\mathbf{q}_i$, this coordinate is said to be *cyclic*, with a corresponding *conserved quantity* $\partial\mathcal{L}/\partial\dot{\mathbf{q}}_i$.

Note that for some general Euclidean metric $g_{ij}$, the Lagrangian can be written explicitly as

$$\mathcal{L} = \frac{1}{2}mg_{ij}\dot{q}^i\dot{q}^j - V(q^i, q^j) \tag{3.164}$$

where we have adopted the Einstein summation convention.

#### Hamiltonian Mechanics

Further to this, we can define the classical Hamiltonian $\mathcal{H}$ as

$$\boxed{\mathcal{H} = \mathcal{H}(\mathbf{q}_i, \mathbf{p}_i, t) = \sum_i \dot{\mathbf{q}}_i\mathbf{p}_i - \mathcal{L}} \tag{3.165}$$

This is a Legendre transform of the Lagrangian. Typically, $\mathcal{H}$ is often found to correspond to the energy of the system. Consider the total differential of the Lagrangian

$$d\mathcal{L} = \frac{\partial \mathcal{L}}{\partial \mathbf{q}_i} \cdot d\mathbf{q}_i + \frac{\partial \mathcal{L}}{\partial \dot{\mathbf{q}}_i} \cdot d\dot{\mathbf{q}}_i + \frac{\partial \mathcal{L}}{\partial t} dt = \frac{\partial \mathcal{L}}{\partial \mathbf{q}_i} \cdot d\mathbf{q}_i + \mathbf{p}_i \cdot d\dot{\mathbf{q}}_i + \frac{\partial \mathcal{L}}{\partial t} dt \tag{3.166}$$

where we have used the definition of $\mathbf{p}_i$, and adopted the implied summation over the $\mathbf{q}_i$'s. We can re-write this as

$$d\mathcal{L} = \left( \frac{\partial \mathcal{L}}{\partial \mathbf{q}_i} \cdot d\mathbf{q}_i + d(\mathbf{p}_i \cdot \dot{\mathbf{q}}_i) - \dot{\mathbf{q}}_i \cdot d\mathbf{p}_i \right) + \frac{\partial \mathcal{L}}{\partial t} dt \tag{3.167}$$

Then,

$$d\mathcal{H} = d\left(\dot{\mathbf{q}}_i \mathbf{p}_i - \mathcal{L}\right) = -\frac{\partial \mathcal{L}}{\partial \mathbf{q}_i} \cdot d\mathbf{q}_i + \dot{\mathbf{q}}_i \cdot d\mathbf{p}_i - \frac{\partial \mathcal{L}}{\partial t} dt \tag{3.168}$$

We can also write the total differential of the Hamiltonian as

$$d\mathcal{H} = \frac{\partial \mathcal{H}}{\partial \mathbf{q}_i} \cdot d\mathbf{q}_i + \frac{\partial \mathcal{H}}{\partial \mathbf{p}_i} \cdot d\mathbf{p}_i + \frac{\partial \mathcal{H}}{\partial t} dt \tag{3.169}$$

Comparison of (3.168) and (3.169) yields

$$\boxed{\frac{\partial \mathcal{H}}{\partial \mathbf{q}_i} = -\dot{\mathbf{p}}_i, \quad \frac{\partial \mathcal{H}}{\partial \mathbf{p}_i} = \dot{\mathbf{q}}_i, \quad \frac{\partial \mathcal{H}}{\partial t} = -\frac{\partial \mathcal{L}}{\partial t}} \tag{3.170}$$

These are known as *Hamilton's equations*. Note that the last of these implies that for a Lagrangian with no explicit time dependence, the corresponding Hamiltonian is conserved.

