Willoughby Seago

Theoretical Physics

Relativity, Nuclear, and Particle Physics —Relativity

September 20, 2021

Course Notes

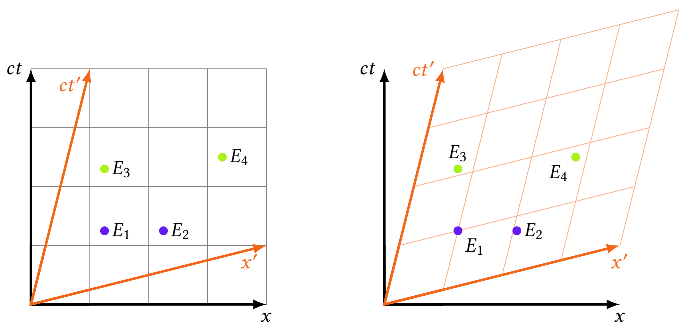
Relativity, Nuclear and Particle Physics (Relativity)

Willoughby Seago

September 20, 2021

These are my notes from the course relativity, nuclear and particle physics (relativity). I took this course as a part of the theoretical physics degree at the University of Edinburgh.

These notes were last updated at 19:49 on November 10, 2021. For notes on other topics see `https://github.com/WilloughbySeago/Uni-Notes`.



Chapters

Page

Contents

Page

One

Introduction

We start this course with a review of classical mechanics, including Newton's laws and Galilean transformations, and conservation laws, . We will then build up special relativity (SR), including length contraction, simultaneity, time dilation, and Lorentz transformations. We end with an introduction to general relativity (GR), including the equivalence principle, the metric tensor, Einstein's field equations, and black holes.

Part I

Classical Mechanics

Two

Newton's Laws and Galilean Transformations

2.1 Newton's Laws

Newton's laws should be familiar, so we state them here in their standard form and expand upon the details later:

- **Newton's first law** (N1) states that a body remains in a state of rest or uniform motion in a straight line unless acted on by an external force.
- **Newton's second law** (N2) states that the rate of change of linear momentum of a body is proportional to the magnitude of the force acting upon the body and is in the direction of the force.
- **Newton's third law** (N3) states that for every action there is an equal and opposite reaction.
- **Newton's law of gravitation** (NG) states that the force between two massive bodies due to gravity is proportional to the product of their masses and inversely proportional to the square of the distance between them. Also the force always acts to bring the bodies closer together.

In these we take "a body" to be a point particle. Uniform motion means that the velocity is constant. This means that the direction of motion doesn't change so, for example, circular motion at a constant speed is *not* uniform motion.

We can regard N1 as a special case of N2 with the force equal to zero. We can *define* force using N2 by choosing appropriate units such that the constant of proportionality is equal to 1 and hence

$$\boldsymbol{F} \coloneqq \dot{\boldsymbol{p}} \tag{2.1.1}$$

where we use the dot notation $\dot{f} \coloneqq \mathrm{d}f/\mathrm{d}t$ for some differentiable quantity, f, and time t. Here $\boldsymbol{p}$ is the linear momentum which, for Newtonian mechanics, is defined to be

$$\boldsymbol{p} \coloneqq m\boldsymbol{v} = n\dot{\boldsymbol{r}} \tag{2.1.2}$$

where $\boldsymbol{v}$ is the velocity vector and $\boldsymbol{r}$ the position vector.

We often deal with the special case where m is constant. In this case $\dot{\boldsymbol{p}} = m\dot{\boldsymbol{v}} = m\ddot{\boldsymbol{r}} = m\boldsymbol{a}$. Unless otherwise specified we will usually assume that mass is a constant.

It seems like N2 defies everyday experience where when we stop pushing an object it stops moving. This is simply because we fail to account for the force of friction opposing the motion.

We can write NG as

$$\boldsymbol{F} = -\frac{GMm}{r^2}\hat{\boldsymbol{r}} \tag{2.1.3}$$

where M and m are the masses of the two bodies, r is their separation, $\hat{\boldsymbol{r}}$ is a unit vector pointing between the bodies, to the body upon which the force acts, and $G = 6.6743\,\mathrm{m^3\,kg^{-2}\,s^{-1}}$ is the universal **gravitational constant**.

As a final remark for this section: we have seen two different ways that mass appears here. In N2 mass is related to the inertia of the body, that is its ability to resist changes in its velocity. In NG mass is a source of a force. We will discus this when we consider GR.

2.2 Frame of Reference

In order to have equations of motion we generally need a coordinate system and a way to specify time. This is what we call a frame of reference. More specifically we can consider some Cartesian coordinate system, (x, y, z), and time measured by some ideal clock, t.

If N1 holds in a frame we call it an **inertial frame**.

We are most interested in the case of two inertial frames moving relative to each other. To aid us in our analysis we define the **standard configuration** in which one frame, S', is moving at a speed V in a straight line relative to another frame, S. We choose a right handed Cartesian coordinate system such that the motion of S' is along the x-axis of S and all axes align. We define time such that $t = t' = 0$ when the origins of the two frames coincide. Unless we state otherwise we will assume frames are in the standard configuration.

2.3 Galilean Transformations

An **event**, for example a light flashing, has a position, (x, y, z), in a given frame as well as occurring at some specified time, t. In Newtonian mechanics the existence of **universal time** is axiomatic. That is once two clocks are synchronised they will agree on all time measurements.

At this point we mention that all of this assumes the existence of an intelligent observer capable of making ideal measurements of position and time.

Consider Figure 2.1 which shows two frames in the standard configuration with frame S' moving at speed V relative to S. Using the lengths x, x', and Vt shown in the diagram we can readily see that $x = x' + Vt$. Combining this with the fact that the other axes are perpendicular to the motion and at one point coincided we get the **Galilean transformation** for position and time:

$$x' = x - Vt, \qquad y' = y, \qquad z' = z, \qquad \text{and} \qquad t' = t. \tag{2.3.1}$$

Equivalently taking $\boldsymbol{V} = (V, 0, 0)$, $\boldsymbol{r} = (x, y, z)$, and $\boldsymbol{r}' = (x', y', z')$ we have

$$\boldsymbol{r}' = \boldsymbol{r} - \boldsymbol{V}t, \qquad \text{and} \qquad t' = t. \tag{2.3.2}$$

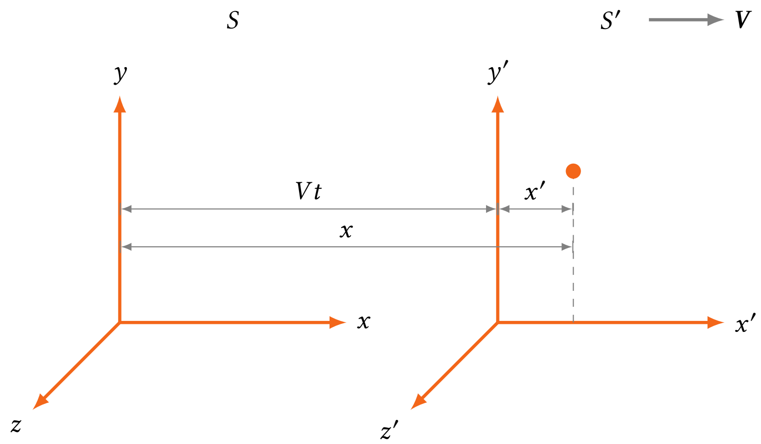


Figure 2.1: Two frames, S and S', in the standard configuration with S' moving at speed V relative to S. Also shown is an event, here a dot, and how this relates to various lengths in the diagram.

Note that in this derivation we have made the assumption that space is isotropic, that is the laws of physics are the same no matter where we are. If we combine this idea with the speed of light being constant we can derive all of SR so it is clearly a very powerful concept.

We can simply differentiate these equations to get the relevant transformation for velocities:

$$\dot{x}' = \dot{x} - V, \qquad \dot{y}' = \dot{y}, \qquad \text{and} \qquad \dot{z}' = \dot{z}, \tag{2.3.3}$$

which implies

$$u'_x = u_x - V, \qquad u'_y = u_y, \qquad \text{and} \qquad u'_z = u_s \tag{2.3.4}$$

or equivalently

$$\boldsymbol{u}' = \boldsymbol{u} - \boldsymbol{V} \tag{2.3.5}$$

where $\boldsymbol{u} = (u_x, u_y, u_z) = (\dot{x}, \dot{y}, \dot{z})$ and similarly $\boldsymbol{u}' = (u'_x, u'_y, u'_z) = (\dot{x}', \dot{y}', \dot{z}')$.

From this we see that the velocity as measured relative to the origin of a given frame is frame dependent. However, suppose we have two objects, moving at speeds $\boldsymbol{u}$ and $\boldsymbol{v}$ in frame S which are $\boldsymbol{u}'$ and $\boldsymbol{v}'$ in frame S'. Then the relative velocity of these two objects in S' is

$$\boldsymbol{v}' - \boldsymbol{u}' = \boldsymbol{v} - \boldsymbol{V} - (\boldsymbol{u} - \boldsymbol{V}) = \boldsymbol{v} - \boldsymbol{u} \tag{2.3.6}$$

which is the relative velocity in frame S. This means that relative velocities are not dependent on the frame. An alternative way of putting this is that relative velocities are Galilean **invariant**, meaning they don't change under a Galilean transformation. This is a powerful concept and invariants under various transformations are incredibly important in many areas of physics.

2.3.1 Invariance of Newton's Laws

Consider some force $\boldsymbol{F}$ and a particle of constant mass, m. Then in some inertial frame, S, we have

$$F_x = m\frac{\mathrm{d}v_x}{\mathrm{d}t} = m\frac{\mathrm{d}^2x}{\mathrm{d}t^2} \tag{2.3.7}$$

and in another initial frame, S', we have

$$F'_x = m\frac{\mathrm{d}v'_x}{\mathrm{d}t} = m\frac{\mathrm{d}}{\mathrm{d}t}(v_x - V) = m\frac{\mathrm{d}v_x}{\mathrm{d}t} = F_x \tag{2.3.8}$$

where we assume that V is constant. Similarly we can show that $F_y = F'_y$ and $F_z = F'_z$. This shows that $\boldsymbol{F'} = \boldsymbol{F}$ for inertial frames and hence $\boldsymbol{a'} = \boldsymbol{a}$. What this means is that N2 is Galilean invariant, by which we mean its form is unaffected by a Galilean transformation.

2.3.2 Newtonian Relativity Principle

The **Newtonian principle of relativity** states that

> The laws of Newtonian mechanics are invariant under Galilean transformations and have the same form in all inertial frames of reference.

This is basically a fancy way of stating that all inertial frames are equivalent and there is no experiment we can do to distinguish between an inertial frame at rest and an inertial frame that is moving relative to some other inertial frame without making measurements in this external frame.

2.4 Accelerating Frames

Consider the setup in Figure 2.1 which shows two frames, S and S', where S' is accelerating along the x-axis of S at some acceleration A.

The same logic we used to derive the Galilean transformations gives us

$$\boldsymbol{r'} = \boldsymbol{r} - \boldsymbol{X}, \qquad \text{and} \qquad t' = t. \tag{2.4.1}$$

where $\boldsymbol{X}$ is the position of the frame S' in frame S. Differentiating we get

$$\boldsymbol{u'} = \boldsymbol{u} - \dot{\boldsymbol{X}} = \boldsymbol{u} - \boldsymbol{V} \tag{2.4.2}$$

where we define $\boldsymbol{V} := \dot{\boldsymbol{X}}$ as the instantaneous velocity of S' with respect to S. Suppose that $\boldsymbol{A}$ is constant. Then basic mechanics tells us that

$$V = V_0 + At \tag{2.4.3}$$

where V_0 is the speed of frame S' in S at $t = 0$. Substituting this into our equation we have

$$\boldsymbol{u'} = \boldsymbol{u} - \boldsymbol{V} = \boldsymbol{u} - \boldsymbol{V_0} - \boldsymbol{A}t \tag{2.4.4}$$

differentiating gives us

$$\boldsymbol{a'} = \boldsymbol{a} - \boldsymbol{A}. \tag{2.4.5}$$

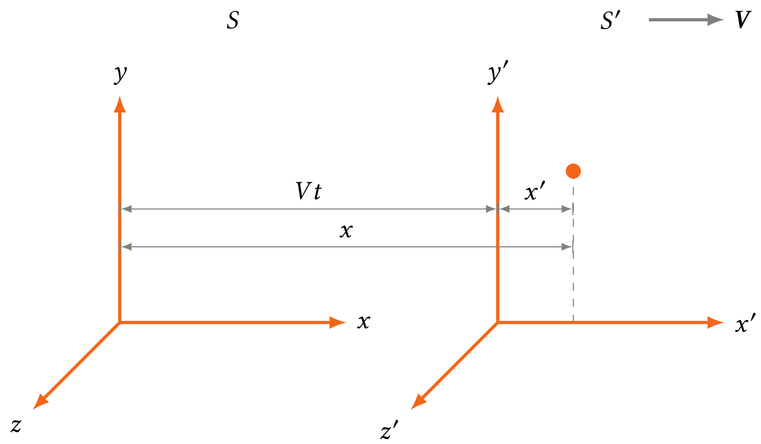


Figure 2.1: Two frames, S and S', in the standard configuration with S' moving at speed V relative to S. Also shown is an event, here a dot, and how this relates to various lengths in the diagram.

Multiplying through by the mass, m, of some particle we have

$$m\boldsymbol{a}' = m\boldsymbol{a} - m\boldsymbol{A}. \tag{2.4.6}$$

The left hand side of this is the force acting on the particle to produce the acceleration observed in frame S'. That is

$$\boldsymbol{F}' = m\boldsymbol{a} - m\boldsymbol{A}. \tag{2.4.7}$$

We see that the force is composed of two parts. The expected $m\boldsymbol{a}$ and an extra term, $m\boldsymbol{A}$, which we call a **fictitious force** or **inertial force** due to the fact that it exists only due to the acceleration of the frame.

This "fictitious" force can have very real effects. For example consider going around a corner in a car. It feels like you are pushed back into your seat. This is a fictitious force due to the acceleration needed to change direction.

▪ **Example 2.4.8 — Pendulum on a Train** Consider a simple pendulum hanging vertically downward on a stationary train. The Newtonian principle of relativity states that if the train is instead moving at constant velocity then the pendulum will again hang vertically downward. The interesting case is when the train is accelerating with constant acceleration, $\boldsymbol{A}$. These three cases are shown in Figure 2.2.

For the accelerating case we first work in the frame of the ground, in which the train is accelerating. Resolving the horizontal and vertical forces gives us

$$T\sin\vartheta = mA, \qquad \text{and} \qquad T\cos\vartheta - mg = 0 \tag{2.4.9}$$

where m is the mass of the pendulum and g the magnitude of the acceleration due to gravity. The left hand side of these equations is the force and the right hand side the resulting acceleration.

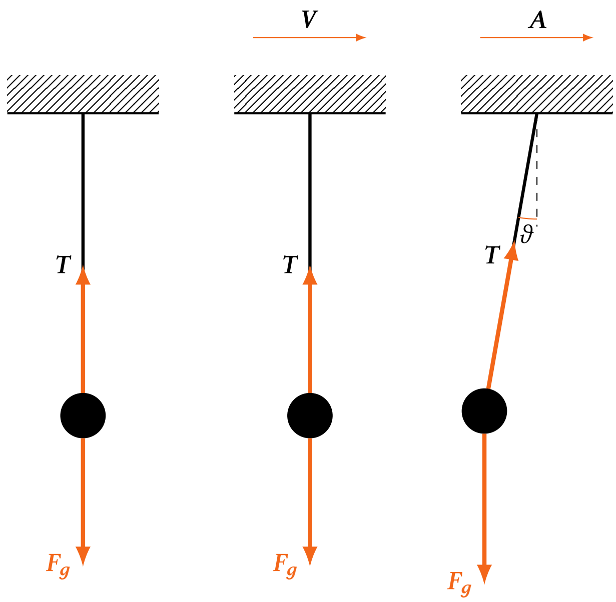


Figure 2.2: A pendulum on a train at rest (left), constant velocity (middle) and accelerating (left).

Working now in the frame of the train and accounting for the resulting fictitious force, $m\boldsymbol{A}$, we can again resolve horizontally and vertically to get

$$T\sin\vartheta - mA, \qquad \text{and} \qquad T\cos\vartheta - mg = 0. \tag{2.4.10}$$

Again we interpret the left hand side as the force and the right hand side as acceleration.
Notice that the equations are the same in both cases, but that the interpretation differs in whether the acceleration of the train is a force or an acceleration.

2.4.0.1 A Brief Note on Gravity

In Example 2.4.8 we were in no doubt that the gravity acted downwards. However, notice that we could interpret the fictitious force as the same as gravity acting in a slightly different direction. This is because both the fictitious force, mA, and the force due to gravity, mg, scale with the mass. Equivalently, both forces satisfy Galileo's principle which states that all bodies, irrespective of their mass, have the same acceleration due to gravity. If gravity and fictitious forces give rise to the same physical effects can we distinguish them? It turns out that the answer is no, and not only this, but we can actually treat gravity as the result of acceleration. This was one of the things that lead Einstein to develop GR.

Three

Two Body Systems

In this chapter we consider two body systems. That is systems of two particles with masses m_1 and m_2 and positions $\boldsymbol{r}_1$ and $\boldsymbol{r}_2$. We define $\boldsymbol{f}$ to be the force on body one due to body two, and by N3 $-\boldsymbol{f}$ is the force on body two due to body one. We also consider the external forces $\boldsymbol{F}_1$ and $\boldsymbol{F}_2$ acting on bodies one and two respectively.

3.1 Conservation of Momentum

Newton's second law gives us the equations of motion

$$\boldsymbol{F}_1 + \boldsymbol{f} = m_1\ddot{\boldsymbol{r}}_1 \tag{3.1.1}$$

$$\boldsymbol{F}_2 - \boldsymbol{f} = m_2\ddot{\boldsymbol{r}}_2 \tag{3.1.2}$$

Adding together these equations we get

$$\boldsymbol{F}_1 + \boldsymbol{F}_2 = m_1\ddot{\boldsymbol{r}}_1 + m_2\ddot{\boldsymbol{r}}_2 = m_1\dot{\boldsymbol{v}}_1 + m_2\dot{\boldsymbol{v}}_2. \tag{3.1.3}$$

In the absence of external forces, i.e. if $\boldsymbol{F}_1 = \boldsymbol{F}_2 = \boldsymbol{0}$, then this reduces to

$$m_1\dot{\boldsymbol{v}}_1 + m_2\dot{\boldsymbol{v}}_2 = 0. \tag{3.1.4}$$

Integrating with respect to time and recognising the definition of linear momentum we have

$$m_1\boldsymbol{v}_1 + m_2\boldsymbol{v}_2 = \boldsymbol{p}_1 + \boldsymbol{p}_2 = \text{constant}. \tag{3.1.5}$$

From this we conclude that in the absence of external forces the total linear momentum is conserved. Importantly this is independent of the interactions between the particles since the $\boldsymbol{f}$s cancel out early on. This suggests that we should consider the system as a whole.

3.2 Conservation of Energy

The **work done** by a force $\boldsymbol{F}$ acting on a particle to move it through an infinitesimal displacement $\mathrm{d}\boldsymbol{r}$ is $\mathrm{d}W = \boldsymbol{F}\cdot\mathrm{d}\boldsymbol{r}$. The work done to move the particle along a path from A to B is then

$$W_{AB} = \int_A^B \boldsymbol{F}\cdot\mathrm{d}\boldsymbol{r} \tag{3.2.1}$$

where the integral is over the path. Evaluating this integral we have

$$
\begin{aligned}
W_{AB} &= \int_A^B \boldsymbol{F} \cdot \mathrm{d}\boldsymbol{r} && (3.2.2)\\
&= \int_A^B \dot{\boldsymbol{p}} \cdot \mathrm{d}\boldsymbol{r} && (3.2.3)\\
&= m \int_A^B \dot{\boldsymbol{v}} \cdot \mathrm{d}\boldsymbol{r} && (3.2.4)\\
&= m \int_A^B \dot{\boldsymbol{v}} \cdot \frac{\mathrm{d}\boldsymbol{r}}{\mathrm{d}t} \,\mathrm{d}t && (3.2.5)\\
&= m \int_A^B \dot{\boldsymbol{v}} \cdot \boldsymbol{v} \,\mathrm{d}t && (3.2.6)\\
&= \frac{m}{2} \int_A^B \frac{\mathrm{d}}{\mathrm{d}t}(v^2) \,\mathrm{d}t && (3.2.7)\\
&= \frac{1}{2} m v_B^2 - \frac{1}{2} m v_A^2 && (3.2.8)\\
&= T_B - T_A. && (3.2.9)
\end{aligned}
$$

Here we have used the product rule to simplify the integrand:

$$
\frac{\mathrm{d}}{\mathrm{d}t}(v^2) = \frac{\mathrm{d}}{\mathrm{d}t}(\boldsymbol{v} \cdot \boldsymbol{v}) = \boldsymbol{v} \cdot \dot{\boldsymbol{v}} + \dot{\boldsymbol{v}} \cdot \boldsymbol{v} = 2\dot{\boldsymbol{v}} \cdot \boldsymbol{v} \implies \dot{\boldsymbol{v}} \cdot \boldsymbol{v} = \frac{1}{2} \frac{\mathrm{d}}{\mathrm{d}t}(v^2). \tag{3.2.10}
$$

We identify T as the **kinetic energy** with subscripts A and B on T and v denoting the kinetic energy or velocity at the relevant location.

3.2.1 Collisions

An **elastic collision** between two bodies is one where the total kinetic energy is conserved. That is

$$
\frac{1}{2} m_1 v_{1i}^2 + \frac{1}{2} m_2 v_{2i}^2 = \frac{1}{2} m_1 v_{1f}^2 + \frac{1}{2} m_2 v_{2f}^2 \tag{3.2.11}
$$

with subscript i and f denoting initial and final quantities from before and after the collision respectively.

Far more realistic in most cases are **inelastic collisions** where the total kinetic energy decreases, so that

$$
\frac{1}{2} m_1 v_{1i}^2 + \frac{1}{2} m_2 v_{2i}^2 > \frac{1}{2} m_1 v_{1f}^2 + \frac{1}{2} m_2 v_{2f}^2. \tag{3.2.12}
$$

Since the total energy is still conserved this energy must be transformed into another form, often as sound or heat, or it is dissipated in deforming the bodies.

3.3 Centre of Mass Coordinates

We mentioned earlier that it can be advantageous to treat a two body system as a whole. Centre of mass coordinates are one way of doing this. The centre of mass is located at $\boldsymbol{R}$ which is such that

$$
\boldsymbol{R} \coloneqq \frac{m_1 \boldsymbol{r}_1 + m_2 \boldsymbol{r}_2}{m_1 + m_2}. \tag{3.3.1}
$$

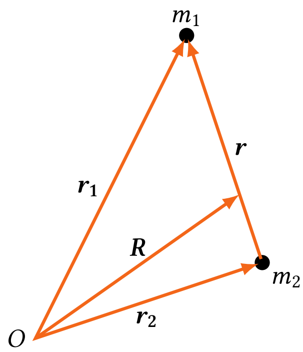


Figure 3.1: Two particles in the centre of mass frame. Drawn as if $m_2 > m_1$ but this needn't be the case.

We also define $\boldsymbol{r}$ as the position of body one from body 2:

$$\boldsymbol{r} := \boldsymbol{r}_1 - \boldsymbol{r}_2. \tag{3.3.2}$$

See Figure 3.1.

Let $M = m_1 + m_2$ be the total mass of the system. Then the centre of mass lies on a line between the two particles a fraction of m_2/M along from particle 1, or m_1/M along from particle 2. Using this, and the fact that $\boldsymbol{r}$ points along this line, from particle 2 to 1, we have

$$\boldsymbol{r}_1 = \boldsymbol{R} + \frac{m_2}{M}\boldsymbol{r}, \qquad \text{and} \qquad \boldsymbol{r}_2 = \boldsymbol{R} - \frac{m_1}{M}\boldsymbol{r}. \tag{3.3.3}$$

We also define the **reduced mass** of the system to be

$$\mu := \frac{m_1 m_2}{m_1 + m_2}. \tag{3.3.4}$$

Using these we can work out Equations (3.1.1) and (3.1.2) in centre of mass coordinates in the absence of external forces. Substituting in the centre of mass and relative position vectors we get

$$\boldsymbol{f} = m_1\ddot{\boldsymbol{r}}_1 = m_1\left(\ddot{\boldsymbol{R}} + \frac{m_2}{M}\ddot{\boldsymbol{r}}\right) = m_1\ddot{\boldsymbol{R}} + \mu\ddot{\boldsymbol{r}}, \tag{3.3.5}$$

$$-\boldsymbol{f} = m_2\ddot{\boldsymbol{r}}_2 = m_2\left(\ddot{\boldsymbol{R}} - \frac{m_1}{M}\ddot{\boldsymbol{r}}\right) = m_2\ddot{\boldsymbol{R}} + \mu\ddot{\boldsymbol{r}}. \tag{3.3.6}$$

Adding these together we get

$$\boldsymbol{0} = m_1\ddot{\boldsymbol{R}} + \mu\ddot{\boldsymbol{r}} + m_2\ddot{\boldsymbol{R}} - \mu\ddot{\boldsymbol{r}} = M\ddot{\boldsymbol{R}}. \tag{3.3.7}$$

Hence $\ddot{\boldsymbol{R}} = \boldsymbol{0}$, which means that the centre of mass moves at a constant velocity, $\dot{\boldsymbol{R}}$. Subtracting the second from the first with $\ddot{\boldsymbol{R}} = \boldsymbol{0}$ we have

$$\boldsymbol{f} = \mu\ddot{\boldsymbol{r}}. \tag{3.3.8}$$

These two equations are very useful. The first describes the motion of the centre of mass, and is independent of any interactions between the particles. The second describes the relative motion of the particles, and is independent of the motion of the centre of mass.

3.4 Scattering

3.4.1 Centre of Mass and Lab Frames

There are two forms of scattering experiment that physicists normally carry out. The first has a fixed target hit by a beam of particles. The second has two beams collide. To make the mathematical treatment of both cases the same we can transform between frames.

We typically consider the two following frames:

- The **LAB frame**, in which we take particle two to be at rest acting as a target before the experiment.
- The **centre of mass frame** (CM), in which the centre of mass is at rest and at the origin, that is $\boldsymbol{R}^* = \boldsymbol{0}$.

> **Notation 3.4.1 — CM Quantities** We denote quantities in the CM frame by an asterisk, *. So, for example $\boldsymbol{r}_1^*$ is the position of particle 1 in the CM frame.

The CM frame is a special case of **centre of momentum frames**, also known as **zero momentum frames** which are frames where the total momentum is zero. In these frames the centre of mass is still stationary, but it may not be at the origin.

We will assume there are no external forces present so both the LAB and CM frames are inertial. In the CM frame the position vectors are

$$\boldsymbol{r}_1^* = \boldsymbol{r}_1 - \boldsymbol{R} = \frac{m_2}{M}\boldsymbol{r} = \frac{\mu}{m_1}\boldsymbol{r}, \qquad \text{and} \qquad \boldsymbol{r}_2^* = \boldsymbol{r}_2 - \boldsymbol{R} = -\frac{m_1}{M}\boldsymbol{r} = -\frac{\mu}{m_2}\boldsymbol{r}. \tag{3.4.2}$$

The separation, $\boldsymbol{r}$, and relative velocity, $\dot{\boldsymbol{r}}$, are Galilean invariant as they are always measured between the two particles and so don't depend on the frame. Hence

$$\boldsymbol{r}^* = \boldsymbol{r}_1^* - \boldsymbol{r}_2^* = \boldsymbol{r}_1 - \boldsymbol{r}_2 = \boldsymbol{r}. \tag{3.4.3}$$

In the CM frame the momenta of the two particles are equal and opposite, that is

$$m_1\dot{\boldsymbol{r}}_1^* = -m_2\dot{\boldsymbol{r}}_2^* = \mu\dot{\boldsymbol{r}} =: \boldsymbol{p}^*. \tag{3.4.4}$$

This shows that the total momentum in the CM frame is zero, and hence it is also a centre of momentum frame.

In any other frame, including, but not limited to, the LAB frame, the centre of mass moves with velocity $\dot{\boldsymbol{R}}$ and the particles have velocities

$$\dot{\boldsymbol{r}}_1 = \dot{\boldsymbol{R}} + \dot{\boldsymbol{r}}_1^*, \qquad \text{and} \qquad \dot{\boldsymbol{r}}_2 = \dot{\boldsymbol{R}} + \dot{\boldsymbol{r}}_2^*. \tag{3.4.5}$$

The momenta of the particles will be

$$\boldsymbol{p}_1 = m_1\dot{\boldsymbol{r}}_r = m_1\dot{\boldsymbol{R}} + \boldsymbol{p}^*, \qquad \text{and} \qquad \boldsymbol{p}_2 = m_2\dot{\boldsymbol{r}}_2 = m_2\dot{\boldsymbol{R}} - \boldsymbol{p}^* \tag{3.4.6}$$

In the LAB frame where particle 2 is initially at rest and hence $\boldsymbol{p}_2 = \boldsymbol{0}$ the centre of mass moves with velocity $\boldsymbol{R} = \boldsymbol{p}^*/m_2$. Therefore the LAB frame in the CM frame moves in the opposite direction with velocity $\dot{\boldsymbol{r}}_2^* = -\boldsymbol{p}^*/m_2$.

Since both the LAB and CM frames are inertial we can move between them with Galilean transformations.

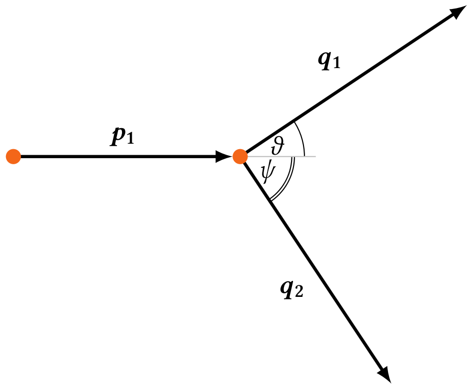


Figure 3.1: Elastic collision of particles in the LAB frame.

3.5 Elastic Collisions

In the absence of external forces the total momentum in any inertial frame must be conserved. We make a further simplification and assume **elastic collisions** in which the total kinetic energy in an inertial frame is conserved. In reality kinetic energy is lost in most collisions, usually as sound or heat, but it can still be a reasonable approximation, for interactions involving small, potentially subatomic, particles collisions are far more likely to be elastic, simply because there are fewer ways for them to lose kinetic energy.

There are typically two ways in which scattering collisions take place. We can either collide two beams of particles or one beam of particles and a target. These are equivalent through a change of frame.

We start by considering an elastic collision in the LAB frame. In this frame one particle is stationary, we refer to this as the target particle, and another particle comes in and collides with it. We denote the momentum of the incoming particle in the LAB frame by $\boldsymbol{p}_1$. The momentum of the target particle is zero. After the collision we denote the momenta of the incoming and target particle by $\boldsymbol{q}_1$ and $\boldsymbol{q}_2$ respectively. We denote the angle through which the incoming and target particles are scattered by ϑ and ψ respectively. Note that when we speak of "before" and "after" what we really mean is sufficiently far away that the interaction between the two particles is negligible. See Figure 3.1.

In the CM frame on the other hand both particles have equal and opposite momenta before the collision, that is the momentum of particle one is $\boldsymbol{p}^*$ and the momentum of particle two is $-\boldsymbol{p}^*$. Conservation of momentum means that the net momentum must be zero after the collision also and so the momenta must be equal and opposite afterwards, we'll denote the momentum of particle 1 after the collision by $\boldsymbol{q}^*$. See Figure 3.2.

We can transform between the two frames with Galilean transformations. Before the collision we have

$$\boldsymbol{p}_1 = m_1\dot{\boldsymbol{r}}_1 = m_1(\dot{\boldsymbol{r}}_1^* + \dot{\boldsymbol{R}}) = \boldsymbol{p}^* + m_1\dot{\boldsymbol{R}}, \tag{3.5.1}$$

$$\boldsymbol{p}_2 = m_2\dot{\boldsymbol{r}}_2 = m_2(\dot{\boldsymbol{r}}_2^* + \dot{\boldsymbol{R}}) = -\boldsymbol{p}^* + m_2\dot{\boldsymbol{R}}. \tag{3.5.2}$$

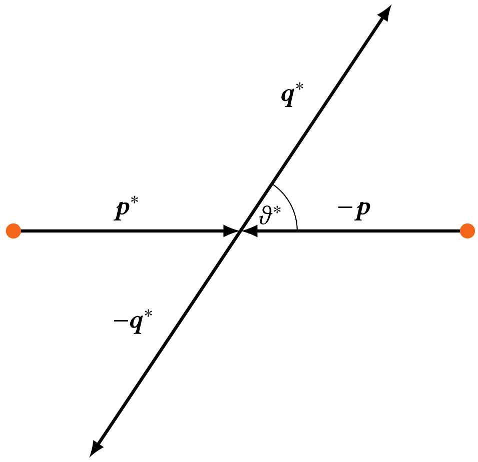


Figure 3.2: Elastic collision of particles in the CM frame.

Similarly after the collision we have

$$\boldsymbol{q}_1 = \boldsymbol{q}^* + m_1\dot{\boldsymbol{R}}, \tag{3.5.3}$$

$$\boldsymbol{q}_2 = -\boldsymbol{q}^* + m_2\dot{\boldsymbol{R}}. \tag{3.5.4}$$

However, $\boldsymbol{q}_2 = \boldsymbol{0}$ in the LAB frame and so $\boldsymbol{p}^* = m_2\dot{\boldsymbol{R}}$. Therefore

$$\boldsymbol{p}_1 = \boldsymbol{p}^*\left(1 + \frac{m_1}{m_2}\right), \tag{3.5.5}$$

$$\boldsymbol{p}_2 = \boldsymbol{q}^* + \boldsymbol{p}^*\frac{m_1}{m_2}, \tag{3.5.6}$$

$$\boldsymbol{q}_2 = \boldsymbol{p}^* - \boldsymbol{q}^*. \tag{3.5.7}$$

Conservation of momentum also gives us $\boldsymbol{p}_1 = \boldsymbol{q}_1 + \boldsymbol{q}_2$. We can combine these relations into a single vector diagram, see Figure 3.3. We can use this diagram to find various relations between quantities. For example we can show that

$$2\psi = \pi - \vartheta^*, \qquad \text{and} \qquad \tan\vartheta = \frac{\sin\vartheta^*}{\cos\vartheta^* + m_1/m_2}. \tag{3.5.8}$$

This last relation in the special case of equal masses gives

$$\tan\vartheta = \frac{\sin\vartheta^*}{\cos\vartheta^* + 1} = \tan\left(\frac{\vartheta^*}{2}\right) \implies \vartheta = \frac{\vartheta^*}{2}. \tag{3.5.9}$$

From this it follows that $\psi + \vartheta = \pi/2$, that is the paths of scattered particles of equal mass are at right angles.

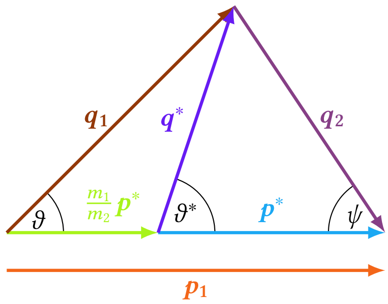


Figure 3.3: A vector diagram showing the relationships between the various momenta and angles in the LAB and CM frames.

Four

Particle Scattering

4.1 Cross Sections

Cross sections allow us to explore interactions between particles. We can think of them as a generalisation of the cross section of macroscopic objects, the shape and position of which tells us whether the two objects will collide. The difference is that particles don't really collide, they interact by various mechanisms, all of which we can take into account with these generalised cross sections.

4.1.1 Differential Cross Sections

For simplicity we will consider a stationary target and an incoming beam of particles, that is we will work in the LAB frame. We characterise the intensity of the incoming beam with the **incident flux**, f, which is defined as the number of particles crossing unit area normal to the beam direction per unit time.

The **impact parameter**, b, is the closest two particles would come if there is no interaction between them. See Figure 4.1. We take the z-axis to be the beam direction and the x and y-axes to be perpendicular to the beam. The scattering angle, ϑ, is the polar angle along which the deflected beam travels after it is sufficiently far from the target such that the interaction between the two is negligible. Due to the symmetry about the z axis the incident particles move in a plane of constant azimuthal angle, φ.

In a scattering experiment we have detectors of a finite size. We therefore need to consider the number of particles scattered into an area. We consider the area that has polar angle between azimuthal angle between φ and $\varphi + \mathrm{d}\varphi$. If we change b slightly, say by $\mathrm{d}b$, then this will change the scattering angle, ϑ, by a small amount, $\mathrm{d}\vartheta$. We can interpret these small intervals of polar coordinates as defining an area, $\mathrm{d}\sigma$. See

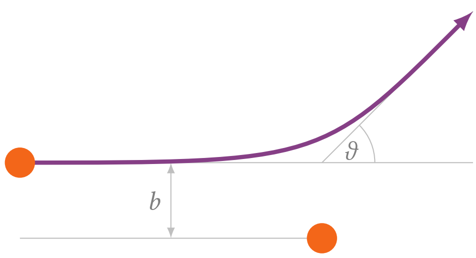


Figure 4.1: Scattering process showing the impact parameter, b, and deflection angle, ϑ.

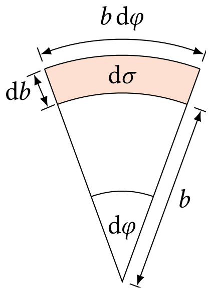


Figure 4.2: The cross section, $\mathrm{d}\sigma$. Seen as if looking along the beam direction.

Figure 4.2. Changing b also changes the **scattering rate**, w, by some small amount, $\mathrm{d}w$, given by

$$\mathrm{d}w = f\,\mathrm{d}\sigma = f b\,\mathrm{d}\varphi|\mathrm{d}b|. \tag{4.1.1}$$

Note that we take the absolute value of $\mathrm{d}b$ since it is possible that b may decrease. From this we can see that the scattering rate depends on the incident flux, f, and the angular size of the detector, $\mathrm{d}\Omega$. We normalise this out by dividing by f and $\mathrm{d}\Omega$.

The angular size, $\mathrm{d}\Omega$, is a **solid angle**. The solid angle subtended at the origin by an element of area $\mathrm{d}A$ at distance L is given by $\mathrm{d}\Omega = \mathrm{d}A/L^2$. In spherical coordinates we have

$$\mathrm{d}A = (L\sin\vartheta\,\mathrm{d}\vartheta)(L\,\mathrm{d}\varphi) \tag{4.1.2}$$

which gives

$$\mathrm{d}\Omega = \sin\vartheta\,\mathrm{d}\vartheta\,\mathrm{d}\varphi. \tag{4.1.3}$$

Solid angles are dimensionless, they are measured in steradians, which are to spheres as radians are to circles. The total solid angle subtended by an entire sphere is

$$\int_{S^2}\int_0^{\pi} \sin\vartheta\,\mathrm{d}\vartheta \int_0^{2\pi} \mathrm{d}\varphi = 4\pi. \tag{4.1.4}$$

Here S^2 is the unit sphere.

It follows that the number of particles scattered in the direction defined by ϑ and φ per second, per unit flux, per unit angular size is

$$\frac{\mathrm{d}w}{f\,\mathrm{d}\Omega} = \frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} = \frac{b}{\sin\vartheta}\left|\frac{\mathrm{d}b}{\mathrm{d}\vartheta}\right|. \tag{4.1.5}$$

We call the quantity $\mathrm{d}\sigma/\mathrm{d}\Omega$ the **differential cross section**. It depends only on the relationship between the impact parameter, b, and the scattering angle, ϑ. This in turn is determined by the nature of the interaction between the particles.

4.1.2 Total Cross Section

Integrating over all scattering directions we get

$$\sigma := \int \frac{\mathrm{d}\sigma}{\mathrm{d}\Omega}\,\mathrm{d}\Omega. \tag{4.1.6}$$

This is the **cross section** for the region over which we integrate. If we integrate over the entire sphere then this is the **total cross section**.

4.2 Conservation Laws

So far we have considered the interaction in the LAB frame, in which the target particle remains stationary throughout. If we want to consider the motion in the CM frame then we need to know how the target particle moves after the interaction. The exact details of course depend on the interaction, but there are some general principles we can use. We can use conservation of angular momentum and kinetic energy conservation. Note that kinetic energy is conserved only when there is no interaction, that is at large distances. When the particles are close some of the kinetic energy will be transferred to potential energy. This is simply energy conservation when potential energy is zero.

In general the particles in the beam will change momentum and this must be balanced by the recoil of the target particle. In the limit of a significantly more massive target particle this recoil velocity is small. The kinetic energy goes with the square of the velocity and therefore the kinetic energy taken away by the target is small and we consider the collision to be elastic.

Inelastic collisions are also possible. In these kinetic energy is lost. After the probe and target particles are sufficiently separated there can be no energy stored as potential energy. So where has this kinetic energy gone? The answer is that the particles must be changed in some way. Possibly by excitation into a more energetic state or, in more extreme cases, the particles may break up after the interaction. We will see later that *all* inelastic collisions result in violation of conservation of mass.

4.3 Hard Sphere Scattering in the CM Frame

Consider two hard, smooth, spheres of radius R, such as billiard balls, in the CM frame. By "hard" here we mean that the surface cannot be deformed and therefore no kinetic energy is lost to deforming the spheres. By "smooth" we mean there is no friction and so kinetic energy is not lost due to friction. These conditions mean we can consider elastic collisions.

In order to find the differential cross section in the CM frame, $\mathrm{d}\sigma^*/\mathrm{d}\Omega^*$, we need to find a relationship between b and ϑ^*. Notice that the spheres collide only if $b < 2R$. See Figure 4.1, from this we see that $b = 2R\sin\alpha^*$. We need to relate α^* to ϑ^* in order to find the cross section. To do this we need to consider the nature of the interaction. In this simple case we have a force acting perpendicular to the surfaces, along the radial direction, at the point of contact. This means that there is no component of force in the tangential direction and hence there is no momentum change along this direction. This means that $p^*\sin\alpha^* = q^*\sin\beta^*$. For an elastic collision $p^* = q^*$ and hence $\alpha^* = \beta^*$. It then follows that $2\alpha^* + \vartheta^* = \pi$ and hence

$$b = 2R\sin\left(\frac{\pi - \vartheta^*}{2}\right) = 2R\cos\frac{\vartheta^*}{2}. \tag{4.3.1}$$

Hence

$$\frac{\mathrm{d}\sigma^*}{\mathrm{d}\Omega^*} = \frac{b}{\sin\vartheta^*}\left|\frac{\mathrm{d}b}{\mathrm{d}\vartheta^*}\right| = R^2. \tag{4.3.2}$$

The differential cross section is independent of the scattering angle in this case and therefore we call it **isotropic**, meaning the same in all directions. The total cross

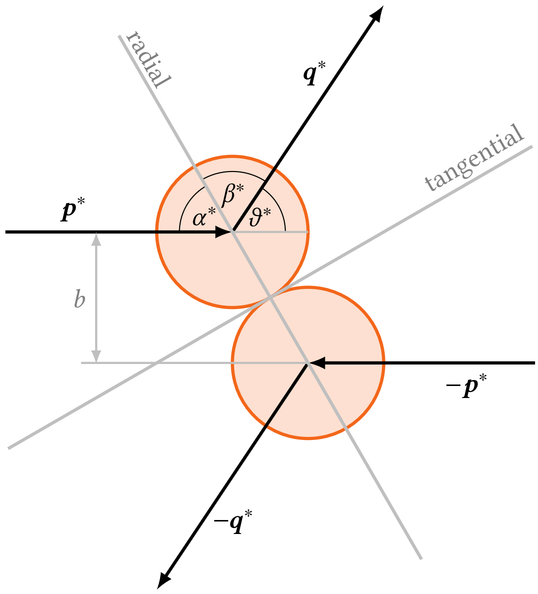


Figure 4.1: Two hard spheres colliding.

section is

$$\sigma^* = \int_{S^2} \frac{d\sigma^*}{d\Omega^*} \, d\Omega^* = R^2 \int_{S^2} d\Omega^2 = 4\pi R^2. \tag{4.3.3}$$

Notice that this is the cross sectional area of the cylinder of radius $2R$ in which the centres of both spheres must be in order for a collision to occur.

4.4 Rutherford Scattering

One of the first scattering experiments ever was Rutherford scattering, in which alpha particles where fired at a thin gold foil. The alpha particles scatter off of the nucleus through the electromagnetic interaction. Suppose the incident alpha particle starts with speed v far from the scattering centre. It's energy at this point is purely kinetic, and given by $E = mv^2/2$. The magnitude of the angular momentum about the target nucleus is $L = mvb$.

Since electrostatic energy decreases with distance and the nucleus is much more massive than the alpha particle conservation of energy means that the initial and final speeds of the alpha particle must be the same.

The direction of the particle changes, so its momentum changes. The change in momentum is

$$\Delta p = 2mv \sin\frac{\vartheta}{2}. \tag{4.4.1}$$

See Figure 4.1 for the direction of $\boldsymbol{\Delta p}$. The change in momentum is also given by the net impulse, which is

$$\Delta p = \int_{-\infty}^{\infty} F \cos\Phi \, dt = \int_{-\infty}^{\infty} \frac{qQ}{4\pi\varepsilon_0 r^2} \cos\Phi \, dt. \tag{4.4.2}$$

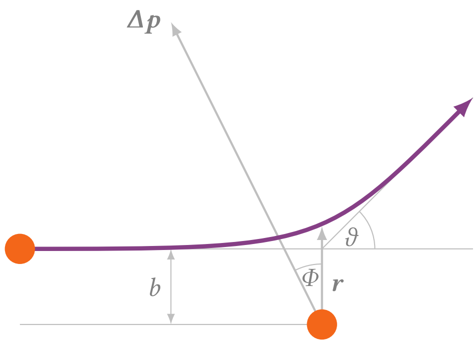


Figure 4.1: Rutherford scattering.

Here q and Q are the charge of the alpha particle and nucleus, r is the distance between them and Φ is the polar angle in plane polar coordinates with the x-axis in the direction of the beam and the origin at the nucleus.

Since the alpha particle is in a central potential its angular momentum is conserved and hence

$$L = mvb = mr^2 \frac{\mathrm{d}\Phi}{\mathrm{d}t} \implies \frac{\mathrm{d}\Phi}{\mathrm{d}t} = \frac{vb}{r^2} \implies \mathrm{d}t = \mathrm{d}\Phi \frac{r^2}{vb}. \tag{4.4.3}$$

Combining this result with the change in momentum, and noticing that $2\Phi_{\max} + \vartheta = \pi$, where Φ is the largest polar angle achieved, we have

$$2mv \sin\frac{\vartheta}{2} = \int_{-\Phi_{\max}}^{\Phi_{\max}} \frac{qQ}{4\pi\varepsilon_0 r^2} \cos\Phi \frac{r^2}{vb}\,\mathrm{d}\Phi = \frac{qQ}{4\pi\varepsilon_0 vb}[\sin\Phi]_{-\Phi_{\max}}^{\Phi_{\max}} = \frac{2qQ}{4\pi\varepsilon_0 vb}\cos\frac{\vartheta}{2}. \tag{4.4.4}$$

From this we have

$$b = \frac{qQ}{4\pi\varepsilon_0 mv^2} \cot\frac{\vartheta}{2}. \tag{4.4.5}$$

We can write this result as $b = a\cot(\vartheta/2)$. It turns out that a is related to r_c, the minimum separation of incident and target particles, by

$$r_c = a + \sqrt{a^2 + b^2}. \tag{4.4.6}$$

It then follows that

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} = \frac{b}{\sin\vartheta}\left|\frac{\mathrm{d}b}{\mathrm{d}\vartheta}\right| = \frac{a^2}{4}\frac{1}{\sin^4(\vartheta/2)}. \tag{4.4.7}$$

This is strongly dependent on the energy of the incoming particle and the scattering angle. It also goes to infinity as $\vartheta \to 0$. The total cross section is also infinite. This is a consequence of the infinite range of the Coulomb force. Even particles with very large impact parameters are scattered, albeit through very small angles.

4.4.1 History

In 1911 Ernest Rutherford, and his students, Hans Geiger and Ernest Marsden carried out a series of scattering experiments with alpha particles and gold foil. They confirmed the predictions above of the cross section. They found that even when the alpha particles where high enough energy that one would expect $r_c \leq 1 \times 10^{-14}$ m they still followed this prediction. From this he concluded that the nucleus must be smaller than 1×10^{-14} m.