### Motion in Electromagnetic Fields

In Special Relativity, a free particle must satisfy the following equation of motion

$$\frac{d}{dt}(\gamma m \dot{\mathbf{x}}) = 0 \tag{3.171}$$

where the dot indicates differentiation with respect to the local frame time $t$. By inspection, it is clear that the Lagrangian that will give this equation of motion is

$$\boxed{\mathcal{L}_{\text{free}} = -\frac{mc^2}{\gamma}} \tag{3.172}$$

(Try differentiating the above expression with respect to $\mathbf{v} = \dot{\mathbf{x}}$ if you are not convinced). One can also show that the Lagrangian corresponding to the interaction of a charged particle with the electromagnetic field is

$$\boxed{\mathcal{L}_{\text{int}} = q(-\phi + \mathbf{v} \cdot \mathbf{A})} \tag{3.173}$$

Substituting $\mathcal{L} = \mathcal{L}_{\text{int}} + \mathcal{L}_{\text{int}}$ into (3.163), and re-arranging, we arrive at

$$\frac{d}{dt}(\gamma m \mathbf{v}) = q\left(-\nabla\phi + \nabla(\mathbf{v} \cdot \mathbf{A}) - \frac{d\mathbf{A}}{dt}\right), \quad \mathbf{p} = \gamma m \mathbf{v} + q\mathbf{A} \tag{3.174}$$

This is the equation of motion of a particle in electromagnetic fields. However, we already know what this should be; the right-hand side of the first expression above should be equal to the Lorentz force (3.34). Remarking that

$$\frac{d\mathbf{A}}{dt} = \frac{\partial \mathbf{A}}{\partial t} + (\mathbf{v} \cdot \nabla)\mathbf{A} \tag{3.175}$$

we can write

$$q\left(-\nabla\phi + \nabla(\mathbf{v}\cdot\mathbf{A}) - \frac{d\mathbf{A}}{dt}\right) = -q\left(\nabla\phi + \frac{\partial\mathbf{A}}{\partial t}\right) + q\left(\nabla(\mathbf{v}\cdot\mathbf{A}) - (\mathbf{v}\cdot\nabla)\mathbf{A}\right) \quad (3.176)$$
$$= -q\left(\nabla\phi + \frac{\partial\mathbf{A}}{\partial t}\right) + q(\mathbf{v}\times(\nabla\times\mathbf{A}))$$

which from (3.12) is clearly the Lorentz force. Thus, the system behaves as we would expect. The Hamiltonian is given by

$$\mathcal{H} = \mathbf{p}\cdot\mathbf{v} + \frac{mc^2}{\gamma} + q\phi - q\mathbf{v}\cdot\mathbf{A} = \gamma mc^2 + q\phi \quad (3.177)$$

However, this is not expressed as a function of the natural Hamiltonian variables $\mathbf{q}_i$ and $\mathbf{p}_i$. As $\mathbf{p} = \gamma m\mathbf{v} + q\mathbf{A}$, we can write that $\gamma m\mathbf{v} = \mathbf{p} - q\mathbf{A}$. Then

$$\gamma mc^2 = E = \left((\gamma m\mathbf{v})^2c^2 + m^2c^4\right)^{1/2} = \left((\mathbf{p} - q\mathbf{A})^2 c^2 + m^2c^4\right)^{1/2} \quad (3.178)$$

meaning that the Hamiltonian for the motion of a charged particle in electromagnetic fields is given by

$$\mathcal{H} = \left((\mathbf{p} - q\mathbf{A})^2 c^2 + m^2c^4\right)^{1/2} + q\phi \quad (3.179)$$

### 3.4.2 Relativistic Lagrangian

Consider some relativistic field $\Phi$, which may be a tensor of any rank. Then, the general *Lagrangian density* is given by

$$\mathcal{L} = \mathcal{L}(\Phi, \partial_\mu\Phi, \tau) \quad (3.180)$$

In a similar way to the classical case, we define the action by

$$\mathcal{S} = \int d^4\mathbf{x}\, \mathcal{L}(\Phi, \partial_\mu\Phi, \tau) \quad (3.181)$$

where the integral is over our entire four-dimensional space. Then, once again invoking Hamilton's principle:

$$\delta\mathcal{S} = \int d^4\mathbf{x}\left[\frac{\partial\mathcal{L}}{\partial\Phi}d\Phi + \frac{\partial\mathcal{L}}{\partial(\partial^\mu\Phi)}d(\partial_\mu\Phi)\right] = \int d^4\mathbf{x}\, d\Phi\left[\frac{\partial\mathcal{L}}{\partial\Phi} - \partial_\mu\left(\frac{\partial\mathcal{L}}{\partial(\partial^\mu\Phi)}\right)\right] = 0 \quad (3.182)$$

where again we have integrated by parts, and recognised a total differential. Then, we obtain the relativistically valid Euler-Lagrange equations.