Subsequent experiments showed deviation from these predictions at higher energies, suggesting that the alpha particles where penetrating the nucleus, and hence the interactions differed. In the simplest case, which is what Rutherford expected, the nucleus is a charged sphere and hence the charge, and by extension force, inside should be zero. It was later shown however that the electromagnetic force alone could not predict the scattering of these higher energy particles. This was some of the early evidence for the existence of the strong nuclear force.

4.5 Moving Between Inertial Frames

Consider Figure 3.3. As $m_1/m_2 \to 0$ we expect that the quantities in the CM frame should approach those in the LAB frame.

In general however the two frames are different. Often it is useful to be able to find the cross section in one frame and convert to the other. To do so is fairly easy. Note that the relative motion of the frames changes the scattering angle, ϑ, and solid angle, Ω, but not the azimuthal angle, φ, or $\mathrm{d}\sigma$. Therefore we have

$$\begin{aligned}
\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} &= \frac{\mathrm{d}\sigma}{\mathrm{d}\Omega^*}\frac{\mathrm{d}\Omega^*}{\mathrm{d}\Omega} && (4.5.1)\\
&= \frac{\mathrm{d}\sigma^*}{\mathrm{d}\Omega^*}\frac{\mathrm{d}\Omega^*}{\mathrm{d}\Omega} && (4.5.2)\\
&= \frac{\mathrm{d}\sigma^*}{\mathrm{d}\Omega^*}\frac{\sin\vartheta^*\,\mathrm{d}\varphi\,\mathrm{d}\vartheta^*}{\sin\vartheta\,\mathrm{d}\varphi\,\mathrm{d}\vartheta} && (4.5.3)\\
&= \frac{\mathrm{d}\sigma^*}{\mathrm{d}\Omega^*}\frac{\sin\vartheta^*}{\sin\vartheta}\frac{\mathrm{d}\vartheta^*}{\mathrm{d}\vartheta^*} && (4.5.4)\\
&= \frac{\mathrm{d}\sigma^*}{\mathrm{d}\Omega^*}\frac{\mathrm{d}(\cos\vartheta^*)}{\mathrm{d}(\cos\vartheta^*)}. && (4.5.5)
\end{aligned}$$

Part II

Special Relativity

Five

The Basics of Special Relativity

5.1 Breakdown of Galilean Relativity

Around the end of the 19th century came the first signs that classical mechanics alone could not explain every observation. At first physicists attempted to fix individual equations to match observations. These fixes came in essentially two parts, fixes for when things were small and fixes for when things were fast. The small fixes eventually developed into quantum mechanics and the fast fixes into relativity. It is the later of these that we study in this course.

We separate relativity into special and general relativity, with special relativity (SR) being a special case of general relativity (GR). Special relativity deals only with inertial frames and is considerably simpler for it. Both special and general relativity start with the principle of relativity, that the laws of physics are the same in all (inertial for SR) frames. From this and a few other assumptions, known as postulates, it is then possible to derive many consequences including length contraction, time dilation, gravitational lensing, and black holes.

The main sign that special relativity was needed was Maxwell's theory of electromagnetism, which he developed in the 19th century. In this theory light propagates as a wave in the electromagnetic field travelling through the vacuum at the speed of light, $3 \times 10^8\,\mathrm{m\,s^{-1}}$.

Maxwell's equations are *not* Galilean invariant. This means they require a unique frame in which light propagates with speed c. It was initially assumed that electromagnetic waves travelled through a medium, called the ether. The Michelson–Morley experiment disproved the existence of such a medium however, by showing that light travelling on perpendicular paths took the same amount of time to travel the same distance. This could not be true in the presence of the ether since Earth moves through the ether and so one direction should be faster than the other.

There were two competing fixes at the time to this problem. Maxwell and others believed that the fixed either frame exists but that the principle of relativity doesn't apply to electromagnetic fields. Einstein however preferred the other explanation, that the principle of relativity applies, but Galilean transformations are not correct. It is this later fix that proved to be consistent and lead to special relativity.

5.2 Postulates of Special Relativity

In 1905 Einstein formulated two postulates from which the rest of special relativity follows.

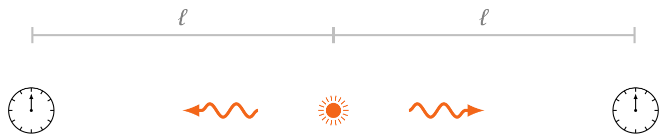


Figure 5.1: Synchronising clocks with a light pulse.

- The **principle of relativity**: All laws of physics have the same form in all inertial frames.
- The speed of light in a vacuum is constant and equal to $c = 2.998 \times 10^8\,\mathrm{m\,s^{-1}}$ in all inertial frames.

It is this second postulate which leads to the weirder consequences of special relativity. The speed of light plays the role of a universal speed limit. Nothing (or more precisely no information) can travel faster than the speed of light. Since in classical mechanics there is no upper speed limit it makes sense that we can recover Newtonian results by taking the limit of $c \to \infty$.

5.3 Time: Synchronisation and Simultaneity

In Galilean relativity we considered time to be universal. If an event was observed in some frame at a given time then an observer in a different frame would agree upon the time at which the event happened. This is not so in special relativity.

In principle there is nothing stopping us from setting up identical clocks at two locations to measure time. We need to synchronise these clocks, in particular we should do so in a way consistent with the postulates of special relativity. Suppose we were to have two clocks, A and B, if we go to the midpoint between A and B are release a flash of light then it will reach both clocks at the same time since the distances are the same and the speed of light is the same by the second postulate. Therefore if we start both clocks at the time they receive the light then they will be synchronised. By extending this method we can synchronise as many clocks as we need. See Figure 5.1.

Now suppose there is an observer at O, the midpoint between A and B, and that O, A, and B are stationary in the inertial frame S. Balls are thrown from A and B to the observer at some speed b, which we take to be much less than c. Both balls arrive at the observer at the same time, which we are free to take as $t = 0$. Suppose that at the same time as the balls reach the observer a train drives past the observer with speed v. Which ball was thrown first according to the stationary observer at O? What about according to an observer on the train?

For the stationary observer in frame S at some short time, Δt, before the balls reach them the balls were a distance $b\Delta t$ away. Extrapolating back the ball from A was thrown a time $t_A = \ell/b$ before it reached the observer, here ℓ is the distance between the observer and A. However, the distance to B from the observer is also ℓ and by the same logic the ball was thrown from B at a time $t_B = \ell/b$ before it reached the observer. That is $t_A = t_B$, and so the stationary observer concludes that both balls were thrown at the same time.

Now consider the same events from the point of view of an observer on the train, we'll call their frame S'. From this frame the train is stationary, we'll call the location of the observer in the train O'. The train observer sees the "stationary" observer, as well as the points A and B, moving. At a short time, $\Delta t'$, before the balls reach the "stationary" observer where were the balls according to the train observer? Applying a Galilean transformation the train observer calculates that the balls are moving with speeds $b - v$ and $b + v$, we'll assume that the train is heading in the direction from A to B, and hence the ball thrown from A has speed $b - v$ and the ball thrown from B has speed $b + v$. The train observer then concludes that a short time, $\Delta t'$, before the balls arrive at the "stationary" observer the ball from A is a distance $(b - v)\Delta t'$ away and the ball from B is a distance $(b + v)\Delta t'$ away. Also at the same time the both points A and B are a distance $v\Delta t'$ further along the AB direction than they are at the time the balls reach the "stationary" observer. Combining this the time taken for the ball to travel from A to O in S' satisfies $\ell - vt'_A = (b - v)t'_A$, and similarly the time taken for the ball to travel from B to O in S' satisfies $\ell + vt'_B = (b + v)t'_B$. Solving these equations the train observer finds $t'_A = t'_B = \ell/b$.

That is the time measured in both frames is the same. Simultaneity is absolute in Galilean relativity.

Repeat the same experiment with pulses of light instead of balls. The same logic in S means that the light arrives at the observer at time $t_A = t_B = \ell/c$.

For the observer on the train things aren't quite so simple. Since light always travels at the same speed in any inertial frame at a time $\Delta t'$ before the light arrives at the "stationary" observer both pulses of light where a distance $c\Delta t'$ away. A careful measurement also reveals that the length of ℓ is not what the "stationary" observer measures but very slightly less. Call it ℓ'. The train observer deduces that the journey time for the two photons where

$$t'_A = \frac{\ell'}{c + v}, \qquad \text{and} \qquad t'_B = \frac{\ell'}{c - v}. \tag{5.3.1}$$

That is, the pulses of light are *not* simultaneous in the S' frame. Simultaneity is *not* absolute in special relativity.

5.4 Lorentz Transforms

Since observers in different frames may disagree about the time at which events happen we know that the Galilean transforms are not correct. So what are the correct transformations? It is the **Lorentz transformation**[1] We will simply state the Lorentz transformation here and derive it later. We assume the system is in the standard configuration and that s' is travelling at a speed v relative to S along the x axis.

[1]we typically speak of a singular Lorentz transformation as we think of applying it to a single four-vector, rather than individual components, more on this later.

$$x' = \gamma(v)[x - \beta ct], \tag{5.4.1}$$
$$y' = y, \tag{5.4.2}$$
$$z' = z, \tag{5.4.3}$$
$$ct' = \gamma(v)[ct - \beta x]. \tag{5.4.4}$$

Here

$$\gamma(v) := \frac{1}{\sqrt{1 - \beta^2}}, \qquad \text{and} \qquad \beta := \frac{v}{c}, \tag{5.4.5}$$

Notice that in the limit of $v \ll c$ the Lorentz transformation reduces to the Galilean transformations as $\beta \to 0$ and hence $\gamma(v) \approx 1$, and $\beta c \to vc/c = v$. We then recover $x' = x - vt$ and $t' = t$.

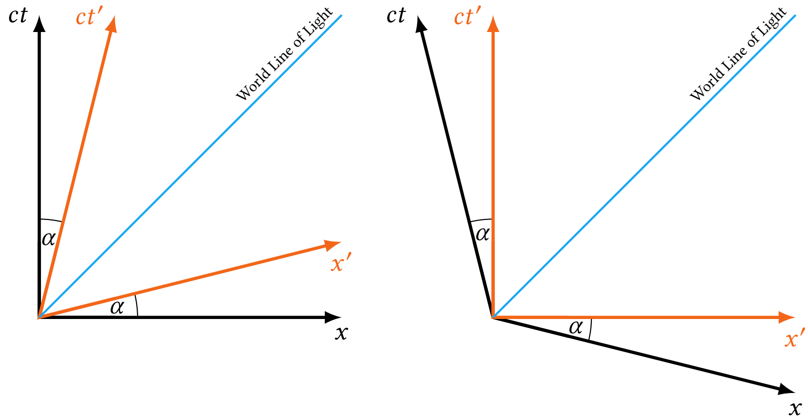


Figure 5.1: Minkowski diagrams showing two frames, S and S', with axes (x, ct) and (x', ct') respectively. The world line of light (or any other light speed particle) is shown. The left hand side shows frame S with orthogonal axes and the right frame S'. The angle between axes is α.

The speed of light appears as a limiting factor in the Lorentz transformation. If $v > c$ then γ ceases to be real and we end up with complex positions, clearly this is non-physical. We will see later that this speed limit has real physical meaning.

The inverse transforms can be found with the principle of relativity, if S' is moving with speed v with respect to S then S is moving with speed $-v$ with respect to S' and since γ is a function of v^2 we have that $\gamma(v) = \gamma(-v)$ and so

$$x = \gamma(v)[x' + \beta ct'], \tag{5.4.6}$$
$$y = y', \tag{5.4.7}$$
$$z = z', \tag{5.4.8}$$
$$ct = \gamma(v)[ct' + \beta x']. \tag{5.4.9}$$

5.5 Minkowski Diagrams

Minkowski diagrams give us a way to graphically visualise space-time. Space-time is four dimensional, fortunately in the standard configuration nothing interesting happens in the dimensions spanned by $\hat{\boldsymbol{y}}$ and $\hat{\boldsymbol{z}}$ and so we can ignore these and consider only a slice at fixed y and z.

Typically we consider two sets of axes, (x, ct) and (x', ct'). The factors of c here playing the double duty of ensuring that the dimensions of the axes are the same, allowing for direct comparison, and scaling the axes so that the scales of events are comparable.

We refer to lines in Minkowski diagrams which correspond to the path of a particle as **world lines**. For example we can see the ct axis as the world line of a particle at rest at the origin, or the diagonal line $x = ct$ as the world line of a particle travelling at the speed of light, starting at the origin at $t = 0$. A simple pair of Minkowski diagrams is shown in Figure 5.1. Both show two frames, the left shows S' in S and the right S in S'. In both the world line of a particle travelling at the speed of light is shown.

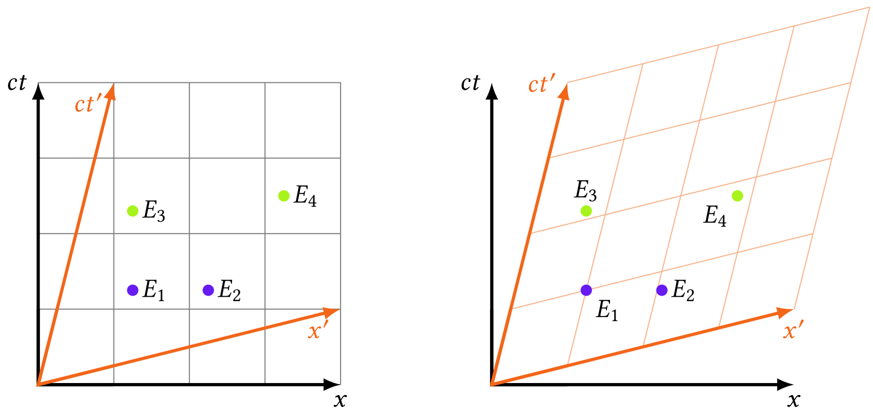


Figure 5.1: Minkowski diagram demonstrating how the order of events can differ between frames.

In frame S frame to origin of S' moves at a speed v relative to frame S. The origin of frame S' has a world line which is a straight line through the origin of S. The angle of the line to the ct vertical axis is α, and satisfies

$$\tan\alpha = \frac{v}{c} = \beta. \tag{5.5.1}$$

The world line of the origin is also the ct' axis.

Similarly the x'-axis in S' is the line for which $t' = 0$. From the Lorentz transformation we can see that this corresponds to the line $ct = \beta x$. This line makes the same angle, α, to the horizontal. Notice that this means that the two axes representing S' quantities are *not* orthogonal.

5.6 Consequences of the Lorentz Transformations

5.6.1 Order of Events

Consider the Minkowski diagrams shown in Figure 5.1. Both show some frame S' in the frame S and four events, represented by dots. The two diagrams differ only by the guidelines drawn on. The first shows lines parallel to the x and ct axes, and the second lines parallel to the x' and ct' axes.

Consider events E_1 and E_2. In frame S these events happen at the same time. This is evidenced by them being on the same line parallel to the x axis. In frame S' these events happen at different times, they aren't on the same line.

Events E_3 occurs *before* event E_4 in frame S. This is evidenced by it being closer to the x-axis. On the other hand event E_3 occurs *after* event E_4 in frame S', since event E_4 is closer to the x'-axis.

5.6.2 Length Contraction

Consider a rod moving with constant velocity according to an observer in an inertial frame, S, in a direction along the length of the rod. In the rest frame of the rod, call this frame S', it is relatively easy to measure the length of the rod. Simply measure the position of both ends and compute the distance between. This can be done without

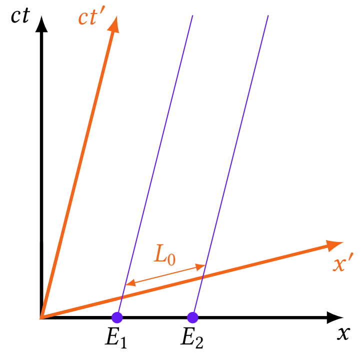


Figure 5.2: The setup for demonstrating length contraction.

considering time since the rod is at rest in S'. Call the measured length L_0, we call this the **proper length** of the rod.

Now consider the process for our observer in S to measure the length of the rod. In this frame the rod is moving and so the observer must measure both ends of the rod *simultaneously*. Say they do so and find that the events corresponding to the rod's ends are (x_1, t_1) and (x_2, t_2) in S, and (x_1', t_1') and (x_2', t_2') in S'. This is shown in Figure 5.2. Notice that E_1 and E_2, the events corresponding to the two ends of the rod, are simultaneous in S but not S', this is fine since measurements in the rest frame don't need to be simultaneous.

Let L be the length as measured in S, that is $L = x_2 - x_1$. How is this related to the proper length as measured in S', $L_0 = x_2' - x_1'$? We can use the Lorentz transformation between S and S' and the fact that $t_2 - t_1 = 0$ as E_1 and E_2 are simultaneous in S.

We find that

$$
\begin{aligned}
L_0 &= x_2' - x_1' && (5.6.1)\\
&= \gamma[x_2 - \beta c t_2] - \gamma[x_1 - \beta c t_1] && (5.6.2)\\
&= \gamma[x_2 - x_1] + \gamma\beta c[t_2 - t_1] && (5.6.3)\\
&= \gamma[x_2 - x_1] && (5.6.4)\\
&= \gamma L. && (5.6.5)
\end{aligned}
$$

This is more commonly written to give the length in some arbitrary frame in terms of the proper length as

$$L = \frac{L_0}{\gamma}. \tag{5.6.6}$$

Since $\gamma \geq 1$ it follows that the observer in S measures the rod to be shorter than its proper length. This is the famous **length contraction**. Moving objects seem shorter along the axis of their motion. Notice that since $\gamma(v) = \gamma(-v)$ it doesn't matter in which way the rod is moving.

5.6.3 Time Dilation

Consider a clock moving at a constant speed wit respect to an inertial frame, S. The world line of the clock is shown in Figure 5.3. The two events, E_1 and E_2, can be thought of as two ticks of the clock. Notice that both occur at the same position in

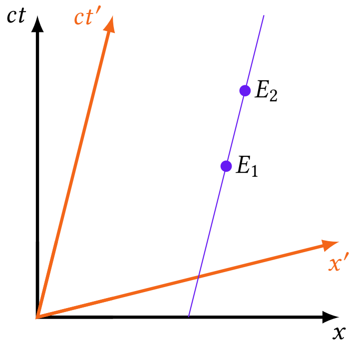


Figure 5.3: The setup for demonstrating time dilation.

S' and hence $x'_2 - x'_1 = 0$. Using the Lorentz transformation for S in terms of S' we have

$$\begin{aligned} ct_2 - ct_1 &= \gamma[ct'_2 + \beta x'_2] - \gamma[ct'_1 + \beta x'_1] && (5.6.7)\\ &= \gamma c[t'_2 - t'_1] + \gamma\beta[x'_2 - x'_1] && (5.6.8)\\ &= \gamma c[t'_2 - t'_1] && (5.6.9)\\ \implies t_2 - t_1 &= \gamma[t'_2 - t'_1]. && (5.6.10) \end{aligned}$$

Here $t'_2 - t'_1 = \tau$ is the time interval measured in the rest frame of the clock, we refer to this as the **proper time**. If the proper time τ passes in some frame then observers in a different frame will see the time t pass with

$$t = \gamma\tau \tag{5.6.11}$$

We see that in S since $\gamma \geq 1$ the time interval is longer. This is the famous **time dilation**. Moving clocks run slow.

5.7 Deriving the Lorentz Transformation

In this section we will derive the Lorentz transformation that we have been using to great effect so far. We will do so using a method called ***k* calculus**.

Consider two inertial frames, S and S', in the standard configuration. An observer, O, stationary at the origin of S emits pulses of light along the positive x-axis at regular intervals, T_0, as measured in S. The light pulses are received by a second observer, O', who is at the origin of S', and moves with it at a constant speed v relative to S. Let T be the time interval that O' measures between receiving light pulses. Define the quantity

$$k := \frac{T}{T_0}. \tag{5.7.1}$$

Similarly imagine that O' is the one who emits light pulses, and that they do so in intervals spaced T apart. Let T_1 be the time between receiving pulses for O. Define

$$k' := \frac{T_1}{T}. \tag{5.7.2}$$

Both of these situations are the same with the frames swapped however and hence by the principle of relativity we must have that $k' = k$.

Now suppose that the observer O emits light and the observer O' simply immediately reflects the pulses back to O. Since the pulses reach O' at intervals of $T = kT_0$ and are immediately reflected back clearly this is equivalent to the observer O' sending out pulses at intervals of T. Hence the observer O sees the reflected pulses at intervals of $T_1 = k'T = kT = k^2T_0$.

This whole process is shown in Figure 5.1. From this we can see that

$$E_2P = \frac{1}{2}E_2R_1 = \frac{c}{2}(T_1 - T_0) = \frac{c}{2}(k^2 - 1)T_0. \tag{5.7.3}$$

We then find that

$$E_1P = E_1E_2 + E_2P = cT_0 + \frac{c}{2}(k^2 - 1)T_0 = \frac{c}{2}(k^2 + 1)T_0. \tag{5.7.4}$$

Trig tells us that

$$PE' = E_1P \tan\alpha = E_1P\beta \tag{5.7.5}$$

where we have used the relationship $\tan\alpha = \beta$ from Equation (5.5.1). We also have $PE' = E_2P$ since these are two sides of an isosceles triangle. Hence,

$$(k^2 + 1)\beta = k^2 - 1 \tag{5.7.6}$$

which gives us

$$k = \sqrt{\frac{1+\beta}{1-\beta}}. \tag{5.7.7}$$

We will see later that this is the relativistic Doppler factor.

Now consider the reflection of a light signal but instead of reflecting from the origin of S' the light reflects from some event $E = (x, t) = (x', t')$. We can work out the time the signal was emitted at E_1 and reflected at E by subtracting from t the propagation time, x/c. Therefore, in S, the time of the event E_1 is $t - x/c$. Similarly we add the propagation time for the reflected signal, also x/c, to t and the result is that event E_4 occurs at $t+x/c$. In S' the signal passes the observer at the time $t'-x'/c$, by the same argument. This corresponds to the event E_2. The reflected signal passes the observer a second time at $t' + x'/c$, corresponding to event E_3. See Figure 5.2.

We can use k-calculus to relate time intervals. The time in S' of event E_2 is k times the time in S of event E_1. The time in S of E_4 is k times the time in S' of E_3. That is

$$t' - \frac{x'}{c} = k\left(t - \frac{x}{c}\right), \qquad \text{and} \qquad t + \frac{x}{c} = k\left(t' + \frac{x'}{c}\right). \tag{5.7.8}$$

We can rewrite this as

$$ct' - x' = k(ct - x), \qquad \text{and} \qquad ct' + x' = \frac{1}{k}(ct + x). \tag{5.7.9}$$

Subtracting the first from the second we get

$$2x' = x\left(\frac{1}{k} + k\right) - ct\left(k - \frac{1}{k}\right). \tag{5.7.10}$$

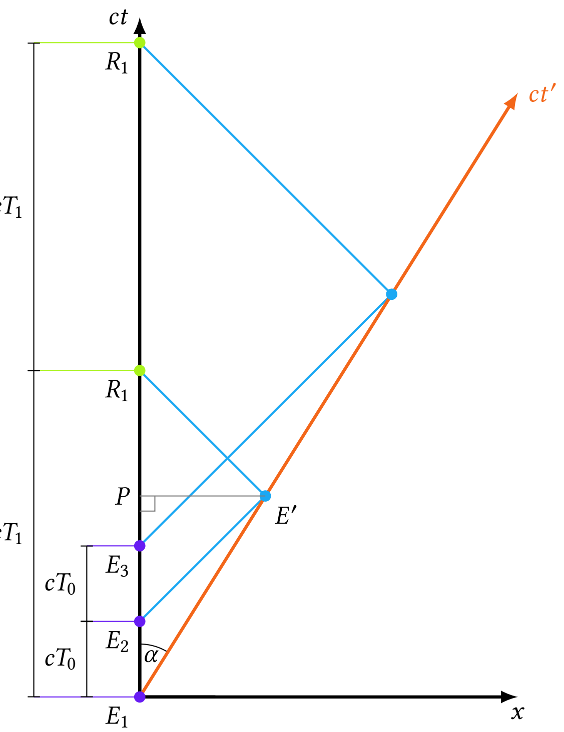


Figure 5.1: Two observers shine lights at each other.

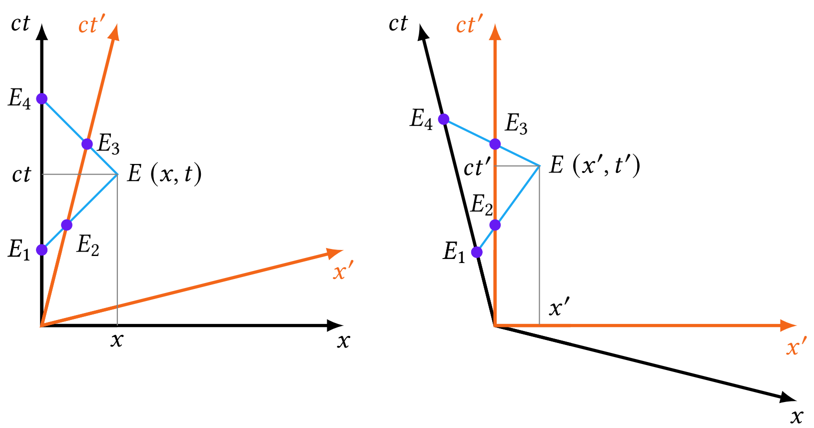


Figure 5.2: Reflection from an arbitrary point in a moving frame,

Using the definition of k we have

$$\frac{1}{k} + k = \frac{1}{k}(k^2 + 1) = \sqrt{\frac{1-\beta}{1+\beta}}\left(\frac{1+\beta}{1-\beta} + 1\right) = \frac{2}{\sqrt{1-\beta^2}} = 2\gamma \tag{5.7.11}$$

$$k - \frac{1}{k} = \frac{1}{k}(k^2 - 1) = \sqrt{\frac{1-\beta}{1+\beta}}\left(\frac{1+\beta}{1-\beta} - 1\right) = \frac{2\beta}{\sqrt{1-\beta^2}} = 2\beta\gamma. \tag{5.7.12}$$

Hence we have

$$x' = \gamma(x = \beta ct). \tag{5.7.13}$$

Similarly if we add the two expressions then we instead get

$$2ct' = ct\left(\frac{1}{k} + k\right) - x(k - \frac{1}{k}) \tag{5.7.14}$$

which gives

$$ct'\gamma(ct - \beta x). \tag{5.7.15}$$

Combining these results with $y' = y$ and $z' = z$ we get the Lorentz transformation.

Six

Relativistic Kinematics

6.1 Proper Time

Consider an observer carrying a clock in some inertial frame, S'. In particular, consider some infinitesimal segment of the observers world line. From frame S we measure time $\mathrm{d}t$ passing between the endpoints of this world-line segment. For the observer the proper time $\mathrm{d}\tau$ passes between the endpoints. These are related by the time dilation formula Equation (5.6.11):

$$\mathrm{d}\tau = \frac{\mathrm{d}t}{\gamma(u)}. \tag{6.1.1}$$

Here u is the observers instantaneous speed compared to the inertial frame S. This is simply the gradient of the observers world line.

A moment later the observer's speed may be different, and hence the gamma factor will be different. The proper time for the world line, between two points start and end, is given by integrating over the world line:

$$\tau = \int_{\text{start}}^{\text{start}} \mathrm{d}\tau = \int_{\text{start}}^{\text{end}} \frac{\mathrm{d}t}{\gamma(u)} = \int_{\text{start}}^{\text{end}} \sqrt{1 - \frac{u^2}{c^2}}\,\mathrm{d}t. \tag{6.1.2}$$

Since $\gamma \geq 1$ we will always find that $\tau \leq t$ where

$$t = \int_{\text{start}}^{\text{end}} \mathrm{d}t \tag{6.1.3}$$

is the time as measured in some other inertial frame. This leads to many seeming paradoxes that can only be resolved with careful consideration. We will see one such example later in the chapter.

6.2 Relativistic Velocity Addition

A moving object has velocity $\boldsymbol{u} = (u_x, u_y, u_z)$ in frame S and $\boldsymbol{u'} = (u'_x, u'_y, u'_z)$ in frame S'. Suppose that the frames are inertial and in the standard configuration such that S' moves with speed v along the x-axis of frame S. We want to find a relationship between $\boldsymbol{u}$ and $\boldsymbol{u'}$. To do so we write the Lorentz transformation in infinitesimal form:

$$\begin{aligned} \mathrm{d}x' &= \gamma(v)[\mathrm{d}x - \beta c\mathrm{d}t], && (6.2.1)\\ \mathrm{d}y' &= \mathrm{d}y, && (6.2.2)\\ \mathrm{d}z' &= \mathrm{d}z, && (6.2.3)\\ c\mathrm{d}t' &= \gamma(v)[c\mathrm{d}t - \beta\mathrm{d}x]. && (6.2.4) \end{aligned}$$

Taking the ratios of these quantities we find that

$$\begin{aligned} u'_x &= \frac{\mathrm{d}x'}{\mathrm{d}t'} && (6.2.5)\\ &= \frac{\gamma(v)[\mathrm{d}x - \beta c\mathrm{d}t]}{\gamma(v)[c\mathrm{d}t - \beta\mathrm{d}x]/c} && (6.2.6)\\ &= \frac{\mathrm{d}x - \beta c\mathrm{d}t}{\mathrm{d}t - \beta\mathrm{d}x/c} && (6.2.7)\\ &= \frac{\frac{\mathrm{d}x}{\mathrm{d}t} - \beta c\frac{\mathrm{d}t}{\mathrm{d}t}}{\frac{\mathrm{d}t}{\mathrm{d}t} - \frac{\beta}{c}\frac{\mathrm{d}x}{\mathrm{d}t}} && (6.2.8)\\ &= \frac{u_x - \beta c}{1 - \frac{\beta}{c}u_x} && (6.2.9)\\ &= \frac{u_x - v}{1 - \frac{vu_x}{c^2}}. && (6.2.10) \end{aligned}$$

Similarly we find that

$$\begin{aligned} u'_y &= \frac{\mathrm{d}y'}{\mathrm{d}t'} = \frac{u_y}{\gamma(v)\left[1 - \frac{vu_x}{c^2}\right]}, && (6.2.11)\\ u'_z &= \frac{\mathrm{d}z'}{\mathrm{d}t'} = \frac{u_z}{\gamma(v)\left[1 - \frac{vu_x}{c^2}\right]}. && (6.2.12) \end{aligned}$$

Similarly using the inverse Lorentz transformations we have

$$\begin{aligned} u_x &= \frac{u'_x + v}{1 + \frac{vu'_x}{c^2}}, && (6.2.13)\\ u_y &= \frac{u'_y}{\gamma(v)\left[1 + \frac{vu'_x}{c^2}\right]}, && (6.2.14)\\ u_z &= \frac{u'_z}{\gamma(v)\left[1 + \frac{vu'_x}{c^2}\right]}. && (6.2.15) \end{aligned}$$

Consider the special case of $\boldsymbol{u} = \boldsymbol{0}$. We have $u'_x = -v$ and $u'_y = u'_z = 0$. Therefore a body at rest in S has a velocity $-v$ in the x' direction in S. Similarly a body at rest in S' has speed v in the x direction in S.

If $u_x = c$ then $u'_x = c$ and vice versa. Therefore a light signal propagating with speed c in the x-direction in frame S is a light signal propagating with speed c in the x'-direction in frame S'. This is consistent with the second postulate, in contrast to the Galilean law for addition of velocities.

Notice that in the limit of $v, u \ll c$ these transformations reduce to the Galilean law for addition of velocities since $\gamma \approx 1$ and $vu_x/c^2 \approx 0$.

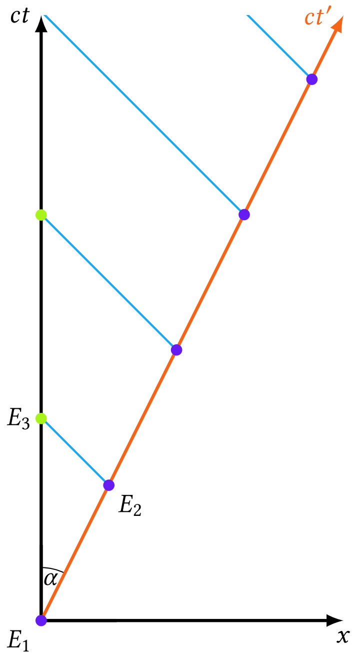


Figure 6.1: Device emitting electromagnetic waves.

6.3 Relativistic Doppler

Consider an electromagnetic wave emitter. The world line for such a device is shown in Figure 6.1. The events labelled along the line correspond to the peak of a wave leaving the emitter. This occurs with period t'_2. The frequency of the wave in the rest frame of the emitter is hence $1/t'_2$.

The events labelled along the ct axis are the wave crests reaching the origin of the frame S. The apparent frequency of the wave to an observer in S is thus $1/t'_3$. Simple geometry gives us

$$ct_3 = ct_2 + \beta ct_2 = ct_2(1+\beta). \tag{6.3.1}$$

Here we have used $\tan\alpha = \beta$. Applying the inverse Lorentz transformation we have

$$ct_2 = \gamma(ct'_2 + \beta x'_2). \tag{6.3.2}$$

We can simplify this since $x'_2 = 0$ and so

$$\begin{align}
ct_3 &= ct'_2\gamma(1+\beta) \tag{6.3.3}\\
&= ct'_2\frac{1+\beta}{\sqrt{1-\beta^2}} \tag{6.3.4}\\
&= ct'_2\frac{1+\beta}{\sqrt{(1-\beta)(1+\beta)}} \tag{6.3.5}\\
&= ct'_2\sqrt{\frac{1+\beta}{1-\beta}}. \tag{6.3.6}
\end{align}$$

Hence the frequencies are related by

$$\nu = \nu' \sqrt{\frac{1-\beta}{1+\beta}}. \tag{6.3.7}$$

This is called the **relativistic longitudinal Doppler** formula. When β is positive the source is receding and the frequency measured in S is lower than the emitted frequency, i.e. there is a relativistic **red shift**.

We considered here the longitudinal case where the source and observer are moving either closer or further apart along the line between them. Resolve the velocity, $\boldsymbol{u}$, into two components, u_r, radially along the line of sight, and u_t, transverse to the line of sight. Defining the angle ϑ to satisfy $\tan\vartheta = u_t/u_r$ and noting that the speed satisfies $u^2 = u_r^2 + u_t^2$ we have

$$\frac{\nu}{\nu'} = \frac{\sqrt{1-\frac{u^2}{c^2}}}{1+\frac{u_r}{c}} = \frac{1}{\gamma\left(1+\frac{u}{c}\cos\vartheta\right)} = \frac{1}{\gamma(1+\beta\cos\vartheta)}. \tag{6.3.8}$$

For purely radial motion this reduces to Equation (6.3.7). Suppose instead that the motion is purely transverse ($\vartheta = \pi/2$). We observe a **relativistic transverse Doppler** effect, which doesn't exist in the standard, non-relativistic case. This is one of the predictions of the new special relativity:

$$\nu = \nu'\sqrt{1-\frac{u^2}{c^2}} = \frac{\nu'}{\gamma}. \tag{6.3.9}$$

6.4 High Speed Travel

Consider some long journey, from A to B, which are a distance ℓ apart on a map. We take S to be the inertial frame that is the rest frame of the map, and therefore the rest frame of A and B. Hence, ℓ is a proper length. Travelling at speed u the journey will be covered in time ℓ/u. Therefore ℓ/c is a lower bound on the journey time, achievable only for massless particles. However, this doesn't account for length contraction/time dilation. In the frame of the map if the traveller moves at speed u then the time for them runs slower by a factor of $1/\gamma$, and so the time taken in frame S is

$$\tau = \frac{\ell}{u\gamma(u)} = \frac{\ell}{u}\sqrt{1-\frac{u^2}{c^2}} \le \frac{\ell}{u}. \tag{6.4.1}$$

So for the traveller the journey takes less time than predicted.

On the other hand for the traveller the journey distance is shortened by a factor of γ to ℓ/γ. The result is the same.

From this we see that length contraction and time dilation are really the same effect but seen from different points of view. This will become clearer later when we introduce four-vectors which allow us to treat both cases the same.

6.5 Twins Paradox

Suppose there are two twins. One of the twins, Stella, gets in a spaceship, flies away from Earth at speed u to some distant planet. After arriving Stella immediately reverses her journey returning back to Earth at speed u. Since she is moving at speed

u the whole time Stella's clock appears to run slowly for the entire journey. Stella's sister, Gaia, who remained on Earth in frame S measures the time of the journey as $2\ell/u$, where ℓ is the distance from Earth to the planet in the Earth/planet rest frame.

Stella on the other hand measures the time taken as $2\ell/(u\gamma)$, which is less than what Gaia measures. Gaia puts this down to time dilation. The problem comes when we consider Stella's rest frame. In this she sees Gaia, and the Earth, fly away at speed u and then return at speed u. It seems like there should be a symmetry between Stella and Gaia's experiences yet there is not. This is the **twins paradox**.

The solution to this paradox is to notice we have skipped over an important detail. We said "since she is moving at speed u the *whole time*", well, Stella *isn't* moving at velocity u the entire time, even if Stella has the ability to turn around instantly she still accelerates. If we account for this then there is no symmetry between Stella and Gaia and hence no paradox.

In the next section we discuss acceleration and then we will come back to this problem.

6.6 Acceleration

It is often said that special relativity can't deal with acceleration. This isn't strictly true. Special relativity can't deal with accelerating frames, only inertial ones. Acceleration as viewed from an inertial frame is within the scope of special relativity.

Consider the simple case of a particle moving along the x-direction in frame S with non-uniform velocity $u(t)$ and acceleration $a = \mathrm{d}u/\mathrm{d}t$. In some other inertial frame S' (i.e. not the frame of the particle) what is the corresponding acceleration $a' = \mathrm{d}u'/\mathrm{d}t'$?

Suppose that S' is moving at speed v along the x-axis of S. Then according to the velocity addition formula Equation (6.2.10):

$$\frac{\mathrm{d}u'}{\mathrm{d}t'} = \frac{1}{\left(1 - \frac{uv}{c^2}\right)} \left[\left(1 - \frac{uv}{c^2}\right) - \frac{(u-v)(-v)}{c^2}\right] \frac{\mathrm{d}u}{\mathrm{d}t'} \tag{6.6.1}$$

$$= \frac{1 - \frac{v^2}{c^2}}{\left(1 - \frac{uv}{c^2}\right)} \frac{\mathrm{d}u}{\mathrm{d}t'}. \tag{6.6.2}$$

Now take S' to be the **instantaneous rest frame** of the particle, also known as the **instantaneously co-moving inertial frame** (ICMF). In this frame the particle is momentarily at rest, i.e. $u' = 0$, and hence $v = u$ at this instant. Setting $v = u$ above we get

$$\frac{\mathrm{d}u'}{\mathrm{d}t'} = \gamma(u)^2 \frac{\mathrm{d}u}{\mathrm{d}t'}. \tag{6.6.3}$$

At the instant being considered S' is the rest frame of the particle and hence t' is a proper time, τ. We then call $\mathrm{d}u'/\mathrm{d}t'$ the **proper acceleration**, a_0. We can then use time dilation to write $\mathrm{d}t' = \mathrm{d}t/\gamma(u)$ and hence

$$\frac{\mathrm{d}u'}{\mathrm{d}t'} = \gamma(u)^3 \frac{\mathrm{d}u}{\mathrm{d}t}. \tag{6.6.4}$$

Thus, we have found a relation between the proper acceleration in the ICMF and the acceleration measured in an arbitrary inertial frame, S:

$$a_0 = \gamma^3 a. \tag{6.6.5}$$

In any frame other than the ICMF the acceleration is less than a_0 since $\gamma^3 \geq 1$.

Given two frames S and S'' in the standard configuration the accelerations in both frames relate to the acceleration in the ICMF the same way and hence

$$\gamma(u'')^3 a'' = \gamma(u)^3 a \tag{6.6.6}$$

where u and u'' are the instantaneous speeds of the particle in S and S'' respectively.

The acceleration differing between frames is new to special relativity. In Galilean relativity we showed that the acceleration, and hence force, is the same in all frames.

Note the slightly different way in which the gamma factor enters here. So far we have been determining the gamma factor from the speed of a frame. Here it enters through the speed of a particle, this will become a more common way to think of the gamma factor.

6.6.1 Constant Proper Acceleration

Consider the special case of constant proper acceleration. Again we restrict motion to the x-direction. Consider the derivative

$$\frac{\mathrm{d}}{\mathrm{d}t}[\gamma(u)u] = \frac{\mathrm{d}\gamma}{\mathrm{d}t}u + \gamma\frac{\mathrm{d}u}{\mathrm{d}t}. \tag{6.6.7}$$

For this we need to find $\mathrm{d}\gamma/\mathrm{d}t$:

$$\frac{\mathrm{d}\gamma}{\mathrm{d}t} = \frac{\mathrm{d}}{\mathrm{d}t}\left(1 - \frac{u^2}{c^2}\right)^{-1/2} \tag{6.6.8}$$

$$= \frac{u}{c^2}\frac{\mathrm{d}u}{\mathrm{d}t}\left(1 - \frac{u^2}{c^2}\right)^{-3/2} \tag{6.6.9}$$

$$= \frac{u}{c^2}\frac{\mathrm{d}u}{\mathrm{d}t}\gamma^3. \tag{6.6.10}$$

Hence,

$$\frac{\mathrm{d}}{\mathrm{d}t}[\gamma(u)u] = \frac{u^2}{c^2}\frac{\mathrm{d}u}{\mathrm{d}t}\gamma^3 + \frac{\mathrm{d}u}{\mathrm{d}t}\gamma. \tag{6.6.11}$$

With some algebra expanding the γ factors we can show that this is equivalent to

$$\frac{\mathrm{d}}{\mathrm{d}t}[\gamma(u)u] = \gamma^3\frac{\mathrm{d}u}{\mathrm{d}t}. \tag{6.6.12}$$

The quantity on the right hand side is simply the proper acceleration, a_0.

Since we are considering the proper acceleration to be constant we can integrate this to get

$$a_0 t = \gamma u = \frac{u}{\sqrt{1 - \frac{u^2}{c^2}}}. \tag{6.6.13}$$

Rearranging this we find that

$$\frac{u}{c} = \frac{a_0 t/c}{\sqrt{1 + \frac{a_0^2 t^2}{c^2}}} = \frac{ct}{\sqrt{\rho^2 + c^2t^2}} \tag{6.6.14}$$

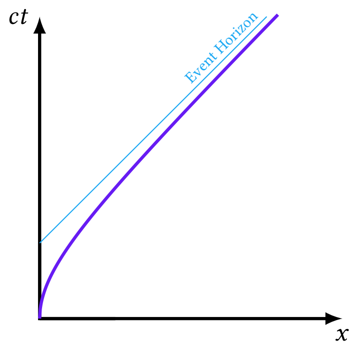


Figure 6.1: The world line of a particle undergoing constant acceleration is a hyperbola. The asymptote is called the event horizon.

where $\rho := c^2/a_0$.

The proper time experienced by an observer undergoing constant acceleration is

$$
\begin{aligned}
c\tau &= c\int_{\text{start}}^{\text{end}} \frac{\mathrm{d}t}{\gamma} && (6.6.15)\\
&= c\int_{\text{start}}^{\text{end}} \sqrt{1-\frac{u^2}{c^2}}\,\mathrm{d}t && (6.6.16)\\
&= c\rho\int_{\text{start}}^{\text{end}} \frac{1}{\sqrt{\rho^2+c^2t^2}}\,\mathrm{d}t && (6.6.17)\\
&= \rho\left[\operatorname{arsinh}\left(\frac{ct}{\rho}\right)\right]_{\text{start}}^{\text{end}}. && (6.6.18)
\end{aligned}
$$

Integrating $u = \mathrm{d}x/\mathrm{d}t$ we get

$$
\begin{aligned}
x &= \int \frac{ct}{\sqrt{\rho^2+c^2t^2}}c\,\mathrm{d}t && (6.6.19)\\
&= \sqrt{\rho^2+c^2t^2}+C && (6.6.20)\\
&= \sqrt{\rho^2+c^2t^2}+x_0-\rho && (6.6.21)
\end{aligned}
$$

where we rewrite the constant of integration C by defining $x = x_0$ at $t = 0$. Rewriting this equation we get

$$(x - x_0 + \rho)^2 - c^2t^2 = \rho^2. \tag{6.6.22}$$

We can identify this as a hyperbola in the (x, ct)-plane. This is shown in Figure 6.1.

6.6.1.1 Non-Relativistic Regime

For any ρ, and hence any a_0, we can always, for some sufficiently small time, approximate the result in Equation (6.6.21) with the binomial expansion. This gives

$$x - x_0 = \frac{c^2t^2}{2\rho} = \frac{1}{2}a_0t^2, \tag{6.6.23}$$

which is a familiar classical result.

6.6.1.2 Highly Relativistic Regime

Rearranging the result of Equation (6.6.22) we have

$$\begin{aligned} ct &= \sqrt{(x - x_0 + \rho)^2 - \rho^2} && (6.6.24)\\ &= \sqrt{(x - x_0)^2 + 2\rho(x - x_0)} && (6.6.25)\\ &= (x - x_0)\sqrt{1 + \frac{2\rho}{x - x_0}}. && (6.6.26) \end{aligned}$$

For any ρ eventually the accelerating object will reach relativistic speed and travel a very large distance. Therefore eventually the term $\rho/(x-x_0)$ becomes small. In this eventuality we can again apply the binomial expansion to get

$$ct \approx x - x_0 + \rho = x - x_0 + \frac{c^2}{a_0}. \tag{6.6.27}$$

This is a line with unit gradient, that is it corresponds to a light signal. It passes through $ct = \rho - x_0$ and is an asymptote to the hyperbola. This is the world line of a light signal passing through $x = x_0$ at $t = c/a_0$. This is called the **event horizon** of the observer.

From Figure 6.1 we can see that a light signal emitted from $x = 0$ at a time later than $t = c/a_0$ will never intersect this line, and hence never intersects the hyperbola. This means that objects whose world lines cross the event horizon to the other side from the hyperbola cannot be seen by the constantly accelerating observer. Of course, if the observer stops accelerating then they will have a speed less than c and the light signal will eventually catch up to them.

6.7 Twins Paradox (Again)

We now have the means to construct a version of the twins paradox where, instead of instantly turning around (almost certainly killing the twin) the twin can undergo a constant acceleration for their entire journey. Suppose Stella sets off with constant proper acceleration, a_0, then at the coordinates (ct_1, x_1) after reaching speed v switches off her engines and coasts along. Later at a time t_2 she switches the engine back on but in reverse, decelerating with constant acceleration, a_0, such that the maximum distance, ℓ, from her twin, is achieved at time t_3. She then continues accelerating until her speed is v, now headed back towards her twin, say this occurs at time t_4. She then travels at this speed until time t_5 where she once again accelerates with constant proper acceleration a_0 arriving back at her twin at time t_6 and travelling at the same speed as her twin.

Stella first reaches speed v when

$$ct_1 = \frac{\rho v/c}{\sqrt{1 - \frac{v^2}{c^2}}}, \qquad \text{and} \qquad x_1 - x_0 + \rho = \frac{\rho}{\sqrt{1 - \frac{v^2}{c^2}}}. \tag{6.7.1}$$

In the interval $[0, t_1]$ we can apply the formula for the proper time to get

$$c\tau_1 = \rho\left[\operatorname{arsinh}\left(\frac{ct}{\rho}\right)\right]_0^{t_2} = \rho \operatorname{arsinh}\left(\frac{v/c}{\sqrt{1 - \frac{v^2}{c^2}}}\right). \tag{6.7.2}$$

By symmetry the deceleration phase for $[t_2, t_3]$ has the same velocity function, but reversed, and so the distance and time taken are the same. Further symmetries mean that the return journey is equivalent to the outgoing journey, but distance decreases instead of increasing.

The total journey time is hence

$$ct_6 = 2c(t_2 - t_1) + 4ct_1 = \frac{2c\ell}{v}\left[1 + \frac{2\rho}{\ell}\left(1 - \sqrt{1 - \frac{v^2}{c^2}}\right)\right]. \tag{6.7.3}$$

In Stella's rest frame we have $c\tau_6 = 2c(\tau_2 - \tau_1) + 4c\tau_1$, and hence

$$c\tau_6 = \frac{2c\ell}{v}\sqrt{1 - \frac{v^2}{c^2}}\left[1 + \frac{2\rho}{\ell}\left(1 - \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}}\right)\right] + 4\rho\,\mathrm{arsinh}\left(\frac{v/c}{\sqrt{1 - v^2/c^2}}\right). \tag{6.7.4}$$

Hence the age difference of the twins is

$$ct_6 - c\tau_6 = \frac{2c\ell}{v}\left(1 + \frac{4\rho}{\ell}\right)\left(1 - \sqrt{1 - \frac{v^2}{c^2}}\right) - 4\rho\,\mathrm{arsinh}\left(\frac{v/c}{\sqrt{1 - \frac{v^2}{c^2}}}\right). \tag{6.7.5}$$

This is determined by the maximum separation, ℓ, the maximum relative speed, v, and Stella's acceleration a_0 (which appears through the hyperbolic radius, ρ).