$$\boxed{\partial_\mu\left(\frac{\partial\mathcal{L}}{\partial(\partial^\mu\Phi)}\right) = \frac{\partial\mathcal{L}}{\partial\Phi}} \quad (3.183)$$

*Suppose that $\Phi = A^\nu$. Consider the Lagrangian density*

$$\mathcal{L} = -\frac{1}{4}\mathbb{F}^{\mu\nu}\mathbb{F}_{\mu\nu}$$

*Show that the equations of motion are of the form of the wave equation, and show that $\mathsf{A} = \mathsf{C}\sin(\mathsf{X}_\rho\mathsf{K}^\rho)$ ($\mathsf{C}$, $\mathsf{K}$ constant four-vectors) is a solution to this equation. Calculate the stress energy tensor $\mathsf{T}$ for this form of solution. In the case that $\mathsf{K} = k(1,0,0,1)$, find the associated electric and magnetic fields.*

This is intended as an illustrative example to give the reader an idea of how to put this machinery into practise. Recalling (3.85), we can write the Lagrangian explicitly as

$$\mathcal{L} = -\frac{1}{4}\mathbb{F}^{\mu\nu}\mathbb{F}_{\mu\nu} = -\frac{1}{4}(\partial^\mu \mathsf{A}^\nu - \partial^\nu \mathsf{A}^\mu)(\partial_\mu \mathsf{A}_\nu - \partial_\nu \mathsf{A}_\mu) \tag{3.184}$$

Performing the differentiation:

$$\frac{\partial \mathcal{L}}{\partial \mathsf{A}^\nu} = 0 \tag{3.185}$$

$$\frac{\partial \mathcal{L}}{\partial(\partial^\mu \mathsf{A}^\nu)} = -(\partial_\mu \mathsf{A}_\nu - \partial_\nu \mathsf{A}_\mu) \tag{3.186}$$

meaning that the Euler-Lagrange equation becomes

$$\partial_\nu(\partial^\mu \mathsf{A}_\mu) - (\partial_\mu \partial^\mu)\mathsf{A}_\nu = 0 \tag{3.187}$$

This is clearly (3.91), which is a wave equation for $\mathsf{A}$. For simplicity, we choose to adopt the Lorentz gauge $\partial^\mu \mathsf{A}_\mu = 0$. Substituting the trial solution $\mathsf{A} = \mathsf{C}\sin(\mathsf{X}_\rho \mathsf{K}^\rho)$ into the resultant wave equation, and the gauge condition allows us to obtain the follow conditions on $\mathsf{C}$ and $\mathsf{K}$:

$$\partial_\mu \mathsf{A}^\mu = 0 \quad \longrightarrow \quad \mathsf{C}_\mu \mathsf{K}^\mu = 0 \tag{3.188}$$

$$\partial_\mu \partial^\mu \mathsf{A}^\nu = 0 \quad \longrightarrow \quad \mathsf{K}_\mu \mathsf{K}^\mu = 0 \tag{3.189}$$

The last of these conditions means that $\mathsf{K}$ is a null four-vector, meaning that the solution must correspond to the propagation of light. Using this trial solution, the field tensor becomes

$$\mathbb{F}^{\mu\nu} = (\mathsf{K}^\mu \mathsf{C}^\nu - \mathsf{K}^\nu \mathsf{C}^\mu)\cos(\mathsf{X}_\rho \mathsf{K}^\rho) \tag{3.190}$$

We now need to calculate the expressions from (3.75):

$$\begin{aligned} -\mathbb{F}^{a\mu}\mathbb{F}^b_\mu &= -(\mathsf{K}^a \mathsf{C}^\mu - \mathsf{K}^\mu \mathsf{C}^a)(\mathsf{K}_\mu \mathsf{C}^b - \mathsf{K}^b \mathsf{C}_\mu)\cos^2(\mathsf{X}_\rho \mathsf{K}^\rho) \\ &= (\mathsf{K}^a \mathsf{K}^b \mathsf{C}^\mu \mathsf{C}_\mu)\cos^2(\mathsf{X}_\rho \mathsf{K}^\rho) \end{aligned} \tag{3.191}$$

$$\begin{aligned} \mathbb{F}^{\mu\nu}\mathbb{F}_{\mu\nu} &= (\mathsf{K}^\mu \mathsf{C}^\nu - \mathsf{K}^\nu \mathsf{C}^\mu)(\mathsf{K}_\mu \mathsf{C}_\nu - \mathsf{K}_\nu \mathsf{C}_\mu)\cos^2(\mathsf{X}_\rho \mathsf{K}^\rho) \\ &= 0 \end{aligned} \tag{3.192}$$

where we have made use of the identities (3.188) and (3.189). This means that our expression for the stress energy tensor is

$$\mathbb{T}^{ab} = \epsilon_0 c^2 (\mathsf{K}^a \mathsf{K}^b \mathsf{C}^\mu \mathsf{C}_\mu)\cos^2(\mathsf{X}_\rho \mathsf{K}^\rho) \tag{3.193}$$

which is clearly the expression for a plane electromagnetic wave (3.83) that we encountered when discussing introducing the stress energy tensor. For $\mathsf{K} = k(1,0,0,1)$, we must have that $\mathsf{C} = a(0,1,0,0)$. Then, we can use (3.190):

$$\mathbb{F}^{\mu\nu} = ka \begin{pmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} \cos(\mathsf{X}_\rho \mathsf{K}^\rho) \tag{3.194}$$

Comparison with (3.39) gives the electric and magnetic fields as $\mathbf{E} = kac(1,0,0)\cos(kz - \omega t)$ and $\mathbf{B} = ka(0,1,0)\cos(kz - \omega t)$, where we have used the fact that $\mathbf{k} = (0,0,1)$.

# 4. *Spinor Fields*

This chapter aims to cover the basic concepts behind spinor fields, including

- An introduction to Spinors
- The Klein-Gordan Equation

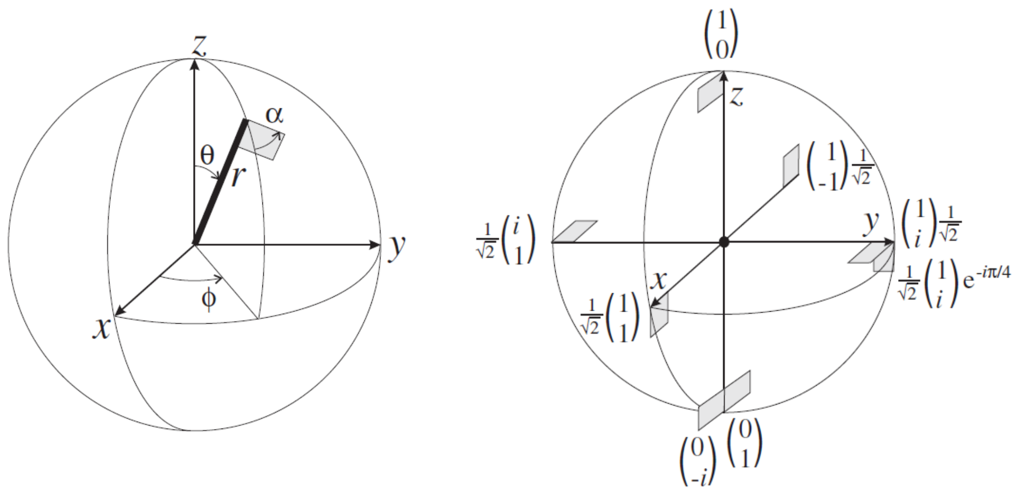
This chapter may seem a little like an after-thought, as the material within it seems quite disconnected from the rest of the course. It has simply been included here for the sake of satisfying the syllabus, and it does not delve into the rich material that is associated with spinors, such as the Dirac equation.