In the limit of $\ell \gg \rho$ we recover the result of Section 6.5. Provided that Stella knows her initial conditions, has a clock to measure proper time, and an accelerometer she can figure out her coordinates, (x, ct), in Gaia's frame rearranging the equations for $[0, t_1]$, to get

$$ct = \rho\sinh\left(\frac{c\tau}{\rho}\right), \qquad \text{and} \qquad x - x_0 + \rho = \rho\cosh\left(\frac{c\tau}{\rho}\right). \tag{6.7.6}$$

Special relativity has allowed for a treatment of an accelerating observer. What we can't do however is work with the frame that accelerates with the observer. For that we need general relativity.

Seven

Space Time Intervals

In this chapter we define space time intervals and discuss their use. To this end we start with a more familiar example of transformations, namely spatial rotations. We then move onto the more general Lorentz transformation. We also introduce the standard notation of relativity, four-vectors. We will do so here in analogy with the vector notation we are all familiar with. The next chapter will formally define four-vectors and discuss their correct use.

7.1 Spatial Rotations

We noticed previously that a Lorentz transformation mixes the time and space coordinates of events. This is, in many ways, analogous to how rotations mix the three Cartesian coordinates. To aid our understanding of coordinate mixing we will consider as an example rotations about the z-axis.

Consider some point P with coordinates (x, y, z). If we rotate the axes (a passive transformation) by ϑ anti-clockwise about the z-axis then the same point has the coordinates (x', y', z'), which are given by

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} \cos\vartheta & \sin\vartheta & 0 \\ -\sin\vartheta & \cos\vartheta & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} \tag{7.1.1}$$

One important fact about passive transformations is that, while they change the individual components of a vector, they don't change the vector. Therefore defining quantities, like the magnitude don't change. We say the magnitude is **invariant** under rotations[1]. We can show this by considering $\boldsymbol{r} \cdot \boldsymbol{r}$:

[1] or "invariant under the action of O(3)" if we're feeling fancy

$$\begin{aligned} \boldsymbol{r} \cdot \boldsymbol{r} &= x'^2 + y'^2 + z'^2 && (7.1.2) \\ &= (x\cos\vartheta + y\sin\vartheta)^2 + (-x\sin\vartheta + y\cos\vartheta)^2 + z^2 && (7.1.3) \\ &= x^2\cos^2\vartheta + 2xy\cos\vartheta\sin\vartheta + y^2\sin^2\vartheta + x^2\sin^2\vartheta && (7.1.4) \\ &\quad - 2xy\sin\vartheta\cos\vartheta + y^2\cos^2\vartheta + z^2 && (7.1.5) \\ &= x^2(\cos^2\vartheta + \sin^2\vartheta) + y^2(\cos^2\vartheta + \sin^2\vartheta) + z^2 && (7.1.6) \\ &= x^2 + y^2 + z^2. && (7.1.7) \end{aligned}$$

Taking the square root of this quantity gives us the magnitude, and clearly it is the same in both frames.

A simple generalisation of this is that the separation of two points is invariant under rotations. This follows since the magnitude of the vector $\boldsymbol{r}_2 - \boldsymbol{r}_1$ is similarly invariant.

Really what these results say is that the dot product is invariant under rotations. Given how useful magnitudes and dot products are clearly invariants are incredibly invaluable when investigating interesting concepts. It would be nice if we could find some for Lorentz transformations. Fortunately we can, and they're very similar to invariants of rotation.

7.2 Lorentz Transformations

7.2.1 Four-Vectors

When considering rotations we saw that even though the coordinates mixed there were certain combinations of coordinates, such as the dot product, which were invariant. With the goal of developing an analogous concept, but including time since this is also mixed with our coordinates, we define the **four-vector**. We will define this rigorously in the next section. For now just think of it as a notational trick keeping all the components in one array:

$$x^\mu = (x^0, x^1, x^2, x^3) = (ct, x, y, z) = (ct, \boldsymbol{r}). \tag{7.2.1}$$

Here μ (and more generally any Greek index) is an index which runs from 0 to 3. The identification $x_1 \leftrightarrow x$, $x_2 \leftrightarrow y$, and $x_3 \leftrightarrow z$ should be familiar. In special relativity we make a distinction between upper and lower indices, hence x^μ, x^0, etc. are not powers of x but x with a superscript index.

What is possibly new at this point is the identification $x^0 \leftrightarrow ct$. We can think of the c as just being there for dimensional consistency.

7.2.2 Lorentz Transformation in Matrix Form

The Lorentz transformation between two frames in the standard configuration is

$$\begin{pmatrix} ct' \\ x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} \gamma & -\beta\gamma & 0 & 0 \\ -\beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} ct \\ x \\ y \\ z \end{pmatrix}. \tag{7.2.2}$$

Now define ω, a quantity we call the **rapidity**, to satisfy $\cosh\omega = \gamma$ (note that both $\cosh\omega$ and γ take values in $[1, \infty)$). Recall the identity

$$\cosh^2 u - \sinh^2 u = 1. \tag{7.2.3}$$

Rearranging this gives us

$$\sinh\omega = \sqrt{\cosh^2\omega - 1} = \sqrt{\gamma^2 - 1} = \sqrt{\frac{1}{1-\beta^2} - 1} = \frac{\beta}{\sqrt{1-\beta^2}} = \beta\gamma. \tag{7.2.4}$$

Hence

$$\tanh\omega = \frac{\sinh\omega}{\cosh\omega} = \frac{\beta\gamma}{\gamma} = \beta \implies \omega = \operatorname{artanh}\beta. \tag{7.2.5}$$

Using these relations we have

$$x' = \gamma x - \beta\gamma ct = \cosh(\omega)x - \sinh(\omega)ct \tag{7.2.6}$$

$$ct' = \gamma ct - \beta\gamma x = \cosh(\omega)ct - \sinh(\omega)x \tag{7.2.7}$$

Notice the similarity to the definition of x and y in polar coordinates. We can write this in a matrix form as

$$\begin{pmatrix} ct' \\ x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} \cosh\omega & -\sinh\omega & 0 & 0 \\ -\sinh\omega & \cosh\omega & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} ct \\ x \\ y \\ z \end{pmatrix}. \tag{7.2.8}$$

From this we identify the matrix that represents the Lorentz transformation:

$$\Lambda := \begin{pmatrix} \gamma & -\beta\gamma & 0 & 0 \\ -\beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} \cosh\omega & -\sinh\omega & 0 & 0 \\ -\sinh\omega & \cosh\omega & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{7.2.9}$$

This allows us to compactly write the Lorentz transformation of x^μ as

$$x'^\mu = \Lambda^\mu{}_\nu x^\nu. \tag{7.2.10}$$

Notation 7.2.11 — Einstein Summation Convention In this course we make use of the **Einstein summation convention**. If an index is repeated twice, once as an upper index and once as a lower index, then it is automatically summed over.
Latin indices (i, j, k, etc.) take the values 1, 2, and 3.
Greek indices (μ, ν, σ, etc.) take the values 0, 1, 2, 3.
For example,

$$\boldsymbol{a} \cdot \boldsymbol{b} = a_i b^i = a_1 b^1 + a_2 b^2 + a_3 b^3. \tag{7.2.12}$$

$$a \cdot b = a_\mu b^\mu a_0 b^0 + a_1 b^1 + a_2 b^2 + a_3 b^3 = a_0 b^0 + \boldsymbol{a} \cdot \boldsymbol{b}. \tag{7.2.13}$$

This last example is the scalar product of two four-vectors, we will see this in the next chapter. The distinction between upper and lower indices will be important later.

We can show that this is really an equivalent way of writing the Lorentz transformation by expanding the sum:

$$x'^\mu = \Lambda^\mu{}_\nu x^\nu = \Lambda^\mu{}_0 x^0 + \Lambda^\mu{}_1 x^1 + \Lambda^\mu{}_2 x^2 + \Lambda^\mu{}_3 x^3 = \Lambda^\mu{}_0 ct + \Lambda^\mu{}_1 x + \Lambda^\mu{}_2 y + \Lambda^\mu{}_3 z. \tag{7.2.14}$$

Now consider specific values of μ:

$$\begin{aligned} x'^0 &= ct' = \Lambda^0{}_0 ct + \Lambda^0{}_1 x + \Lambda^0{}_2 y + \Lambda^0{}_3 z = \gamma ct' - \beta\gamma x + 0y + 0z = \gamma ct - \beta\gamma x \\ x'^1 &= x' = \Lambda^1{}_0 ct + \Lambda^1{}_1 x + \Lambda^1{}_2 y + \Lambda^1{}_3 z = -\beta\gamma ct' + \gamma x + 0y + 0z = \gamma x - \beta\gamma ct \\ x'^2 &= y' = \Lambda^2{}_0 ct + \Lambda^1{}_2 x + \Lambda^2{}_2 y + \Lambda^2{}_3 z = 0ct' + 0x + 1y + 0z = y \\ x'^3 &= z' = \Lambda^3{}_0 ct + \Lambda^3{}_2 x + \Lambda^3{}_2 y + \Lambda^3{}_3 z = 0ct' + 0x + 0y + 1z = z \end{aligned}$$

So we have recovered the Lorentz transformations between two frames in the standard configuration.

7.2.3 Space-Time Interval

Consider the proper time elapsed for an infinitesimal segment of a world line, $\mathrm{d}\tau$:

$$
\begin{aligned}
c\,\mathrm{d}\tau &= \frac{c\,\mathrm{d}t}{\gamma} && (7.2.15)\\
&= c\,\mathrm{d}t\sqrt{1-\frac{u^2}{c^2}} && (7.2.16)\\
&= c\,\mathrm{d}t\sqrt{1-\frac{1}{c^2}\left(\frac{\mathrm{d}x}{\mathrm{d}t}\right)^2-\frac{1}{c^2}\left(\frac{\mathrm{d}y}{\mathrm{d}t}\right)^2-\frac{1}{c^2}\left(\frac{\mathrm{d}z}{\mathrm{d}t}\right)^2} && (7.2.17)\\
&= \sqrt{c^2\,\mathrm{d}t^2-\mathrm{d}x^2-\mathrm{d}y^2-\mathrm{d}z^2}. && (7.2.18)
\end{aligned}
$$

It follows that

$$c^2\,\mathrm{d}\tau^2 = c^2\,\mathrm{d}t^2-\mathrm{d}x^2-\mathrm{d}y^2-\mathrm{d}z^2, \tag{7.2.19}$$

or, in a more condensed notation

$$\mathrm{d}s^2 := c^2\,\mathrm{d}\tau^2 = g_{\mu\nu}\,\mathrm{d}x^\mu\,\mathrm{d}x^\nu. \tag{7.2.20}$$

Here $\mathrm{d}x^\mu = (c\,\mathrm{d}t, \mathrm{d}x, \mathrm{d}y, \mathrm{d}z)$, and g is the **metric tensor**, specifically the **Minkowski metric**, defined as

$$g := \begin{pmatrix} 1 & 0 & 0 & 0\\ 0 & -1 & 0 & 0\\ 0 & 0 & -1 & 0\\ 0 & 0 & 0 & -1 \end{pmatrix}. \tag{7.2.21}$$

The Minkowski metric in particular is often written as η, and g is saved for more general metrics.

The metric tensor plays the role of the Kronecker delta, δ, in the normal dot product, $\boldsymbol{a}\cdot\boldsymbol{b} = \delta_{ij}a^ib^j = a_ib^j$, we will come to the raising and lowering of indices via the metric tensor in the next chapter. We can think of the Kronecker delta as the metric tensor for Euclidean space.

! There are two contradictory conventions when it comes to the metric, the one we have been using is that the time part is positive and the other three parts are negative, or succinctly $(+---)$. The other convention is that the time part is negative and the other three parts are positive, or $(-+++)$. Make sure to check which convention is being used.

7.2.4 Invariance of the Space-Time Interval

The quantity $\mathrm{d}s^2$ defined above is the invariant that we have been searching for in this chapter. In general any interval, s, defined by $s^2 = g_{\mu\nu}x^\mu x^\nu$ is invariant. This is

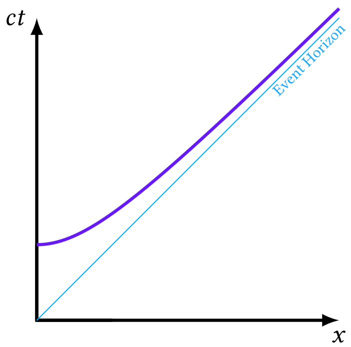


Figure 7.1: A line of constant space-time interval, $c^2t^2 - x^2 = \text{constant}$. Notice the gradient is less than 1 and so this represents an object moving faster than the speed of light.

fairly easy to show for frames in the standard configuration:

$$\begin{aligned}
s'^2 &= g_{\mu\nu}x'^\mu x'^\nu && (7.2.22)\\
&= (ct')^2 - x'^2 - y'^2 - z'^2 && (7.2.23)\\
&= \gamma^2(ct - \beta x)^2 - \gamma^2(x - \beta ct)^2 - y^2 - z^2 && (7.2.24)\\
&= \gamma^2(ct)^2 - \gamma^2\beta ctx + \gamma^2\beta^2x^2 - \gamma^2x^2 + \gamma^2\beta xct - \gamma^2\beta^2(ct) - y^2 - z^2 && (7.2.25)\\
&= (\gamma^2 - \gamma^2\beta^2)(ct)^2 - (\gamma^2 - \gamma^2\beta) - y^2 - z^2 && (7.2.26)\\
&= (ct)^2 - x^2 - y^2 - z^2 && (7.2.27)\\
&= s^2. && (7.2.28)
\end{aligned}$$

Here we have used

$$\gamma^2 - \gamma^2\beta^2 = \frac{1}{1 - \frac{u^2}{c^2}} - \frac{u^2/c^2}{1 - \frac{u^2}{c^2}} = \frac{1 - \frac{u^2}{c^2}}{1 - \frac{u^2}{c^2}} = 1. \tag{7.2.29}$$

Similarly one can show that the interval between any two points is invariant.

Note that we say the interval is *invariant*, and not *constant*. The value of the interval can, and will, change along the world line. It's just that at a given point in space-time the value of a space-time interval doesn't depend on the frame from which we measure it. In fact, the line $c^2t^2 - x^2 = s^2 = \text{constant}$ is a hyperbola and cannot be a world line as it would require faster than light travel. See Figure 7.1.

7.2.5 Interval Classification

One place where the invariant space-time interval differs from the vector length considered in Section 7.1 is that s^2 can be negative. We typically classify space-time intervals into one of three types:

$s^2 > 0$ We say that the interval is **time-like**. We have $ct > |\boldsymbol{x}|$.

$s^2 = 0$ We say that the interval is **light-like**. We have $ct = |\boldsymbol{x}|$.

$s^2 < 0$ We say that the interval is **space-like**. We have $ctM|\boldsymbol{x}|$.

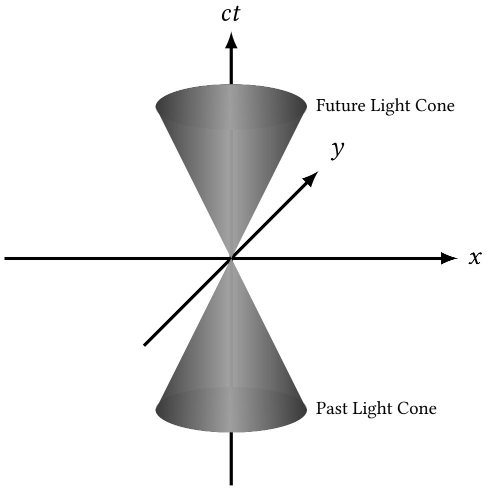


Figure 7.2: Light cones drawn in two spatial dimensions and time.

Graphically we separate space time into two cones (actually four-dimensional cones, although we usually only draw them in three dimensions) and an area outside of the cones. Both cones have their points at the origin, one extends forward in time and the other back. The edges of each cone are at 45°, i.e. the edges are the world lines of photons.

Events in the cones have time-like space intervals. Events separated by a time-like interval cannot be simultaneous in any reference frame, but we can find a frame in which their spatial separation is zero, in which measured times are proper times.

Similarly events outside the cones have space-like intervals. Events separated by a space-like interval have a reference frame where their separation is purely spatial, and we can find a proper distance between them. We can also find a frame such that the temporal order of the events is reversed.

Events on the cone have light-like intervals. Only photons, and other massless particles, can travel on the light cone.

7.2.5.1 Causality

The idea that cause comes before effect is called **causality**. For causality to be preserved it must be that information cannot travel faster than light. Essentially, the order of events can change a bit, but not so much that causality is violated. An observer at the origin can only send information to points in the future light cone, and can only receive information from the past light cone. This means that the observer can only effect changes in the future light cone and can only be effected by events in the past light cone.

Eight

Four-Vectors

Definition 8.0.1 — Four-Vector A **four-vector** is any four dimensional vector whose components transform by the Lorentz transformation.

8.1 Position

The prototypical four-vector is the position, or **four-position**. Given a spatial position, $\boldsymbol{r} = (x, y, z)$, and a time, t, the position four-vector is

$$x^\mu = (x^0, x^1, x^2, x^3) = (ct, x, y, z) = (ct, \boldsymbol{r}). \tag{8.1.1}$$

As usual Greek indices go from 0 to 3 and we identify the 0th component with time. When discussing four-vectors we commonly call "normal" vectors like $\boldsymbol{r}$ three-vectors.

Recall that the magnitude of a three-vector is invariant under rotations. The magnitude is simply the square root of the dot product. In index notation we write the dot product as

$$|\boldsymbol{r}|^2 = \boldsymbol{r} \cdot \boldsymbol{r} = x_i x^j = \delta_{ij} x^i x^j = x^2 + y^2 + z^2, \tag{8.1.2}$$

or in matrix form:

$$|\boldsymbol{r}|^2 = \boldsymbol{r} \cdot \boldsymbol{r} = \boldsymbol{r}^\top \boldsymbol{r} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}^\top \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x & y & z \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = x^2 + y^2 + z^2 \tag{8.1.3}$$

We define the length of a four vector in a similar way, replacing Latin indices with Greek, and the Kronecker delta with the Minkowski metric tensor:

$$x^2 = x \cdot x = x_\mu x^\mu = g_{\mu\nu} x^\nu x^\mu = (ct)^2 - x^2 - y^2 - z^2 \tag{8.1.4}$$

or in matrix form

$$x^2 = x \cdot x = g \begin{pmatrix} ct \\ x \\ y \\ z \end{pmatrix}^\top \begin{pmatrix} ct \\ x \\ y \\ z \end{pmatrix} = \begin{pmatrix} ct & -x & -y & -z \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = (ct)^2 - x^2 - y^2 - z^2. \tag{8.1.5}$$

Here we have used the matrix form of the metric tensor:

$$g = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \tag{8.1.6}$$

Looking at these we can make the identification

$$x_\mu = g_{\mu\nu} x^\nu = (x_0, x_1, x_2, x_3) = (ct, -x, -y, -z) = (ct, -\boldsymbol{r}). \tag{8.1.7}$$

Formally x^μ is a **contravariant** four-vector, and x_μ, is a **covariant** four-vector. We can think of the act of raising and lowering indices by transposing and multiplying by the metric[1].

[1] if the metric is not orthogonal and symmetric then sometimes you will need to multiply by the inverse of the metric

R The same notation, x^μ, is often used to talk about a four-vector and the components of a four-vector, this can cause conceptual difficulties but is rarely an issue in calculations as we mostly work with components.

8.2 Scalar Product

Having defined the length of a four-vector the obvious thing to do is to generalise to a scalar product, an analogue of dot-product for three-vectors.

Definition 8.2.1 — Scalar Product The **scalar product** of the four-vectors a^μ and b^μ is defined as

$$a \cdot b = a_\mu b^\mu = g_{\mu\nu} a^\nu b^\mu = a^0 b^0 - a^1 b^1 - a^2 b^2 - a^3 b^3 = a^0 b^0 - \boldsymbol{a} \cdot \boldsymbol{b}. \tag{8.2.2}$$

Here $a^\mu = (a^0, a^1, a^2, a^3) = (a^0, \boldsymbol{a})$, and similarly for b^μ.

Theorem 8.2.3 — Lorentz Invariance of Scalar Products. The scalar product of two four-vectors is Lorentz invariant.

Proof. Let a^μ and b^μ be four-vectors in some frame S. The same four-vectors in frame S have the components

$$a'^0 = \gamma(a^0 - \beta a^1), \quad a'^1 = \gamma(a^1 - \beta a^0), \quad a'^2 = a^2, \quad \text{and} \quad a'^3 = a^3, \tag{8.2.4}$$

and similar for b^μ.
Hence,

$$\begin{aligned} a \cdot b &= a'_\mu b'^\mu && (8.2.5) \\ &= g_{\mu\nu} a'^\nu b'^\mu && (8.2.6) \\ &= a'^0 b'^0 - a'^1 b'^1 - a'^2 b'^2 - a'^3 b'^3 && (8.2.7) \\ &= \gamma^2 (a^0 - \beta a^1)(b^0 - \beta b^1) - \gamma^2 (a^1 - \beta a^0)(b^1 - \beta b^0) - a^2 b^2 - a^3 b^3 && (8.2.8) \\ &= \gamma^2 [a^0 b^0 (1 - \beta^2) - a^1 b^1 (\beta^2 - 1)] - a^2 b^2 - a^3 b^3 && (8.2.9) \\ &= a^0 b^0 - a^1 b^1 - a^2 b^2 - a^3 b^3. && (8.2.10) \end{aligned}$$

Here we have used

$$\gamma^2(1-\beta^2) = \frac{1}{1-\beta^2}(1-\beta^2) = 1, \tag{8.2.11}$$

and hence $\gamma^2(\beta^2-1) = -\gamma^2(1-\beta^2) = -1$. □

Corollary 8.2.12 — Inverse Lorentz Transformation The Lorentz transformation is orthogonal. That is, $\Lambda^{-1} = \Lambda^\top$.

Proof. Consider the scalar product of a^μ and b^μ. In index notation we have

$$\begin{aligned}
a \cdot b &= a'_\mu b'^\mu &(8.2.13)\\
&= g_{\mu\nu} a'^\nu b'^\mu &(8.2.14)\\
&= g_{\mu\nu} \Lambda^\nu{}_\sigma a^\sigma \Lambda^\mu{}_\rho b^\rho &(8.2.15)\\
&= \Lambda_{\mu\sigma} a^\sigma \Lambda^\mu{}_\rho b^\rho &(8.2.16)\\
&= g_{\sigma\kappa} \Lambda_\mu{}^\kappa a^\sigma \Lambda^\mu{}_\rho b^\rho &(8.2.17)\\
&= g_{\sigma\kappa} \Lambda_\mu{}^\kappa \Lambda^\mu{}_\rho a^\sigma b^\rho &(8.2.18)\\
&= g_{\sigma\kappa} (\Lambda^\top \Lambda)^\kappa{}_\rho a^\sigma b^\rho &(8.2.19)
\end{aligned}$$

By the invariance of scalar products, Theorem 8.2.3, we know that this must be equal to

$$a \cdot b = a_\rho b^\rho = g_{\sigma\rho} a^\sigma b^\rho. \tag{8.2.20}$$

So we can identify

$$g_{\sigma\kappa} (\Lambda^\top \Lambda)^\kappa{}_\rho a^\sigma b^\rho = g_{\sigma\rho} a^\sigma b^\rho. \tag{8.2.21}$$

This implies that $g_{\sigma\kappa}(\Lambda^\top\Lambda)^\kappa{}_\rho = g_{\sigma\rho}$. That is, the action of $\Lambda^\top\Lambda^\kappa{}_\rho$ is to exchange the index κ for the index ρ when κ is summed over. This property is unique to the Kronecker delta and so we must have $\Lambda^\top\Lambda^\kappa{}_\rho = \delta^\kappa{}_\rho$ which means that $\Lambda^\top\Lambda = I$, the identity. Therefore $\Lambda^\top = \Lambda^{-1}$. □

8.3 Four-Velocity

We have already seen how the three-velocity, $\boldsymbol{u} = \dot{\boldsymbol{r}}$ transforms in Section 6.2. Can we define a similar quantity that transforms as a four-vector (i.e. under Lorentz transformations)? Yes, we can. Any operation that is Lorentz invariant can be combined with a four-vector to give another four-vector. Recall that the infinitesimal line element is given by

$$\mathrm{d}s^2 = c^2\,\mathrm{d}t^2 - \mathrm{d}x^2 - \mathrm{d}y^2 - \mathrm{d}z^2 = c^2\,\mathrm{d}\tau^2. \tag{8.3.1}$$

Since $\mathrm{d}s^2$ is invariant and c is invariant by the postulates this means that $\mathrm{d}\tau$ is an invariant interval. Hence differentiating with respect to τ is invariant.