As they will be relevant throughout this chapter, we include the Pauli matrices here for reference:

$$\sigma^x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma^y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma^z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad {\sigma^j}^2 = I$$

## 4.1 An Introduction to Spinors

Spinors are tensor-like objects that are members of the proper Lorentz group. For every tensor of rank-$r$ within this group, there is a corresponding spinor of rank-$2r$. For example, a general four-vector would correspond to a Hermitian spinor of rank-2, which can be represented as a $2 \times 2$ Hermitian matrix of complex numbers.




(a) (b)

Figure 4.1: Spinors (a) Coordinates used to specify the spinor (b) Some examples of basic spinors

We define a 1st-rank spinor $\mathbf{s}$ as a pair of complex numbers

$$\mathbf{s} = (a, b)^T = s e^{i\alpha/2} \begin{pmatrix} \cos(\theta/2) e^{-i\phi/2} \\ \sin(\theta/2) e^{i\phi/2} \end{pmatrix} \tag{4.1}$$

This has a direction is space (a ‘flagpole’), and orientation about this axis (a ‘flag’), and an overall sign. This means that the set of parameters $(r, \theta, \phi, \alpha)$ is enough to uniquely describe the spinor state, up to a sign. The first three fix the length and direction of the flagpole by

$$r = s^2 = |a|^2 + |b|^2, \quad \mathbf{r} = \begin{pmatrix} r\sin\theta\cos\phi \\ r\sin\theta\sin\phi \\ r\cos\theta \end{pmatrix} = \begin{pmatrix} ab^* + a^*b \\ i(ab^* - a^*b) \\ |a|^2 - |b|^2 \end{pmatrix} \tag{4.2}$$

where the second expression for $\mathbf{r}$ has been obtained by inverting the definition (4.1). If we define the vector of Pauli matrices $\boldsymbol{\sigma} = (\sigma_x, \sigma_y, \sigma_z)$, then the flagpole vector can clearly be written as

$$\mathbf{r} = \langle \mathbf{s} | \boldsymbol{\sigma} | \mathbf{s} \rangle \tag{4.3}$$

### 4.1.1 Spinors and Four-vectors

Suppose that we have some four-vector $\mathsf{U}^\mu = (t, x, y, z)$. With this, we associate the matrix

$$U = \mathbf{s}\mathbf{s}^\dagger = \begin{pmatrix} |a|^2 & ab^* \\ a^*b & |b|^2 \end{pmatrix} \tag{4.4}$$

It is clear from the above expression that $U = U^\dagger \longrightarrow UU^\dagger = I$, meaning that $U$ is Hermitian. We can write any Hermitian matrix in the form

$$U = \begin{pmatrix} t+z & x-iy \\ x+iy & t-z \end{pmatrix} \tag{4.5}$$

where $t, x, y, z$ are all real scalars. Comparing (4.4) and (4.5), we can see that

$$\begin{pmatrix} t \\ x \\ y \\ z \end{pmatrix} = \frac{1}{2} \begin{pmatrix} |a|^2 + |b|^2 \\ ab^* + a^*b \\ i(ab^* - a^*b) \\ |a|^2 - |b|^2 \end{pmatrix} = \mathsf{U}^\mu \tag{4.6}$$

Defining $\sigma^\mu = (I, \boldsymbol{\sigma})$, it is clear that the four-vector associated with our spinor $\mathbf{s}$ can be written as

$$\boxed{\mathsf{U}^\mu = \frac{1}{2} \langle \mathbf{s} | \sigma^\mu | \mathbf{s} \rangle = \frac{1}{2} \mathbf{s}^\dagger \sigma^\mu \mathbf{s}} \tag{4.7}$$

Note that the inclusion of the factor of a 1/2 is a matter of convention; it does not change the calculations if it is neglected. It is clear from the explicit form of $\mathsf{U}^\mu$ in (4.6) that this is a null four-vector. We note that the determinant of $U$ also gives the invariant associated with this four-vector; it is clear from (4.4) that this is null.

### 4.1.2 Transformation of Spinors

Like other members of the Lorentz group, spinors transform linearly under both rotations and Lorentz transformations, except that the resulting spinor depends on the sequence of operations that was performed. Unlike vectors and tensors, a spinor transforms to its negative when the space is rotated through 0 to $2\pi$; this is a characteristic property of spinors.

Under some frame transformation, a general 1st-rank spinor $\mathbf{s}$ will transform according to

$$\mathbf{s}' = \Lambda \mathbf{s} \tag{4.8}$$

where $\Lambda$ is some $2 \times 2$ matrix. Consider the transformation $U' = \Lambda U \Lambda^\dagger$. To keep the determinant (i.e the invariant) unchanged, we must have that

$$\det(\Lambda) \det\left(\Lambda^\dagger\right) = 1 \quad \longrightarrow \quad \det(\Lambda) = e^{i\lambda} \tag{4.9}$$

We choose the case of $\lambda = 0$, meaning that we are restricted to the group of complex $2 \times 2$ matrices with unit determinant. The most general matrix satisfying such conditions is

$$\boxed{\Lambda = \exp\left(i\boldsymbol{\sigma} \cdot \boldsymbol{\theta}/2 - \boldsymbol{\sigma} \cdot \boldsymbol{\rho}/2\right)} \tag{4.10}$$

where $\rho$ is the rapidity, and $\theta$ is the rotation angle. $\Lambda$ is unitary if the transformation is a rotation in space; if $\Lambda$ is Hermitian then it corresponds to a boost. Using Taylor expansions, we can write these transformations in the more useful forms of

$$e^{i(\theta/2)\sigma^\mu} = \cos(\theta/2) I + i \sin(\theta/2) \sigma^\mu \tag{4.11}$$

$$e^{-(\rho/2)\sigma^\mu} = \cosh(\rho/2) I - \sinh(\rho/2) \sigma^\mu \tag{4.12}$$

Note that it often helps to further simplify these by writing them in their $2 \times 2$ matrix form. If one contracts a spinor into a four-vector as described in the previous section, the transformation of the spinor will contract to the transformation of the four-vector. This means that we can find the transformations of physical quantities by representing them as spinors, and finding the transformation of the spinor, before transforming back using

$$\mathsf{U}'^{\mu} = \frac{1}{2}\left\langle \mathbf{s}'\right| \sigma^{\mu} \left|\mathbf{s}'\right\rangle \tag{4.13}$$

which is often conveniently done component wise. An example of such a calculation is performed in 4.1.4.

### 4.1.3 Chirality

The concept of *chirality* concerns the way in which these spinors transform, and is very similar to the idea of contravariant and covariant four-vectors. Specifically, we have two different types of spinors:

$$\text{Contraspinor:} \quad \mathbf{s}'_R = \Lambda \mathbf{s}_R \tag{4.14}$$

$$\text{Cospinor:} \quad \mathbf{s}'_L = \left(\Lambda^{\dagger}\right)^{-1} \mathbf{s}_L \tag{4.15}$$

These are often referred to as being right and left handed respectively. When contracted with $\sigma^{\mu}$, the former produces a contravariant four-vector, while the latter produces a covariant four-vector, as we would anticipate from these transformations. The idea is that we regard the transformation $\Lambda$ as a kind of rotation; for a given $\rho$ and $\theta$, the contraspinor 'rotates' on way, while the cospinor 'rotates' the other way. They are thus said to posses opposite chirality.

### 4.1.4 A Worked Example

*A two-component spinor* $\mathbf{s}$ *may be used to represent the four-momentum* $\mathsf{P}$ *of a massless particle by using the relationship* $\mathsf{P}^{\mu} = \mathbf{s}^{\dagger}\sigma^{\mu}\mathbf{s}$*, where* $\sigma^{\mu}$ *are the Pauli spin matrices. Obtain the four-momentum for*

$$\mathbf{s}_1 = \begin{pmatrix}1\\0\end{pmatrix}, \quad \mathbf{s}_2 = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\1\end{pmatrix}, \quad \mathbf{s}_3 = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\i\end{pmatrix}, \quad \mathbf{s}_4 = \begin{pmatrix}e^{i|\alpha|}\\0\end{pmatrix}$$