This means that

$$u^\mu := \frac{\mathrm{d}x^\mu}{\mathrm{d}\tau} \tag{8.3.2}$$

is a four-vector, and it is this quantity that we define as the **four-velocity**. In an inertial frame, S, the four-velocity has components

$$u^\mu = \frac{\mathrm{d}}{\mathrm{d}\tau}(ct, \boldsymbol{r}) = \left(c\frac{\mathrm{d}t}{\mathrm{d}\tau}, \frac{\mathrm{d}\boldsymbol{r}}{\mathrm{d}\tau}\right) = \left(\gamma c, \gamma\frac{\mathrm{d}\boldsymbol{r}}{\mathrm{d}t}\right) = (\gamma c, \gamma\boldsymbol{u}). \tag{8.3.3}$$

Here we have used $t = \gamma\tau$ and the standard three-velocity, $\boldsymbol{u} = \dot{\boldsymbol{r}}$.

Since the four-velocity is a four velocity it transforms under a Lorentz transformation as

$$u'^0 = \gamma(v)[u^0 - \beta u^1] \tag{8.3.4}$$
$$u'^1 = \gamma(v)[u^1 - \beta u^0] \tag{8.3.5}$$
$$u'^2 = u^2 \tag{8.3.6}$$
$$u'^3 = u^3. \tag{8.3.7}$$

where v is the speed of the S' frame with respect to the S frame.

The scalar product of the four-velocity is invariant. We are therefore free to choose a frame to evaluate it. The simplest choice is the rest frame of the particle, where $\boldsymbol{u} = \boldsymbol{0}$ and $\gamma = 1$:

$$u^2 = u \cdot u = u_\mu u^\mu = g_{\mu\nu}u^\mu u^\nu = \gamma^2(u)(c^2 - \boldsymbol{u}\cdot\boldsymbol{u}) = c^2. \tag{8.3.8}$$

Notice that we are using γ in two ways here, as a function of the speed of a particle and as a function of the relative speed of the frames. If we were to calculate the scalar product in a different frame we would get

$$\begin{aligned} u^2 = u \cdot u = u'_\mu u'^\mu &= \gamma^2(u)(c^2 - \boldsymbol{u}\cdot\boldsymbol{u}) = \gamma^2(u)(c^2 - \boldsymbol{u}^2) \\ &= \gamma^2(u)(1 - \frac{\boldsymbol{u}^2}{c^2}) = \gamma^2(u)c^2\frac{1}{\gamma^2(u)} = c^2, \end{aligned} \tag{8.3.9}$$

so the result is the same, it is just a bit more work to get.

8.4 Four-Acceleration

Just as the four-velocity is defined as the proper time derivative of the position we can define the **four-acceleration** as the second proper time derivative of the position:

$$A^\mu := \frac{\mathrm{d}^2x^\mu}{\mathrm{d}\tau^2} = \frac{\mathrm{d}u^\mu}{\mathrm{d}\tau}. \tag{8.4.1}$$

This has components

$$A^\mu = \frac{\mathrm{d}}{\mathrm{d}\tau}(\gamma c, \gamma\boldsymbol{u}) = \gamma\frac{\mathrm{d}}{\mathrm{d}t}(\gamma c, \gamma\boldsymbol{u}). \tag{8.4.2}$$

Using Equation (6.6.10) we have

$$\frac{\mathrm{d}\gamma}{\mathrm{d}t} = \frac{u}{c^2}\frac{\mathrm{d}u}{\mathrm{d}t}\gamma^3 = \gamma^3\frac{\boldsymbol{u}\cdot\boldsymbol{a}}{c^2} \tag{8.4.3}$$

where $\boldsymbol{a} = \dot{\boldsymbol{u}}$ is the three-acceleration. We then have

$$\frac{\mathrm{d}}{\mathrm{d}t}(\gamma \boldsymbol{u}) = \gamma \boldsymbol{a} + \gamma^3 \frac{\boldsymbol{u} \cdot \boldsymbol{a}}{c^2} \boldsymbol{u}. \tag{8.4.4}$$

From this we get that the four-acceleration in some inertial frame is

$$A^\mu = \left(\gamma^4 \frac{\boldsymbol{u} \cdot \boldsymbol{a}}{c}, \gamma^2 \boldsymbol{a} + \gamma^4 \frac{\boldsymbol{u} \cdot \boldsymbol{a}}{c^2} \boldsymbol{u} \right). \tag{8.4.5}$$

Since the four-acceleration is a four-vector it transforms as we would expect under Lorentz transformations.

One frame that we may want to consider is the ICMF frame, in which the velocity of the particle is zero and $\boldsymbol{a}_0$ is the proper acceleration. In this frame the four-acceleration and four-velocity are

$$A^\mu_{\text{ICMF}} = (0, \boldsymbol{a}_0), \qquad \text{and} \qquad u^\mu_{\text{ICMF}} = (c, \boldsymbol{0}). \tag{8.4.6}$$

Clearly in this frame the four-acceleration and four-velocity are orthogonal, in the usual sense that

$$A_{\text{ICMF}} \cdot u_{\text{ICMF}} = 0. \tag{8.4.7}$$

Since this is a scalar product, and hence invariant, in any inertial frame the four-acceleration and four-velocity are orthogonal. Calculating in the ICMF frame again we have

$$A_\mu A^\mu = -\boldsymbol{a}_0 \cdot \boldsymbol{a}_0 = -a_0^2 < 0. \tag{8.4.8}$$

As well as this $u_\mu u^\mu = c^2$. We conclude that four-velocity is time-like and four-acceleration space-like.

8.5 Four-Momentum

In Newtonian mechanics the momentum, defined as $m\boldsymbol{v}$, is conserved in the absence of external forces. However, this is not so in relativistic mechanics. Instead we find that the **four-momentum**, defined as

$$p^\mu = m u^\mu = m \frac{\mathrm{d}x^\mu}{\mathrm{d}\tau}, \tag{8.5.1}$$

is a conserved quantity in the absence of external forces. This has components

$$p^\mu = (\gamma m c, \gamma m \boldsymbol{u}) \tag{8.5.2}$$

Consider the scalar product of the four-momentum with itself, this is easiest to compute in the rest frame of the particle where we have

$$p_\mu p^\mu = m^2 c^2 \tag{8.5.3}$$

since $\gamma = 1$ and $\boldsymbol{u} = \boldsymbol{0}$ in this frame. Comparing this result to Einstein's famous $E_{\text{rest}} = mc^2$, where E_{rest} is the rest energy, we see that $p^0 c = E_{\text{rest}}/c$. For a particle not at rest we have $p^0 c = \gamma m c^2 = E$ and $p^0 = E/c$, where E is the **relativistic energy**. Hence

$$p^\mu = \left(\frac{E}{c}, \boldsymbol{p} \right), \tag{8.5.4}$$

where $\boldsymbol{p} := \gamma m \boldsymbol{u}$ is the **relativistic momentum**. We will discuss four-momentum, and the relativistic energy more in the next chapter.

8.5.1 Notes

The definition of the relativistic momentum differs from the Newtonian definition by a factor of γ, $\boldsymbol{p} = \gamma m \boldsymbol{u}$. In the non-relativistic limit we have $\gamma \to 1$ and so recover the Newtonian momentum, $m\boldsymbol{u}$. This relativistic three-momentum *is* conserved in the absence of forces.

Some sources will speak of a "relativistic mass", which is given by γm, this makes the modification of the relativistic energy and momentum formulas above more obvious, but is not physically meaningful and is somewhat outdated since it is an attempt to fix classical ideas to work with relativity rather than using purely relativistic concepts. The relevant relativistic quantity is the relativistic energy, which, when using units where $c = 1$, is $E = \gamma m$, and this concept is more meaningful.

8.5.2 Lorentz Transformation of Four-Momentum

The four-momentum is a four-vector and so transforms as

$$p'^{\mu} = \Lambda^{\mu}{}_{\nu} p^{\nu}, \tag{8.5.5}$$

Writing out the components for two frames in the standard configuration, and using $\boldsymbol{P}$ for the relativistic three-momentum, we get

$$\frac{E'}{c} = \gamma\frac{E}{c} - \gamma\beta P^1, \tag{8.5.6}$$

$$P'^1 = \gamma P^1 - \gamma\beta\frac{E}{c}, \tag{8.5.7}$$

$$P'^2 = P^2, \tag{8.5.8}$$

$$P'^3 = P^3. \tag{8.5.9}$$

This should be no suprise at this point.

We can obtain a useful form of the Lorentz transformation if we split $\boldsymbol{P}$ into two components, $\boldsymbol{P}_{\parallel}$ and $\boldsymbol{P}_{\perp}$, which are parallel and perpendicular to the velocity, $\boldsymbol{v}$, of S' with respect to S. We get

$$\frac{E'}{c} = \gamma\frac{E}{c} - \gamma\beta P_{\parallel} \tag{8.5.10}$$

$$P'_{\parallel} = \gamma P_{\parallel} - \gamma\beta\frac{E}{c}, \tag{8.5.11}$$

$$\boldsymbol{P}'_{\perp} = \boldsymbol{P}_{\perp}. \tag{8.5.12}$$

This form of the Lorentz transformation is called a **Lorentz boost**. We often talk about "boosting" between frames when transforming between inertial frames.

Nine

Relativistic Dynamics

9.1 Relativistic Energy

In the previous chapter we stated that the zeroth component of the four-momentum is the energy, or rather E/c. We based this on the assumption that $E = mc^2$ is valid. If this interpretation is correct then we should be able to recover Newtonian mechanics in the non-relativistic limit. In particular when $u \ll c$ we can expand the γ factor as a binomial expansion in u^2/c^2:

$$\gamma = \left(1 + \frac{u^2}{c^2}\right)^{-1/2} \approx 1 + \frac{1}{2}\frac{u^2}{c^2} + \frac{3}{8}\frac{u^4}{c^4} + \cdots, \tag{9.1.1}$$

having used the binomial expansion

$$(1 + x)^p \approx 1 + px + \frac{1}{2}(p - 1)px^2 + \cdots. \tag{9.1.2}$$

We then have

$$p^0 c = \gamma(u)mc^2 = mc^2 + \frac{1}{2}\frac{u^2}{c^2}mc^2 + \cdots \approx mc^2 + \frac{1}{2}mu^2. \tag{9.1.3}$$

We can recognise the second term as the kinetic energy of a slow moving particle. The first term, mc^2, we interpret as the **rest energy** due to the mass. There is no Newtonian equivalent but since we are almost always considering energy only defined up to a constant it does not appear in Newtonian mechanics. Putting this together we find that at relativistic speeds the total energy, E, the **relativistic kinetic energy**, T, and the rest energy, $E_{\mathrm{rest}} = mc^2$, are related by

$$E = \gamma mc^2 = mc^2 + T = E_{\mathrm{rest}} + T \tag{9.1.4}$$

and

$$T \coloneqq (\gamma - 1)mc^2. \tag{9.1.5}$$

We can think of the relativistic kinetic energy as keeping all of the higher order terms that we discard in the non-relativistic limit.

9.2 Energy-Momentum Relation

The four-momentum is

$$p^\mu = \left(\frac{E}{c}, \boldsymbol{p}\right) = (\gamma mc, \gamma m\boldsymbol{u}). \tag{9.2.1}$$

Computing the square of this in some arbitrary inertial frame we have

$$p^2 = \frac{E^2}{c^2} - \boldsymbol{p}^2. \tag{9.2.2}$$

Evaluating this in the rest frame of the particle we have

$$p^2 = m^2c^2 \tag{9.2.3}$$

since $\boldsymbol{p} = \boldsymbol{0}$ and $\gamma = 1$ in the particles rest frame. Combining these we get the relativistic **energy-momentum relation**, relating the total energy, E, momentum, p, and mass, m, of a particle:

$$E^2 = m^2c^4 + p^2c^2. \tag{9.2.4}$$

This shows how energy and momentum are closely linked in special relativity, even more so than in Newtonian mechanics.

Considering the non-relativistic limit and keeping only the first term we have

$$\begin{aligned} E &= \sqrt{m^2c^4 + \boldsymbol{p}^2c^2} && (9.2.5)\\ &= mc^2\sqrt{1 + \frac{\boldsymbol{p}^2c^2}{m^2c^4}} && (9.2.6)\\ &= mc^2\left(1 + \frac{\boldsymbol{p}^2}{m^2c^2}\right)^{1/2} && (9.2.7)\\ &\approx mc^2 + \frac{1}{2}mc^2\frac{\boldsymbol{p}^2}{m^2c^2} && (9.2.8)\\ &= mc^2 + \frac{\boldsymbol{p}^2}{2m}. && (9.2.9) \end{aligned}$$

So we again recover the total energy is the rest energy plus kinetic energy, this time with the kinetic energy of a slow particle in terms of the momentum, $\boldsymbol{p}^2/(2m)$.

9.2.1 Quantum Mechanics

In quantum mechanics we replace the momentum, $\boldsymbol{p}$, with the momentum operator, $\hat{\boldsymbol{p}} \coloneqq -i\hbar\nabla$. Doing so in the non-relativistic limit with energy $\boldsymbol{p}^2/(2m)$ (i.e. neglecting the constant rest energy) we get

$$\hat{H} = \frac{\hbar^2}{2m}\nabla^2. \tag{9.2.10}$$

From this we can get the Schrödinger equation for a free particle by demanding that

$$(\hat{H} - E)\psi(x, t) = 0, \tag{9.2.11}$$

or as its more normally written in terms of an eigenvalue equation

$$\hat{H}\psi(x, t) = E\psi(x, t). \tag{9.2.12}$$

If instead we use the relativistic form, $E^2 = p^2c^2 + m^2c^4$, and also recognising $\hat{E} = i\hbar\partial/\partial t$ we get

$$(\hat{E}^2 - c^2\hat{\boldsymbol{p}}^2 - m^2c^2)\psi(x, t) = \left(-\hbar^2\frac{\mathrm{d}^2}{\mathrm{d}t^2} + \hbar^2\nabla^2 - m^2c^4\right)\psi(x, t) = 0, \tag{9.2.13}$$

or, as its usually written,

$$\left(\frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \nabla^2 + \frac{m^2c^2}{\hbar^2}\right)\psi(x,t) = 0. \tag{9.2.14}$$

This is the **Klein–Gordon equation**. It applies to spinless relativistic particles. It can be written more compactly as

$$(\Box^2 + \mu^2)\psi(x,t) = 0 \tag{9.2.15}$$

where[1] $\mu = mc/\hbar$ and $\Box^2$ is the d'Alembert operator, which can be thought of as the four-vector version of the Laplacian and is defined as

[1] in natural units, where $c = \hbar = 1$, this simplifies to $\mu = m$.

$$\Box^2 := g^{\mu\nu}\partial_\mu\partial_\nu = \frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \frac{\partial^2}{\partial x^2} - \frac{\partial^2}{\partial y^2} - \frac{\partial^2}{\partial z^2} = \frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \nabla^2 \tag{9.2.16}$$

where $\partial_\mu := \partial/\partial x^\mu$ represents differentiation with respect to x^μ.

9.2.2 Quantities as Functions of Momentum

The quantities γ, β, and $\beta\gamma$, which turn up frequently in special relativity, can all be rewritten using the energy-momentum relation:

$$\boldsymbol{u} = c^2\frac{\boldsymbol{p}}{E} \implies \beta = \frac{pc}{E}, \tag{9.2.17}$$

$$E = \gamma mc^2 \implies \gamma = \frac{E}{mc^2}, \tag{9.2.18}$$

$$p = \beta\gamma mc \implies \beta\gamma = \frac{pc}{mc^2} = \frac{p}{mc} \tag{9.2.19}$$

9.3 Zero Mass Particles

A problem arises with our definition for massless particles. Since massless particles travel at the speed of light γ becomes $1/0$, and $m = 0$ so γm is undefined and this appears all components of the four momentum. We circumvent this issue by just assuming that the relativistic energy-momentum relation still holds when $m = 0$, in which case it reduces to

$$E = pc. \tag{9.3.1}$$

This relation should hold as we can derive the same relation in quantum mechanics by considering $E = \hbar\omega$ and $p = \hbar k$ and $\omega = 2\pi c/\lambda = kc$ from which we get $E = \hbar kc = pc$. We can also derive it in electromagnetism by considering the radiation pressure of light, although this derivation is a bit more involved[2]. This demonstrates nice agreement between special relativity, quantum mechanics, and electromagnetism.

[2] see the section on radiation pressure in the electromagnetism course for details.

9.4 Relativistic Newton's Second Law

Newton's second law, $\dot{\boldsymbol{p}} = \boldsymbol{F}$, is a postulate defining the force. It would be nice if we could derive a similar statement from special relativity, and indeed we can. The obvious definition being to define the **four-force** as the proper time derivative of the four momentum:

$$F^{\mu} := \frac{\mathrm{d}p^{\mu}}{\mathrm{d}\tau} = ma^{\mu} \tag{9.4.1}$$

where p^{μ} is the four momentum, τ is proper time, m is the mass and a^{μ} is the for-acceleration. Using $\tau = t/\gamma$ we can write this as

$$F^{\mu} = \gamma\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{E}{c}, \boldsymbol{p}\right) = \gamma\left(\frac{P}{c}, \boldsymbol{F}\right) \tag{9.4.2}$$

where $P = \dot{E}$ is the power and $\boldsymbol{F} = \dot{\boldsymbol{p}}$ is the three-force.

So, the zeroth component of the four-force is related to the power and the spatial components, $\gamma\dot{p}^i$, reduce to $\dot{p}^i$ in the non-relativistic limit so we recover Newton's second law.

In the ICMF we have

$$P = \frac{\mathrm{d}E}{\mathrm{d}t} = 0, \qquad \text{and} \qquad \boldsymbol{F}_0 = \frac{\partial\boldsymbol{p}}{\partial t} = m\boldsymbol{a}_0 \tag{9.4.3}$$

where $\boldsymbol{F}_0$ is the **proper-force** and $\boldsymbol{a}_0$ is the proper acceleration.

9.5 Conservation of Four-Momentum

Consider a collision between two particles, A and B, in the absence of external forces. Let p_A^{μ} and p_B^{μ} be their momenta before the collision and q_a^{μ} and q_B^{μ} their momenta after the collision. The components of any one of these four momenta in two inertial frames are related by a Lorentz transformation, so taking p_A^{μ} as an example we have

$$p_A'^0 = \gamma p_A^0 - \beta\gamma p_A^1, \qquad \text{and} \qquad p_A'^1 = \gamma p_A^1 - \beta\gamma p_A^0. \tag{9.5.1}$$

Considering the total momentum we then have

$$p_A'^1 + p_B'^1 = \gamma(p_A^1 + p_B^1) - \beta\gamma(p_A^0 + p_B^0) \tag{9.5.2}$$

before the collision and

$$q_A'^1 + q_B'^1 = \gamma(q_A^1 + q_B^1) - \beta\gamma(q_A^0 + q_B^0) \tag{9.5.3}$$

after the collision. Assuming that the relativistic three-momentum is conserved each component of the total momentum, $p_A^i + p_B^i$, must be conserved and conservation must hold in any frame, therefore we have

$$p_A^1 + p_B^1 = q_A^1 + q_B^1, \qquad \text{and} \qquad q_A'^1 + q_B'^1 = q_A^1 + q_B^1. \tag{9.5.4}$$

This can only hold in general if we also have

$$p_A^0 + p_B^0 = q_A^0 + q_B^0. \tag{9.5.5}$$

This proves that requiring the relativistic-three momentum to be conserved and the four-momentum to transform under the Lorentz transformation as a four vector implies that the zeroth component of the four-momentum is conserved. We have already seen that the zeroth component of the four-momentum is closely related to the energy and hence this is simply conservation of energy.

The relation between conservation of momentum and conservation should not be too surprising given the relation between energy and momentum in special relativity. From the point of view of Noether's theorem conservation of momentum arises from invariance under spatial translation, and conservation of energy from invariance under temporal translation. In special relativity we combine the idea of spatial and temporal translations into space-time translations, or four-translations, and hence combine the conservation of energy and momentum into the translation four-momentum.

Ten

Particle Decays

10.1 Two-Body Decays

Consider an unstable particle of mass M decaying into two daughter particles of masses m_i. We work in the rest frame of the parent particle for simplicity. In this frame the four-momentum of the parent particle is $P = (MC, \mathbf{0})$ and the four-momenta of the daughter particles are $p_i = (E_i/c, \boldsymbol{p_i})$.

Conservation of four momentum gives us

$$P^\mu = p_1^\mu + p_2^\mu \tag{10.1.1}$$
$$Mc^2 = E_1 + E_2 \tag{10.1.2}$$
$$\mathbf{0} = \boldsymbol{p}_1 + \boldsymbol{p}_2. \tag{10.1.3}$$

From this we see that the two daughter particles must have equal and opposite momenta, $\boldsymbol{p} = \boldsymbol{p}_1 = -\boldsymbol{p}_2$. The relativistic energy-momentum relation tells us that

$$E_i - \sqrt{m_i^2c^4 + p^2c^2} \geq m_1c^2. \tag{10.1.4}$$

Substituting this into the conservation of energy we get

$$Mc^2 \geq m_1c^2 + m_2c^2 \implies M \geq m_1 + m_2. \tag{10.1.5}$$

This gives us a condition for the masses of the two daughter particles. We can state this as a threshold condition by defining

$$Q \coloneqq M - \sum_i m_i \tag{10.1.6}$$

and noting that a decay process is possible only if $Q \geq 0$. In the case of $Q = 0$ the daughter particles are at rest, for $Q > 0$ the excess energy becomes kinetic energy of the daughter particles.

We can use various invariant four-vector scalar products to derive the energies and momenta of the daughter particles. We can evaluate these in any inertial frame and choose to do so in the CM frame of the two daughter particles since this is easiest. We start with the invariant mass calculated in the rest frame of each particle, namely

$$P \cdot P = M^2c^2, \qquad \text{and} \qquad p_i \cdot p_i = m_i^2c^2. \tag{10.1.7}$$

We then compute these quantities in the CM frame and we have

$$\begin{aligned}
p_1^\mu &= P^\mu - p_2^\mu, && (10.1.8)\\
p_1^2 &= (P^\mu - p_2^\mu)^2 && (10.1.9)\\
&= P^2 + p_2^2 - 2P \cdot p_2 && (10.1.10)\\
m_1^2 c^2 &= M^2 c^2 + m_2^2 c^2 - 2ME_2, && (10.1.11)\\
\implies E_2 &= \frac{c^2}{2M}(M^2 - m_1^2 + m_2^2). && (10.1.12)
\end{aligned}$$

Similarly we have

$$\begin{aligned}
p_2^\mu &= P^\mu - p_1^\mu, && (10.1.13)\\
p_2^2 &= (P^\mu - p_1^\mu)^2 && (10.1.14)\\
&= P^2 + p_1^2 - 2P \cdot p_1 && (10.1.15)\\
m_2^2 c^2 &= M^2 c^2 + m_1^2 c^2 - 2ME_1, && (10.1.16)\\
\implies E_1 &= \frac{c^2}{2M}(M^2 + m_1^2 - m_2^2) && (10.1.17)
\end{aligned}$$

Rearranging the energy-momentum relation we have

$$p = \sqrt{\frac{E^2}{c^2} - m^2 c^2} \tag{10.1.18}$$

from which we get

$$p = |\boldsymbol{p}_1| = |\boldsymbol{p}_2| = \frac{c}{2M}\sqrt{[M^2 - (m_1 + m_2)^2][M^2 - (m_1^2 - m^2)^2]}. \tag{10.1.19}$$

Consider the special case where both daughter particles have identical mass, $m_1 = m_2 = m$, the particles then have energy

$$E_1 = E_2 = \frac{1}{2}Mc^2 \tag{10.1.20}$$

and momentum

$$p = \frac{c}{2}\sqrt{M^2 - 4m^2}. \tag{10.1.21}$$

Consider the even more special case where $m = 0$ and we have energy

$$E_1 = E_2 = \frac{1}{2}Mc^2 \tag{10.1.22}$$

and momentum

$$p = \frac{1}{2}Mc \tag{10.1.23}$$

from which we get

$$E_1 = E_2 = pc \tag{10.1.24}$$

as expected.