*Starting from a spinor orientated along* $\mathbf{e}_z$*, and applying suitable rotations, confirm the answers obtain for the above cases. Then, by applying a rotation to the previous result, find a spinor representing a particle whose energy is* $E$ *and whose momentum is along* $(1,1,0)$ *in some frame* $S$*. Find the spinor components for this same particle when it is observed in a frame* $S'$ *moving at a relative velocity* $(15/17)c$ *to* $S$ *along* $\mathbf{e}_z$*. Hence obtain the four-momentum, and use it to obtain the angle between the particle velocity and* $\mathbf{e}_z$ *in the new frame.*

Recalling the definition $\mathbf{s} = (a,b)^T$ and - noting the factor of two difference -(4.6), it follows that

$$\mathsf{P}^0 = |a|^2 + |b|^2, \quad \mathsf{P}^x = a^*b + ab^*, \quad \mathsf{P}^y = i(ab^* - ba^*), \quad P^z = |a|^2 - |b|^2 \tag{4.16}$$

From it is easy to read off that

$$\mathsf{P}_1 = (1,0,0,1), \quad \mathsf{P}_2 = (1,1,0,0), \quad \mathsf{P}_3 = (1,0,1,0), \quad \mathsf{P}_4 = (1,0,0,1) \tag{4.17}$$

To rotate spinors, we make use of (4.11). Now, $\mathbf{s}_1 \equiv \mathbf{s}_z = (1,0)^T$ is already orientated along $\mathbf{e}_z$. $\mathbf{s}_2$ is orientated along $\mathbf{e}_x$, so we rotate by $-\pi/2$ around $\mathbf{e}_y$:

$$\mathbf{s} = e^{i(-\pi/4)\sigma^y}\mathbf{s}_1 = \left[\begin{pmatrix}1&0\\0&1\end{pmatrix}\cos(\pi/4) - i\begin{pmatrix}0&-i\\i&0\end{pmatrix}\sin(\pi/4)\right]\begin{pmatrix}1\\0\end{pmatrix} = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\1\end{pmatrix} = \mathbf{s}_2 \tag{4.18}$$

Similarly, $\mathbf{s}_3$ is orientated along $\mathbf{e}_y$, so rotate the coordinates by $\pi/2$ around $\mathbf{e}_x$:

$$\mathbf{s} = e^{i(\pi/4)\sigma^x}\mathbf{s}_1 = \left[\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}\cos(\pi/4) + i\begin{pmatrix}0 & 1\\ i & 0\end{pmatrix}\sin(\pi/4)\right]\begin{pmatrix}1\\0\end{pmatrix} = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\1\end{pmatrix} = \mathbf{s}_3 \tag{4.19}$$

The normalisation for a spinor of energy $\mathbf{E}$ is simply $\sqrt{E}$ times the relevant spinor. Start with $\mathbf{s}_2$, and rotate the coordinates around $\mathbf{e}_z$ by $-\pi/4$ to obtain a spinor along $(1,1,0)$:

$$\mathbf{s}_{\text{new}} = \left[\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}\cos(\pi/8) - i\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}\sin(\pi/4)\right]\frac{1}{\sqrt{2}}\begin{pmatrix}1\\1\end{pmatrix} = \frac{1}{\sqrt{2}}\begin{pmatrix}e^{-i\pi/8}\\ e^{i\pi/8}\end{pmatrix} \tag{4.20}$$

This means that the required spinor is given by

$$\mathbf{s}_E = \sqrt{\frac{E}{2}}\begin{pmatrix}1\\ e^{i\pi/4}\end{pmatrix} \tag{4.21}$$

Now, in order to perform the boost, we need to find an expression for $\rho$. For this, we use (2.65):

$$\tanh\rho = \beta \quad \longrightarrow \quad e^{\rho} = \sqrt{\frac{1+\beta}{1-\beta}} = \sqrt{\frac{32}{2}} = 4 \tag{4.22}$$

The spinor in the new frame is then given by

$$\mathbf{s}'_E = e^{-(\rho/2)\sigma^z}\mathbf{s}_E = \sqrt{\frac{E}{2}}\begin{pmatrix}e^{-\rho/2} & 0\\ 0 & e^{\rho/2}\end{pmatrix}\begin{pmatrix}1\\ e^{i\pi/4}\end{pmatrix} = \sqrt{\frac{E}{2}}\begin{pmatrix}1/2\\ 2e^{i\pi/4}\end{pmatrix} \tag{4.23}$$

Then, the components of the new four-momentum are given by

$$\mathsf{P}'^0 = E' = \frac{17}{8}E, \quad \mathsf{P}'^x = \frac{E}{\sqrt{2}}, \quad \mathsf{P}'^y = \frac{E}{\sqrt{2}}, \quad \mathsf{P}'^z = -\frac{15}{8}E \tag{4.24}$$

As this four-vector is null, we have that $\mathsf{P}^0 = \left|\mathsf{P}^i\right|^2$, allowing us to write that $\mathsf{P}'^0\cos\theta = \mathsf{P}'^z$. Then, the angle $\theta$ that we are interested in is given by

$$\theta = \cos^{-1}(\mathsf{P}^z/\mathsf{P}^0) = \cos^{-1}(-15/17) \approx 151.9^\circ \tag{4.25}$$

## 4.2 The Klein-Gordan Equation

To construct a relativistic version of quantum mechanics, we need a new fundamental equation to replace the Time-Dependant Schrödinger Equation that has some spatial and time dependence. The form of this equation cannot be derived, much as the Schrödinger equation cannot be derived from anything more fundamental. It can, however, be motivated from classical considerations. For example, in the case of the Schrödinger equation, we can demand that we obtain the energy dispersion relation

$$E = \frac{p^2}{2m} \tag{4.26}$$

Replacing the physical observables with the associated operators $E \to i\hbar\partial_t$, $p_i \to -i\hbar\partial_i$, we obtain the Schrödinger equation for a free particle.

This gives an obvious strategy for deducing a possible form for a relativistic version of the Schrödinger equation for a free particle. The relativistic analogue of (4.26) is

$$E = p^2c^2 + m^2c^4 \tag{4.27}$$

If we again use the identification $E \to i\hbar\partial_t$, $p_i \to -i\hbar\partial_i$, and let these operators act on some function $\phi(\mathbf{x}, t)$, we obtain

$$-\hbar^2\partial_t^2\phi = -c^2\hbar^2\nabla^2\phi + m^2c^4\phi \tag{4.28}$$

This can clearly be written in a manifestly covariant form as

$$\boxed{\left(-\partial^\mu\partial_\mu + \frac{m^2c^2}{\hbar^2}\right)\phi = 0} \tag{4.29}$$

This is known as the *Klein-Gordan Equation* (KGE), though it is actually due to Schrödinger. Note that the notation $\mu = mc/\hbar$ is often used.

Let us solve the KGE for a simple case. Assume that we have a spherically symmetric field $\phi = \phi(r, t)$, and that we are in the low frequency limit, such that time derivatives have a magnitude much smaller than spatial derivatives. Then, the KGE becomes

$$(\nabla^2 - \mu^2)\phi = 0 \quad \longrightarrow \quad \frac{1}{r}\frac{\partial^2}{\partial r^2}(r\phi) = \mu^2\phi \tag{4.30}$$

This has general solution

$$\phi = \frac{c_1}{r}e^{-\mu r} + \frac{c_2}{r}e^{\mu r} \tag{4.31}$$

for constants $c_1$ and $c_2$. Evidently, for properly bounded solutions, we set $c_2 = 0$. The quantity $1/\mu \backsim 10^{-15}$ clearly defines some length-scale; it turns out that this is the length-scale of the weak interaction, with which we can associate a mass $m = \mu c/\hbar$, which turns out to be the mass of the pion, a force carrier (well, sort of) associated with the strong nuclear force. These concepts will be covered in greater detail in the BIV notes.

Defining the four-current

$$\mathsf{J}^\mu = i(\phi\partial^\mu\phi^* - \phi^*\partial^\mu\phi) \tag{4.32}$$

it is easily to show that

$$\partial_\mu\mathsf{J}^\mu = i(\phi\partial_\mu\partial^\mu\phi^* - \phi^*\partial_\mu\partial^\mu\phi) \tag{4.33}$$

which for the KGE becomes $\partial_\mu\partial^\mu = 0$; that is, four-current conservation is conserved by the fields $\phi$ that satisfy the KGE, as we would expect from a free particle.