10.1.1 Units

In nuclear and particle physics SI units are usually far too big for the quantities we want to discuss. It is far more common to use electron volts for energy. Using $E = mc^2$ we can then measure mass in eV/c^2. Similarly using $E = pc$ we can measure momentum in eV/c^2. Energy scales are typically such that MeV or GeV are the appropriate units. For example the mass of a proton is $m_\mathrm{p} = 938\,\mathrm{MeV}/c^2$.

It should also be noted that it is common to work in natural units, in which c (and $\hbar$ for that matter) are set to 1. In this case energy, mass, and momentum are all measured with the same units. It is simple to convert back to non-natural units by dimensional analysis to find out where the factors of c need to be added.

10.2 Three-Body Decays

Suppose now that the parent particle of mass M decays into three daughter particles with masses m_i and momenta $p_i^\mu = (E_i/c, \boldsymbol{p_i})$. In the rest frame of the parent particle it has momentum $P^\mu = (Mc, \boldsymbol{0})$. Conservation of four-momentum gives

$$Mc^2 = E_1 + E_2 + E_3, \qquad \text{and} \qquad \boldsymbol{0} = \boldsymbol{p_1} + \boldsymbol{p_2} + \boldsymbol{p_3}. \tag{10.2.1}$$

Here we have four equations (one for each component of the momentum), but six unknowns (E_i and $|\boldsymbol{p_i}|$). We aren't just lacking information though, the particles produced will have a range of energies and momenta. We still have the condition that $M \ge m_1 + m_2 + m_3$ from energy conservation.

To explore the range of the energy spectrum we define the invariant mass of a pair of daughter particles, m_{ij} by

$$m_{ij}^2c^2 := (p_i^\mu + p_j^\mu)^2 = m_i^2c^2 + m_j^2c^2 + 2p_i \cdot p_j. \tag{10.2.2}$$

This is Lorentz invariant as it is defined as the square of a four-vector. In the CM frame of the parent particle we have

$$m_{ij}^2c^2 = (P^\mu - p_k^\mu)^2 = M^2c^2 + m_k^2c^2 - 2ME_k \tag{10.2.3}$$

where $k \ne i, j$. It then follows that particle k has energy

$$E_k = \frac{c^2}{2M}(M^2 + m_k^2 - m_{ij}^2). \tag{10.2.4}$$

Evaluating $p_i \cdot p_j$ in the CM frame of the i-j subsystem we have

$$p_i^{*\mu} = \left(\frac{E_i^*}{c}, \boldsymbol{p}^*\right), \qquad \text{and} \qquad p_j^{*\mu} = \left(\frac{E_j^*}{c}, -\boldsymbol{p}^*\right). \tag{10.2.5}$$

Where we are now using the asterisk to denote quantities in the i-j CM frame. We then have

$$p_i \cdot p_j = p_i^* \cdot p_j^* = \frac{E_i^*E_j^*}{c^2} + p^{*2}, \tag{10.2.6}$$

where $p^* = |\boldsymbol{p}^*|$. It follows that the minimum value of m_{ij}^2 occurs when both particles i and j are at rest in the i-j CM frame. That is, $p_i^{*\mu} = (m_ic, \boldsymbol{0})$ and $p_j^{*\mu} = (m_jc, \boldsymbol{0})$. We then have

$$\min m_{ij} = m_i + m_j \tag{10.2.7}$$

and so it follows that the maximum value E_k can take is

$$\max E_k = \frac{c^2}{2M}(M^2 + m_k^2 0(m_i + m_j)^2). \tag{10.2.8}$$

The minimum possible energy for particle k is simply when its at rest, $E_k = m_k c^2$.

■ **Example 10.2.9** The charged kaon, K^+, has a mass of $494\,\mathrm{MeV}/c^2$ and can decay into three π mesons:

$$\mathrm{K}^+ \longrightarrow 2\pi^+ + \pi^-. \tag{10.2.10}$$

Each pion has a mass of approximately $140\,\mathrm{MeV}/c^2$. The maximum energy of any given pion is then

$$\frac{M_{\mathrm{K}^+}^2 + m_\pi^2 - (2m_\pi)^2}{2M_{\mathrm{K}^+}^2} = 187.5\,\mathrm{MeV}/c^2. \tag{10.2.11}$$

10.2.1 Notes

In β-decay a neutron in a nucleus transforms into a proton and an electron is ejected. Measuring the energy of the electrons reveals that they have a spectrum of energies. This was early evidence that there is another particle involved in β-decay, and indeed there is, the electron antineutrino, $\bar{\nu}_\mathrm{e}$.

It can be tempting to think of decay processes as the decay products being contained within the parent particle and it simply breaking apart into the decay products. However, this idea doesn't explain how many particles have multiple decay modes or the exchange of kinetic energy into mass.

Eleven

Relativistic Scattering

11.1 Relativistic Two-Body Collisions

Consider two particles with masses m_i and four-momenta p_i^μ in the absence of external forces. Collisions can be described in any frame, but we typically work in the LAB and CM frames. In the LAB frame we take particle 2 to be at rest and refer to it as the target, and particle 1 as the beam particle or projectile. In the CM frame the total three-momentum is zero and so the two particles have equal and opposite three-momenta. The two frames are, of course, related by a Lorentz transform.

A useful quantity is the invariant mass of the two particles, defined as for decay processes, by

$$m_{12}^2c^2 = (p_1^\mu + p_2^\mu)^2 = p_1^2 + p_2^2 + 2p_1 \cdot p_2. \tag{11.1.1}$$

Since this is invariant we can evaluate it in any frame. In the CM frame $\boldsymbol{p}_1^* = -\boldsymbol{p}_2^*$ and so

$$p_1^{\mu *} + p_2^{\mu *} = \left(\frac{E_1^*}{c} + \frac{E_2^*}{c}, \boldsymbol{p}_1^* + \boldsymbol{p}_2^*\right) = \left(\frac{E_1^* + E_2^*}{c}, \boldsymbol{0}\right). \tag{11.1.2}$$

Hence the invariant mass is

$$m_{12}^2c^2 = \frac{1}{c^2}(E_1^* + E_2^*). \tag{11.1.3}$$

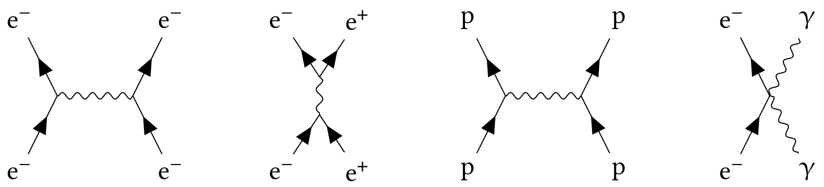
So, in the CM frame, $m_{12}c^2$ is just the total energy. Since $E_i^* \geq m_ic^2$ we find that the minimum value for the invariant mass is

$$\min m_{12} = m_1 + m_2. \tag{11.1.4}$$

In the LAB frame the beam particle has momentum $p_1^\mu = (E_1/c, \boldsymbol{p}_1)$, and the target particle has momentum $p_2^\mu = (m_2c, \boldsymbol{0})$. The invariant mass is then

$$m_{12}^2c^2 = m_1^2c^4 + m_2^2c^4 + 2E_1m_2c^2. \tag{11.1.5}$$

Notice that this is *not* the total energy in the LAB frame. However, since m_{12} is invariant we have shown that the relativistic energy is *not* Lorentz invariant. That is the energy differs depending on the frame in which it is measured. Notice that it is still possible that energy be a conserved quantity, meaning it doesn't change with time, within a given frame, its just that when we move frames the value changes.



(a) $e^- + e^- \to e^- + e^-$ (b) $e^- + e^+ \to e^- + e^+$ (c) $p + p \to p + p$ (d) $e^- + \gamma \to e^- + \gamma$

Figure 11.1: Various elastic scattering processes.

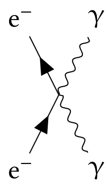


Figure 11.2: Compton scattering.

11.2 Elastic Scattering

An **elastic collision** is defined as one where, after the particles are sufficiently separated, the total kinetic energy is conserved. A consequence of this is that the total mass of the particles is conserved since

$$T = E - \sum_i m_i c^2 \tag{11.2.1}$$

and so if both T and E are conserved $\sum_i m_i$ must be too.

Elastic collisions most often occur between fundamental particles at energies below a few hundred MeV. We can think of these collisions as the particles bouncing off each other, although there may well be internal stages including other particles that we don't see. Common examples are electron-electron scattering, electron-positron annihilation and production, proton proton scattering, and **Compton scattering**, which is a photon scattering off of an electron. Possible interactions for these cases are shown in Figure 11.1. We treat all of these processes in pretty much the same way. We will consider as an example here Compton scattering.

11.2.1 Compton Scattering

Consider a photon being scattered from an electron. In the LAB frame the incident photon has frequency ν, and the scattered photon frequency ν'. Suppose that the photon is scattered at an angle ϑ to its incoming direction, and the electron recoils at an angle ψ. See Figure 11.2

Denote the momentum of the photon and electron by p_1^μ and p_2^μ before the scattering and q_1^μ and q_2^μ after the scattering. In the LAB frame the photon and electron have four-momenta

$$p_1^\mu = \left(\frac{E_1}{c}, p_1, 0, 0\right) = \left(\frac{h\nu}{c}, \frac{h\nu}{c}, 0, 0\right) \tag{11.2.2}$$

$$p_2^\mu = (m_e c, 0, 0, 0) \tag{11.2.3}$$

$$q_1^\mu = \left(\frac{E_1'}{c}, q_1\cos\vartheta, q_1\sin\vartheta, 0\right) = \left(\frac{h\nu'}{c}, \frac{h\nu'}{c}\cos\vartheta, \frac{h\nu'}{c}\sin\vartheta, 0\right) \tag{11.2.4}$$

$$q_1 = \left(\frac{E_{e^-}}{c}, p_{e^-}\cos\psi, -p_{e^-}\sin\psi, 0\right) \tag{11.2.5}$$

Here we have used $E = pc$ and $E = h\nu$ for a photon. Momentum conservation gives us

$$p_1^\mu + p_2^\mu = q_1^\mu + q_2^\mu \implies q_2^\mu = p_1^\mu + p_2^\mu - q_1^\mu. \tag{11.2.6}$$

We therefore have

$$m_{e^-}^2 c^2 = p_1\cdot p_1 + p_2\cdot p_2 + q_1\cdot q_1 + 2p_1\cdot p_2 - 2p_1\cdot q_1 - 2p_2\cdot q_1 \tag{11.2.7}$$

$$= 0 + m_{e^-}^2c^2 + 0 + 2m_e h\nu - 2\left(\frac{h^2\nu\nu'}{c^2} - \frac{h^2\nu\nu'}{c^2}\cos\vartheta\right) - 2m_e h\nu'. \tag{11.2.8}$$

Rearranging this we get

$$\frac{1}{\nu'} - \frac{1}{\nu} = \frac{h}{m_e c^2}(1-\cos\vartheta). \tag{11.2.9}$$

Rewriting this in terms of the wavelengths, $\lambda = c/\nu$ and $\lambda' = c/\nu'$ we find that the change in wavelength of a scattered photon is

$$\Delta\lambda = \lambda' - \lambda = \frac{h}{mc}(1-\cos\vartheta). \tag{11.2.10}$$

The quantity $h/(mc) = 2.4\times 10^{-12}\,\mathrm{m}$ is called the **Compton wavelength** of the electron.

11.3 Antiparticles

We take a moment now to "discover" a remarkable fact that arises as a consequence of mixing special relativity and quantum mechanics. Consider a particle, in the state $|\varphi\rangle$. This particle could be scattered into a new states, $|\psi\rangle$, by some potential, $\hat{V}$. The amplitude for this to happen is $\langle\psi|\hat{V}|\varphi\rangle$. To compute the final state of the electron we must consider all possible scatterings, we can do so by considering perturbation theory to all orders or by considering a path integral[1].

[1] see the quantum theory notes for path integrals

Importantly we must also consider scattering to events *outside* of the light cone of the particle. This implies we are considering faster than light travel, and we are, which is fine as this isn't a real state that we can measure the particle in, instead it is a virtual state and no information can be transmitted. When this is the case we have two events, the start and end of the scattering process, which are separated by a space-like interval. Therefore we can find a frame in which the order of the events is reversed, that is the scattering ends *before* it starts. In this case we can view the particle as going *backwards* in time. The particle travelling backwards in

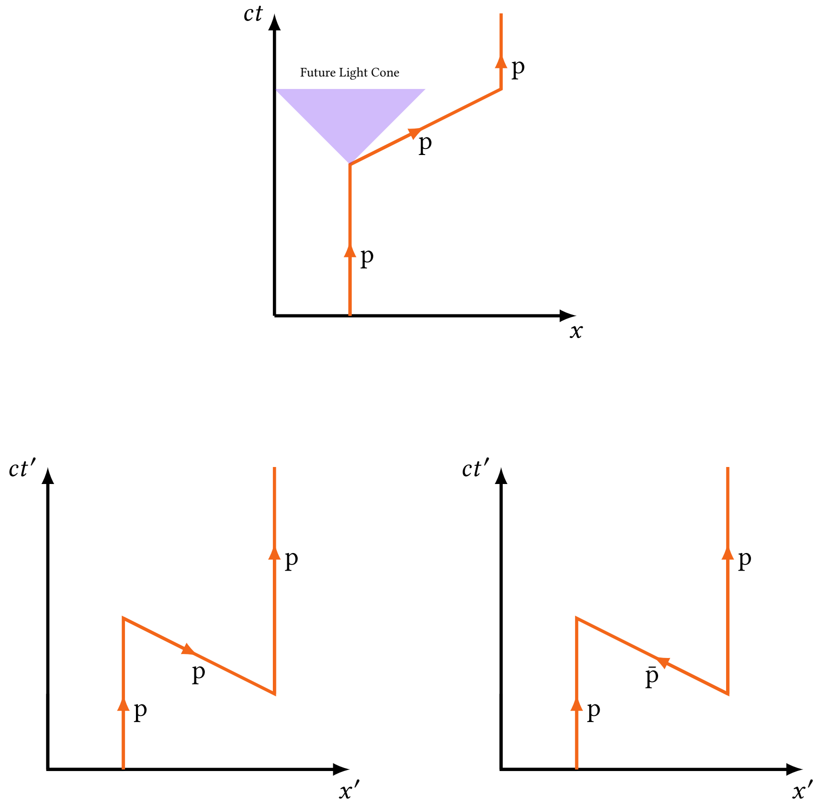


Figure 11.1: A particle is scattered backwards in time, or is it an antiparticle?

time is identical to the forward travelling particle. Mathematically this particle travelling backwards in time is identical to a particle of the opposite charge (and parity) travelling forwards in time. This is the **antiparticle**!

Figure 11.1 shows three views of the same scattering. The first is in a frame in which the particle is scattered forwards in time. The other two are both in the same frame, different to the first in that the order of scattering events is reversed. The second shows a particle scattered backwards in time. The third shows a particle coming in the bottom, some time later particle-antiparticle pair production occurs. The antiparticle goes on to annihilate with the original particle and the particle we see at the end is the one produced in the particle-antiparticle pair.

11.4 Inelastic Scattering

Kinetic energy need not be, and indeed often isn't, conserved in scattering interactions. If the invariant mass of the two incoming particles is greater than the sums of their masses then there is excess energy available, which we see as the kinetic energy of the incoming particles. This energy can be used to create new particles in particle-antiparticle pair production, the particles and antiparticles produced this way are real (as opposed to the virtual antiparticle of the last section).

Examples of inelastic scattering are electron-positron scattering with some excess energy converted into a photon:

$$e^- + e^+ \to e^- + e^+ + \gamma. \tag{11.4.1}$$

Electron-positron pair production in a nucleus, N:

$$\gamma + \mathrm{N} \to e^- + e^+ + \mathrm{N}. \tag{11.4.2}$$

Electron-positron annihilation creating a photon which later creates a muon-antimuon pair:

$$e^- + e^+ \to \mu^- + \mu^+. \tag{11.4.3}$$

A proton-proton interaction where excess energy is converted into a proton-antiproton pair:

$$\mathrm{p} + \mathrm{p} \to \mathrm{p} + \mathrm{p} + \mathrm{p} + \bar{\mathrm{p}}. \tag{11.4.4}$$

Gluon-gluon fusion into a Higgs boson and other particles, "X":

$$\mathrm{g} + \mathrm{g} \to \mathrm{H}^0 + X, \tag{11.4.5}$$

this is the most common Higgs-producing interaction in the LHC, which uses beams of protons, which produce gluons when interacting via the strong force.

Consider proton-proton scattering producing a π^0 meson:

$$\mathrm{p} + \mathrm{p}top + \mathrm{p} + \pi^0. \tag{11.4.6}$$

How much energy does the projectile need in the LAB frame in order for this process to be able to occur. We know from Equation (11.1.5) that the invariant mass gives us

$$(m_{12}c^2)^2 = m_1^2c^4 + m_2^2c^4 + 2E_1m_2c^2 \tag{11.4.7}$$

and this is minimised when all three final particles are at rest, giving

$$m_{12\,\mathrm{min}} = \min m_{12} = m_1 + m_2 + m_3. \tag{11.4.8}$$

So, in the LAB frame the minimum energy required is

$$\min E_1 = \frac{(m_{12\,\mathrm{min}})^2c^4 - (m_1^2 + m_2^2)c^4}{2m_2c^2} = \frac{(m_1 + m_2 + m_3)^2c^4 - (m_1^2 + m_2^2)c^4}{2m_2c^2}. \tag{11.4.9}$$

For our example $m_1 = m_2 = m_\mathrm{p} = 938\,\mathrm{MeV}/c^2$, and $m_3 = m_\pi = 135\,\mathrm{MeV}/c^2$, so we have

$$\min E_1 = 1218\,\mathrm{MeV}. \tag{11.4.10}$$

This is the **threshold energy**, the minimum energy required for a process to be energetically possible.

Recall that we defined the kinetic energy of a particle with mass m and total energy E as

$$T := E - mc^2 = E - E_\mathrm{rest}. \tag{11.4.11}$$

[2] this looks very similar to the formula for $\min E_1$, notice that the squared is outside of the last bracket here but each term is squared for E_1.

We can write the threshold kinetic energy as[2]

$$\min T_1 = \min E_1 - m_1c^2 = \frac{(m_1 + m_2 + m_3)^2c^4 - (m_1 + m_2)^2c^4}{2m_2c^2}. \tag{11.4.12}$$

For the proton-proton process creating a π^0 meson we have $\min T_1 = 280\,\mathrm{MeV}$.

We can generalise this formula easily, let $m_{j,\mathrm{f}}$ be the final mass of the jth particle and $m_{k,\mathrm{i}}$ the initial mass of the kth particle. Then

$$\min T_1 = \frac{(\sum_j m_{j,\mathrm{f}})^2c^2 - (\sum_k m_{k,\mathrm{i}})^2c^2}{2m_{\mathrm{target}}} \tag{11.4.13}$$

where m_{target} is the mass of the target particle. Note that the ranges of both sums may differ if particles are created or annihilated.

Consider the interaction

$$\mathrm{p} + \mathrm{p} \longrightarrow \mathrm{p} + \mathrm{p} + \mathrm{p} + \bar{\mathrm{p}}. \tag{11.4.14}$$

In this case $m_{j,\mathrm{f}} = m_{k,\mathrm{i}} = m_{\mathrm{target}} = m_{\mathrm{p}} = 938\,\mathrm{MeV}/c^2$, and we have 2 initial particles and four final particles so the minimum required kinetic energy is

$$\min T = \frac{(4m_{\mathrm{p}})^2c^2 - (2m_{\mathrm{p}})^2c^2}{2m_{\mathrm{p}}} = 6m_{\mathrm{p}}c^2 = 5.63\,\mathrm{GeV}. \tag{11.4.15}$$

It should be noted at this point that energetic feasibility is not the only condition for a given interaction to be allowed. For example the following reaction cannot occur as baryon number isn't conserved:

$$\mathrm{p} + \mathrm{p} \not\rightarrow \mathrm{p} + \mathrm{p} + \bar{\mathrm{p}}. \tag{11.4.16}$$

In general there are many quantities that must be conserved, such as baryon number, electron/muon/tauon lepton number, charge, …. We will see more in the particle physics part of the course.

Part III

General Relativity

Twelve

The Basics of General Relativity

12.1 Mach's Principle

Ernst Mach is credited with formulating what is now known as **Mach's principle**:

> Absolute motion has no meaning, there is only motion relative to other objects.

This seemingly simple idea has deep ramifications. In the simplest example a single particle, alone in the universe, cannot have any motion, this means no kinetic energy, momentum, or angular momentum. Likewise if the Earth where the only body in the universe it would not make sense to discuss its rotation about its axis. From this we can deduce that the Coriolis force and centrifugal force are ultimately not due to Earth's rotation but due to the net gravitational force of the rest of the universe.

Put another way Mach's principle states that inhomogeneity is required for translation and anisotropy for rotation. We used to take distance stars as a fixed frame in which to measure motion. Nowadays we can take the cosmic microwave background (CMB) as our fixed frame, using slight inhomogeneities due to quantum fluctuations expounded by inflation in the early universe.

12.2 Inertial and Gravitational Mass

There are two types of mass that we commonly consider. The first, which we shall refer to as **inertial mass**, is the mass as it appears in Newton's second law, and is a measure of how difficult it is to accelerate an object:

$$\boldsymbol{F} = m_{\mathrm{I}}\boldsymbol{a}. \tag{12.2.1}$$

The second type of mass is **gravitational mass**. It is a measure of how strongly a given gravitational field pulls on the object, its a coupling constant, the same as charge is to electromagnetism, colour to the strong force, or isospin to the weak force.

$$\boldsymbol{F} = -G\frac{M_{\mathrm{G}}m_{\mathrm{G}}}{r^2}\hat{\boldsymbol{r}}. \tag{12.2.2}$$

There is no reason why resistance to motion and strength of gravitational interactions should be related but we see from Equations (12.2.1) and (12.2.2) that they

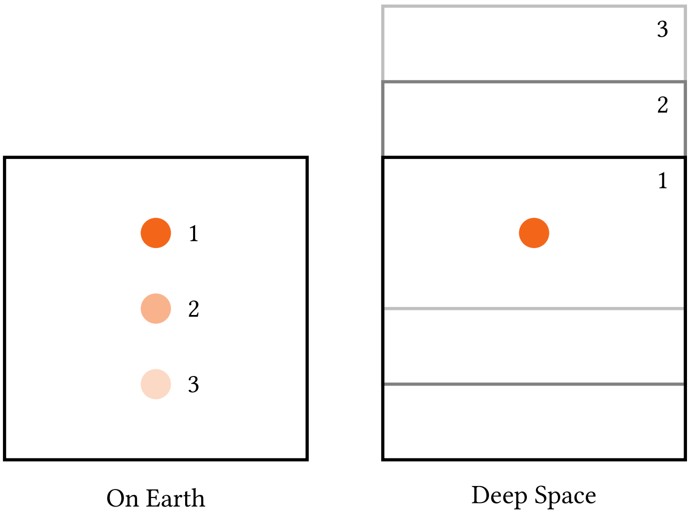


Figure 12.1: Left: A ball falls in a box on Earth, three snapshots. Right: A box accelerates upwards in deep space, three snapshots. Inside the box the two cases are indistinguishable.

are indeed connected. These equations only require that $m_{\mathrm{I}} \propto m_{\mathrm{G}}$, but we define the constant G such that $m_{\mathrm{I}} = m_{\mathrm{G}}$. We then get that $\boldsymbol{a} = GM/r^2$. This holds for all objects in this gravitational field, which leads to the **weak equivalence principle**:

> All objects experience the same acceleration in a given gravitational field.

This result is not immediately obvious, especially given our everyday experiences, where a feather will fall slower than a stone, of course the difference here is we aren't accounting for air resistance. The weak equivalence principle was first tested famously by Galileo who dropped balls of different masses but the same size from the leaning tower of Pisa. For this reason the weak equivalence principle is sometimes referred to as Galileo's principle. It has no been verified to a precision of 1 part in 1×10^{12}.

12.3 The Strong Equivalence Principle

Einstein proposed a series of *gedanken* (thought) experiments, involving a person in a box with no way to see out in order to measure relative motion. First, suppose that the box is on Earth. Every object will therefore accelerate downwards with acceleration g. Now, suppose that the box is in deep space, so that the net gravitational force is negligible. Accelerate the box upwards with acceleration g. Objects in the box will seem to "fall" to the floor, when from a different frame we might see the floor being accelerated up to the object. However, from inside the box there is no way to tell.

A similar thought experiment considers the box in free fall on Earth. In this case objects in the box *and* the box fall with an acceleration g, and so have no acceleration relative to each other, meaning objects will seem to float in the box, when in another frame we see them falling along with the box. Now returning to deep space if the box is not accelerating objects will float, meaning from inside the box one cannot tell the difference between free fall and no gravity.

From these two thought experiments Einstein formulated the **strong equivalence principle**:

> No local experiment can distinguish between the free fall of a body in a gravitational field and the uniform motion of the same body in the absence of a gravitational field.

This principle is fundamental to general relativity. The requirement that the experiments be "local" is to explicitly disallow experiments that measure a change in gravitational field, rather than absolute field, for example if the box was sufficiently tall one could drop an object at the top. If the box is on Earth then the acceleration at the start of the journey would be measurably lower than the acceleration at the end. If the box was accelerating with uniform acceleration in space however then the acceleration would be the same for the whole journey. To aid in this definition we define a **local inertial frame** (LIF).

Alternatively if the box is sufficiently wide then dropping objects at either end one would find they accelerate in slightly different directions on Earth, as both accelerate towards the centre of the Earth. This doesn't happen in space.

In the absence of a gravitational field of course we can drop the locality requirement since if there is no field there can certainly be no variations in the field.

A more formal statement of the strong equivalence principle involves treating each point as a LIF, and we can then state the strong equivalence principle as

> A linearly accelerating frame is locally identical to a rest frame in a gravitational field.

That is an instantaneously co-moving frame (ICMF) has the same properties as a local inertial frame (LIF) in a gravitational field.

12.4 Gravitational Redshift

Consider a box accelerating with a magnitude g in deep space. A photon is emitted going downwards from the top of the box, in the opposite direction to the acceleration. The photon takes time t_H to reach the bottom of the box, which is a distance H from the top in the rest frame of the box. In time t_H the photon covers a distance ct_H, and the box moves a distance $gt_H^2/2$. Therefore

$$H = ct_H + \frac{1}{2}gct_H^2 \implies t_H = \frac{-c \pm \sqrt{c^2 + 2gH}}{g} = \frac{c}{g}\left(\pm\sqrt{1 + \frac{2gH}{c}} - 1\right). \tag{12.4.1}$$

Taking the positive time speed that the box acquires in this time is

$$v_H = gt_H = c\left(\sqrt{1 - \frac{2gH}{c^2}} - 1\right) \approx \frac{gH}{c} \tag{12.4.2}$$

where we have used the Taylor expansion

$$\left(1 + \frac{2gH}{c^2}\right)^{1/2} \approx 1 + \frac{gH}{c^2}, \tag{12.4.3}$$

which assumes that $gH \ll c^2$.

An observer on the floor of the box is moving toward the source and therefore observers a Doppler shift, with the source frequency, ν_{source}, appearing to the observer to be

$$\nu_{\text{obs}} = \nu_{\text{source}}\sqrt{\frac{1-\beta}{1+\beta}} \approx \nu_{\text{source}}(1-\beta) = \nu_{\text{source}}\left(1 - \frac{v}{c}\right) \tag{12.4.4}$$

where we have again applied a Taylor expansion assuming $\beta \ll 1$:

$$(1-\beta)^{1/2}(1-\beta)^{-1/2} \approx \left(1 - \frac{\beta}{2}\right)\left(1 - \frac{\beta}{2}\right) = 1 - \beta + \frac{\beta^2}{2} \approx 1 - \beta. \tag{12.4.5}$$

In this case this is a blue shift since we are considering an observer moving toward the source, beware that the Doppler shift formula is for an observer moving *away* from the source, so the sign of v, and hence β, needs to be reversed, that is we take $v = -v_H$. The relative increase in the frequency is then

$$\frac{\Delta\nu}{\nu} = \frac{\nu_{\text{obs}} - \nu_{\text{source}}}{\nu_{\text{source}}} \tag{12.4.6}$$

$$= \frac{(1 + v_H/c)\nu_{\text{source}} - \nu_{\text{source}}}{\nu_{\text{source}}} \tag{12.4.7}$$

$$= \frac{v_H}{c} \tag{12.4.8}$$

$$= \frac{gH}{c^2}. \tag{12.4.9}$$

By the strong equivalence principle this behaviour of an accelerating box must be the same as a box in a gravitational field, so we conclude that gravity causes a Doppler shift without the need for relative motion of the observer and source. This is called **gravitational redshift**.

12.5 Curved Light

Consider a photon crossing a box of width d which is accelerating upwards at g. When the photon reaches the other side, which takes time $t = d/c$, the box has moved upwards a distance $gt^2/2 = gd^2/(2c^2)$. Alternatively, to an observer in the box it appears as if the photon drops a distance $gd^2/(2c^2)$.

We can calculate the approximate radius of curvature, R, of the lights path by assuming it follows a circular path. For small deflections we have $2R\varphi \approx d$, where φ is the angle between the tangents to the path at either end and a straight line joining the endpoints (see Figure 12.1). We therefore have

$$\frac{d}{2R} \approx \varphi \approx \tan\varphi = \frac{gt^2}{2d} = \frac{gd}{2c^2}, \tag{12.5.1}$$

and so

$$R \approx \frac{c^2}{g}. \tag{12.5.2}$$

By the strong equivalence principle a gravitational field of strength g will produce the same curvature of light. For Earth the radius of curvature of light is $R \approx 9{\times}10^{15}$ m,

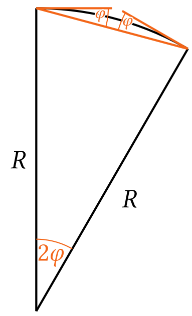


Figure 12.1: The geometry for calculating the radius of curvature of light in an accelerating box.

which is much greater than the radius of Earth, which is just $R_\oplus = 6 \times 10^6$ m, and so the effect of Earth's gravity on lights path isn't noticeable, over scales that fit on Earth. During a solar eclipse a noticeable bend of $\varphi \approx 0.001°$ was measured for light passing by the Sun, which we predict causes light to bend with a radius of curvature of $R \approx 3 \times 10^{14}$ m.

12.6 Curved Space

In the previous section we said that light curves when passing a massive object. This is true only when we assume that space is Euclidean. A better description is that gravity bends space-time into a non-Euclidean geometry. In Euclidean geometry everything works as expected, parallel lines never meet, angles in a triangle add to 180°, the circumference of a circle of radius r is $2\pi r$ etc. In non-Euclidean geometry these properties may not hold, one example being the surface of a sphere, which is a non-Euclidean two-dimensional surface (embedded in three-dimensional Euclidean space).

The geometry of Newtonian mechanics is a three-dimensional Euclidean space. The geometry of special relativity is a four-dimensional Minkowski space with zero curvature. The geometry of general relativity is significantly more complicated, to describe it we need significantly more maths. For now we will stick to the idea that it is curved.

The curvature of a space is captured by the invariant space-time interval, which, in its most general form, is given by

$$\mathrm{d}s = g_{\mu\nu}\,\mathrm{d}x^\mu\,\mathrm{d}x^\nu \tag{12.6.1}$$

where g is the metric tensor. In Newtonian mechanics indices run from 1 to 3 and $g_{ij} = \delta_{ij}$, where δ is the Kronecker delta. In special relativity indices run from 0 to 4 and $g_{\mu\nu} = \eta_{\mu\nu}$ is the Minkowski metric, given by

$$\eta_{\mu\nu} = \mathrm{diag}(1, -1, -1, -1). \tag{12.6.2}$$

In general relativity indices again run from 0 to 4, but now the metric is far less restricted. In general it will be a function of the position in space-time, and also need not be diagonal.

Light follows geodesics in space-time. A **geodesic** being the shortest path, as measured by integrating the invariant space-time interval along the path. In particular for light $\mathrm{d}s^2 = 0$, and we call the path a **null geodesic**.

12.7 Clocks in a Gravitational Field

Recall the setup for introducing the gravitational redshift. We can measure the frequency of the photon and from there work out the time, however, the frequency is effected by gravity, so the time must also be effected by gravity. In particular the photon blue-shifts, and therefore the frequency increases, which means that if we use each oscillation to define one unit of time things will take more units of time to complete. This is **gravitational time dilation**.

In particular the time, T_H, measured on a clock at a height H above the surface of Earth runs faster than a clock, measuring T_A, at the surface of the Earth, by

$$T_A = \left(1 - \frac{gH}{c^2}\right) T_H. \tag{12.7.1}$$

We can derive this result another way by considering two clocks, A and B at the bottom and top of a lift, which is accelerating upwards with acceleration g relative to some clock C which is at rest. When B passes C it has speed v_B and by special relativity B runs slower than C. When A passes C it has speed $v_A > v_B$, since the whole lift is accelerating, and A runs slower again. The times are related by the special relativity time dilation equation, $t = \tau/\gamma$:

$$T_B = T_C\sqrt{1 - \frac{v_B^2}{c^2}}, \qquad \text{and} \qquad T_A = T_C\sqrt{1 - \frac{v_A^2}{c^2}}. \tag{12.7.2}$$

Consider now a lift on Earth, stationary with clocks A and B at the bottom and top, and the clock C in free fall. By the strong equivalence principle we must have the same relation between times measured. Making a binomial expansion of the square roots above and combining the equations we get

$$T_B \approx T_A\left(1 + \frac{v_A^2 - v_B^2}{2c^2}\right). \tag{12.7.3}$$

We can calculate the speed of clock C using energy conservation. If its mass is m then we have

$$\frac{1}{2}m(v_a^2 - v_B^2) + m(U(x_A) - U(x_B)) = \text{constant} \tag{12.7.4}$$

where $U(x) = -GM/r$ is the Newtonian gravitational potential. We can choose for the constant to 0 and so we find that

$$T_B = T_A\left(1 - \frac{\Delta U}{c^2}\right). \tag{12.7.5}$$

Near the surface of Earth the acceleration, g, is a constant and $\Delta U = -gH$, so we recover the same result.

The first test of gravitational time dilation was in 1971 when atomic clocks were carried on planes in opposite directions around the world and it was found that, after accounting for special relativity, there was still a noticeable difference compared to clocks that had remained stationary.

The global positioning system (GPS) has to account for the effects of both special and general relativity on the clocks on its satellites, which have to be very accurate for GPS to work. GPS satellites orbit Earth twice per day at a height of $R_S = 26\,600\,\mathrm{km}$ from the centre of Earth. The difference in gravitational potential between the Earth's surface and the satellites is[1]

[1] $R_\oplus = 6370\,\mathrm{km}$
$M = M_\oplus = 5.972 \times 10^{24}\,\mathrm{kg}$
$G = 6.6738 \times 10^{-11}\,\mathrm{m^3\,kg^{-1}\,s^{-2}}$

$$\frac{\Delta U}{c^2} = \frac{GM}{c^2}\left(\frac{1}{R_\oplus} - \frac{1}{R_S}\right) = 5.3 \times 10^{-10}. \tag{12.7.6}$$

This is the fraction by which the satellite clocks run faster. If we didn't correct for the effect over a day the GPS clocks would be off by 45.8 µs, which corresponds to a error of 13.7 km in the position they give.

We also have to account for time dilation due to the speed of the satellites, which is $v_S = 3874\,\mathrm{m\,s^{-1}}$, giving

$$\frac{1}{\gamma} = \sqrt{1 - \frac{v_S^2}{c^2}} \approx 1 - \frac{v_S^2}{2c^2} \approx 1 - 8.35 \times 10^{-11}. \tag{12.7.7}$$

The result is that over a day GPS clocks would be off by −7.2 µs, which is an error of −2.2 km in the position.

The result is that we need a net correction of 38.6 µs per day. This correction is made by using a frequency of 10.23 MHz, but actually setting each clock to a frequency of 10.229 999 995 4 MHz on Earth, which counters the effects of relativity.

12.8 Lengths in a Gravitational Field

It should not be too surprising at this point that an effect for time has a related effect for length. In this case it can be show that two lengths, L_A and L_B measured in a gravitational field are related by

$$L_A = L_B\left(1 - \frac{\Delta U}{c^2}\right). \tag{12.8.1}$$

The result is that the stronger the gravitational field is the shorter the length is. This can be seen as a consequence of space-time being curved.

Thirteen

More General Relativity

13.1 Einstein's Field Equations

Einstein proposed that the curvature of space-time is related to the distribution of matter (and therefore energy). In particular the **gravitational field**, $G^{\mu\nu}$ is a 4-dimensional, rank 4 tensor given by

$$G^{\mu\nu} := R^{\mu\nu} - \frac{1}{2}g^{\mu\nu}R = \frac{8\pi G}{c^2}T^{\mu\nu}. \tag{13.1.1}$$

These are the **Einstein's field equations**. Here $R^{\mu\nu}$ is the **Ricci curvature tensor**, which is defined in terms of derivatives of the metric tensor, $g^{\mu\nu}$. R is the local curvature, which is given by $R = g^{\mu\nu}R_{\mu\nu}$. G is Newton's gravitational constant. $T^{\mu\nu}$ is the **energy-momentum tensor**, sometimes the word stress is added in as well, such as the stress-energy tensor. The element T^{00} corresponds to the relativistic energy density and T^{0i} correspond to the energy flow in direction x^i. T^{ii} corresponds to the ith component of the relativistic 3-momentum, and T^{ij} to the relativistic momentum flow of the ith component of the three-momentum in the x^j direction.

The left hand side of the equation, $R^{\mu\nu} - g^{\mu\nu}R/2$, corresponds to the curvature of space-time, and the right hand side, $8\pi GT^{\mu\nu}/c^2$, to the distribution of matter. Einstein's field equations are the equivalent of Poisson's equations in electromagnetism, they relate the potentials to the distribution of charges and current, or in this case the gravitational potential to the distribution of matter.

The general goal is to solve the Einstein field equations for the metric, $g^{\mu\nu}$. This is a difficult task and we won't do so here, we'll just consider a few examples.

13.2 The Metric Tensor

We are usually interested in spherically symmetric situations, such as point particles. For this reason it is advantageous to be able to work in spherical coordinates. We can do so by considering

$$x = r\cos\varphi\sin\vartheta, \qquad y = r\sin\varphi\sin\vartheta, \qquad \text{and} \qquad z = r\cos\vartheta. \tag{13.2.1}$$

Differentiating these we can find $\mathrm{d}x$, $\mathrm{d}y$, and $\mathrm{d}z$ in terms of spherical quantities. For example

$$\frac{\partial x}{\partial r} = \cos\varphi\sin\vartheta \implies \mathrm{d}x = \cos\varphi\sin\vartheta\,\mathrm{d}r. \tag{13.2.2}$$

Doing so for all possible pairs of Cartesian and spherical coordinates we find that we can write the invariant space-time interval as

$$\mathrm{d}s^2 = c^2\,\mathrm{d}t^2 - \mathrm{d}r^2 - r^2\,\mathrm{d}\vartheta^2 - r^2\sin^2\vartheta\,\mathrm{d}\varphi^2 = g_{\mu\nu}\,\mathrm{d}x^\mu\,\mathrm{d}x^\nu, \tag{13.2.3}$$

where now $\mathrm{d}x^\mu = (c\,\mathrm{d}t, \mathrm{d}r, \mathrm{d}\varphi, \mathrm{d}\vartheta)$. We can then identify

$$g_{\mu\nu} = \mathrm{diag}(1, -1, -r^2, -r^2\sin^2\vartheta). \tag{13.2.4}$$

This is still the Minkowski metric, just in spherical coordinates.

We expect that the presence of matter will result in space-time becoming curved. If the matter is spherically symmetric then we make the ansatz that

$$\mathrm{d}s^2 = e(r)c^2\,\mathrm{d}t^2 - f(r)\mathrm{d}r^2 - r^2\,\mathrm{d}\vartheta^2 - r^2\sin^2\vartheta\,\mathrm{d}\varphi^2 \tag{13.2.5}$$

where e and f are some functions of r to be determined from the exact distribution of matter.

13.3 Schwarzschild Solution

Outside of a spherical mass at the origin the solution to Einstein's field equations, known as the **Schwarzschild solution**, gives

$$e(r) = 1 + \frac{2U}{c^2} = 1 - \frac{2GM}{c^2 r}, \qquad \text{and} \qquad f(r) = \frac{1}{1 + 2U/c^2} = \frac{1}{1 - 2GM/(c^2 r)}. \tag{13.3.1}$$

The resulting invariant space-time interval is called the **Schwarzschild line element**:

$$\mathrm{d}s^2 = \left(1 - \frac{2GM}{c^2 r}\right)c^2\,\mathrm{d}t^2 - \frac{1}{1 - 2GM/(c^2 r)}\,\mathrm{d}r^2 - r^2\,\mathrm{d}\vartheta^2 - r^2\sin^2\vartheta\,\mathrm{d}\varphi^2. \tag{13.3.2}$$

Some properties of this solution are

- **Static** The functions e and f don't depend on time, so the solution doesn't depend on time.
- **Singularity** There is a singularity at $r = 0$ where $U \to -\infty$ and so $e \to -\infty$ and $f \to 0$.
- **Far-Field Limit** In the limit $r \to \infty$ $e, f \to 1$, so we recover flat, Minkowski, space-time.
- **Curvature** The factor $-2U/c^2 = 2GM/(c^2 r)$ can be seen as a measure of the curvature a distance r from the origin.
- **Schwarzschild Radius** The function f diverges as r approaches

  $$\mathcal{R} := \frac{2GM}{c^2}. \tag{13.3.3}$$

  The curvature therefore diverges also. We call this quantity the **Schwarzschild radius**. For the Sun $\mathcal{R}_\odot = 3\,\mathrm{km}$ and for Earth $\mathcal{R}_\oplus = 9\,\mathrm{mm}$, both much *much* smaller than the actual size. For a neutron star of radius R_N we find that $\mathcal{R}_\mathrm{N}/R_\mathrm{N} \approx 0.3$.

The proper time interval for the Schwarzschild solution can be obtained by considering a zero spatial interval with $\mathrm{d}r = \mathrm{d}\vartheta = \mathrm{d}\varphi = 0$. This gives

$$\mathrm{d}s^2 = c^2\mathrm{d}\tau^2 = \left(1 - \frac{2GM}{c^2 r}\right) c^2\,\mathrm{d}t^2 \implies \mathrm{d}\tau = \sqrt{1 - \frac{2GM}{c^2 r}}\mathrm{d}t \approx \left(1 + \frac{U}{c^2}\right)\mathrm{d}t. \quad (13.3.4)$$

Similarly we can derive the proper radial length by considering a distance $\mathrm{d}l$ defined by setting $\mathrm{d}t = \mathrm{d}\varphi = \mathrm{d}\vartheta = 0$ and setting $\mathrm{d}l = \sqrt{-\mathrm{d}s^2}$.

$$\mathrm{d}s^2 = -\mathrm{d}l^2 = -\frac{\mathrm{d}r^2}{1 - 2GM/(c^2 r)} \implies \mathrm{d}l = \frac{1}{\sqrt{1 - 2GM/(c^2 r)}}\mathrm{d}r \approx \left(1 - \frac{U}{c^2}\right)\mathrm{d}r. \quad (13.3.5)$$

We see that these are equivalent to the gravitational time dilation and length contraction equations derived in the last chapter.

Considering a photon in this gravitational field we know it will follow a null geodesic with $\mathrm{d}s^2 = 0$. It is possible to show from the Schwarzschild solution that the gravitational redshift is, as expected,

$$\frac{\Delta\nu}{\nu} = -\frac{GM}{c^2}\left(\frac{1}{r_1} - \frac{1}{r_2}\right) \quad (13.3.6)$$

where r_1 is the position of the source and r_2 the position of the observer. Close to Earth we recover the previous result $\Delta\nu/\nu \approx gH/c^2$.

13.4 Black Holes

When $r \to \mathcal{R}$ we see that $\mathrm{d}\tau \to 0$ and $\mathrm{d}l \to \infty$. This means that no information can be transmitted from inside the sphere $r = \mathcal{R}$ to outside of this sphere. We call this sphere the **event horizon** and the region inside a **black hole**, as even light can't escape.

13.5 Cosmology

General relativity has many consequences for cosmology. One of the first demonstrated was that an isotropic universe cannot be static, it must either expand or contract. In 1920 astronomers measured the expansion of the universe, by measuring the rate at which distant galaxies recede, which was given by the red shift of spectral lines. They found out that not only is the universe expanding but the furthest points are expanding fastest. It is believed that the expansion started after the big bang and continued up to now. Probably with a period early on of rapid expansion, known as inflation.

We assume that the universe is, on the largest scales, homogenous and isotropic. Such a universe can be described by the **Robertson–Walker metric**, which gives

$$\mathrm{d}s^2 = c^2\mathrm{d}t^2 - a^2(t)\left(\frac{\mathrm{d}r^2}{1 - kr^2} + r^2\,\mathrm{d}\vartheta^2 + r^2\sin^2\vartheta\,\mathrm{d}\varphi^2\right) \quad (13.5.1)$$

where a is a function of time known as the **scale factor**, and k is a constant related to the curvature:

- A **flat universe** has $k = 0$. The matter in a flat universe is at a critical density, ρ_c.
- A **closed universe** has $k = 1$. The matter density is $\rho > \rho_c$. Eventually the expansion will stop and the universe will start to contract.
- A **open universe** has $k = -1$. The matter density is $\rho < \rho_c$. The universe will expand indefinitely.

The Robertson–Walker metric is the only homogenous, isotropic metric, and so we can solve for a using Einstein's field equations. We do so by assuming uniform density, ρ, and pressure, p. In this case Einstein's field equations reduce to two differential equations for a, called the **Friedmann equations**:

$$\left(\frac{\dot{a}}{a}\right)^2 + \frac{k}{a^2} = \frac{8\pi G}{3c^4}\rho, \qquad \text{and} \qquad \frac{\ddot{a}}{a} = -\frac{8\pi G}{3}(\rho + 3p). \tag{13.5.2}$$

The rate of expansion is measured by the **Hubble constant** (which is not constant, but varies slowly), which is

$$H(t) \coloneqq \frac{\dot{a}(t)}{a(t)}. \tag{13.5.3}$$

This is related to the expansion rate by Hubble's law, which says that the velocity that a point at position $\boldsymbol{r}$ expands away from the origin is

$$\boldsymbol{v} = \frac{\mathrm{d}\boldsymbol{r}}{\mathrm{d}t} = H\boldsymbol{r}. \tag{13.5.4}$$

Our current measurements give

$$H = 68.3(12)\,\mathrm{km\,s^{-1}\,Mpc^{-1}} \tag{13.5.5}$$

where $1\,\mathrm{pc} = 3.08 \times 10^{16}\,\mathrm{m} = 3.26$ light years.

Setting $k = 0$, for a flat universe, we find that the critical density satisfies

$$\frac{\rho_c}{c^2} = \frac{3}{8\pi G}H^2 \approx 1 \times 10^{-29}\,\mathrm{g\,cm^{-3}}. \tag{13.5.6}$$

This corresponds to about five atoms per cubic metre. However, luminous stars only contribute about 0.2 atoms per cubic metre. We define the density of a type of matter, i, relative to the critical density as $\Omega_i = \rho_i/\rho_c$. We expect that $\sum_i \Omega_i = 1$. While this isn't true if we consider only luminous matter if we include dark matter and dark energy it could well be true. The current best guess as to the density of various types of matter is

$$\Omega(\text{stars}) = 0.4\,\%, \tag{13.5.7}$$
$$\Omega(\text{intergalactic gas}) = 3.6\,\%, \tag{13.5.8}$$
$$\Omega(\text{dark matter}) = 22\,\%, \tag{13.5.9}$$
$$\Omega(\text{dark energy}) = 74\,\%. \tag{13.5.10}$$

There are still many open questions, such as "what is dark matter/energy?" and "How can we make general relativity work with quantum mechanics?", but we finish here.

Acronyms

Index