

Willoughby Seago

**Theoretical Physics**

# **Relativity, Nuclear, and Particle Physics —Relativity**

September 20, 2021

**COURSE NOTES**

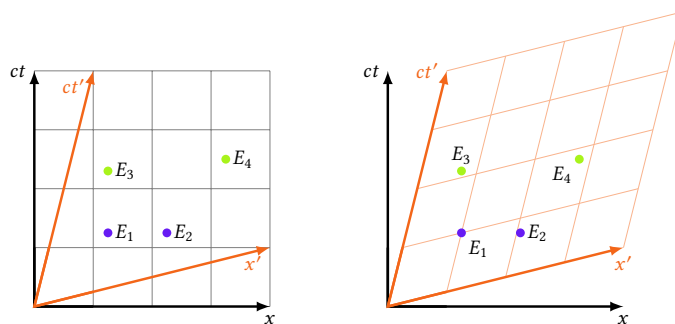
# Relativity, Nuclear and Particle Physics (Relativity)

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 17:08 on October 19, 2021. For notes on other topics see <https://github.com/WilloughbySeago/Uni-Notes>.



---

# Chapters

---

	Page
<b>Chapters</b>	<b>ii</b>
<b>Contents</b>	<b>iii</b>
<b>1 Introduction</b>	<b>1</b>
<b>I Classical Mechanics</b>	<b>2</b>
<b>2 Newton's Laws and Galilean Transformations</b>	<b>3</b>
<b>3 Two Body Systems</b>	<b>9</b>
<b>4 Particle Scattering</b>	<b>16</b>
<b>II Special Relativity</b>	<b>22</b>
<b>5 The Basics of Special Relativity</b>	<b>23</b>
<b>6 Relativistic Kinematics</b>	<b>33</b>
<b>7 Space Time Intervals</b>	<b>42</b>
<b>8 Four-Vectors</b>	<b>48</b>
<b>Acronyms</b>	<b>54</b>
<b>Index</b>	<b>55</b>

---

# Contents

---

	Page
<b>Chapters</b>	<b>ii</b>
<b>Contents</b>	<b>iii</b>
<b>1 Introduction</b>	<b>1</b>
<b>I Classical Mechanics</b>	<b>2</b>
<b>2 Newton's Laws and Galilean Transformations</b>	<b>3</b>
2.1 Newton's Laws . . . . .	3
2.2 Frame of Reference . . . . .	4
2.3 Galilean Transformations . . . . .	4
2.3.1 Invariance of Newton's Laws . . . . .	6
2.3.2 Newtonian Relativity Principle . . . . .	6
2.4 Accelerating Frames . . . . .	6
<b>3 Two Body Systems</b>	<b>9</b>
3.1 Conservation of Momentum . . . . .	9
3.2 Conservation of Energy . . . . .	9
3.2.1 Collisions . . . . .	10
3.3 Centre of Mass Coordinates . . . . .	10
3.4 Scattering . . . . .	12
3.4.1 Centre of Mass and Lab Frames . . . . .	12
3.5 Elastic Collisions . . . . .	13
<b>4 Particle Scattering</b>	<b>16</b>
4.1 Cross Sections . . . . .	16
4.1.1 Differential Cross Sections . . . . .	16
4.1.2 Total Cross Section . . . . .	17
4.2 Conservation Laws . . . . .	18
4.3 Hard Sphere Scattering in the cm Frame . . . . .	18
4.4 Rutherford Scattering . . . . .	19
4.4.1 History . . . . .	21
4.5 Moving Between Inertial Frames . . . . .	21

<b>II Special Relativity</b>	<b>22</b>
<b>5 The Basics of Special Relativity</b>	<b>23</b>
5.1 Breakdown of Galilean Relativity . . . . .	23
5.2 Postulates of Special Relativity . . . . .	23
5.3 Time: Synchronisation and Simultaneity . . . . .	24
5.4 Lorentz Transforms . . . . .	25
5.5 Minkowski Diagrams . . . . .	26
5.6 Consequences of the Lorentz Transformations . . . . .	26
5.6.1 Order of Events . . . . .	26
5.6.2 Length Contraction . . . . .	28
5.6.3 Time Dilation . . . . .	29
5.7 Deriving the Lorentz Transformation . . . . .	29
<b>6 Relativistic Kinematics</b>	<b>33</b>
6.1 Proper Time . . . . .	33
6.2 Relativistic Velocity Addition . . . . .	33
6.3 Relativistic Doppler . . . . .	35
6.4 High Speed Travel . . . . .	36
6.5 Twins Paradox . . . . .	36
6.6 Acceleration . . . . .	37
6.6.1 Constant Proper Acceleration . . . . .	38
6.7 Twins Paradox (Again) . . . . .	40
<b>7 Space Time Intervals</b>	<b>42</b>
7.1 Spatial Rotations . . . . .	42
7.2 Lorentz Transformations . . . . .	43
7.2.1 Four-Vectors . . . . .	43
7.2.2 Lorentz Transformation in Matrix Form . . . . .	43
7.2.3 Space-Time Interval . . . . .	45
7.2.4 Invariance of the Space-Time Interval . . . . .	45
7.2.5 Interval Classification . . . . .	46
<b>8 Four-Vectors</b>	<b>48</b>
8.1 Position . . . . .	48
8.2 Scalar Product . . . . .	49
8.3 Four-Velocity . . . . .	50
8.4 Four-Acceleration . . . . .	51
8.5 Four-Momentum . . . . .	52
8.5.1 Notes . . . . .	53
8.5.2 Lorentz Transformation of Four-Momentum . . . . .	53
<b>Acronyms</b>	<b>54</b>
<b>Index</b>	<b>55</b>

# 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

$$F := \dot{\mathbf{p}} \tag{2.1.1}$$

where we use the dot notation  $\dot{f} := df/dt$  for some differentiable quantity,  $f$ , and time  $t$ . Here  $\mathbf{p}$  is the linear momentum which, for Newtonian mechanics, is defined to be

$$\mathbf{p} := m\mathbf{v} = m\dot{\mathbf{r}} \tag{2.1.2}$$

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

We often deal with the special case where  $m$  is constant. In this case  $\dot{\mathbf{p}} = m\dot{\mathbf{v}} = m\ddot{\mathbf{r}} = m\mathbf{a}$ . 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

$$\mathbf{F} = -\frac{GMm}{r^2}\hat{\mathbf{r}} \quad (2.1.3)$$

where  $M$  and  $m$  are the masses of the two bodies,  $r$  is their separation,  $\hat{\mathbf{r}}$  is a unit vector pointing between the bodies, to the body upon which the force acts, and  $G = 6.6743 \text{ m}^3 \text{ kg}^{-2} \text{ s}^{-1}$  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 discuss 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:

$$x' = x - Vt, \quad y' = y, \quad z' = z, \quad \text{and} \quad t' = t. \quad (2.3.1)$$

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

$$\mathbf{r}' = \mathbf{r} - \mathbf{V}t, \quad \text{and} \quad t' = t. \quad (2.3.2)$$

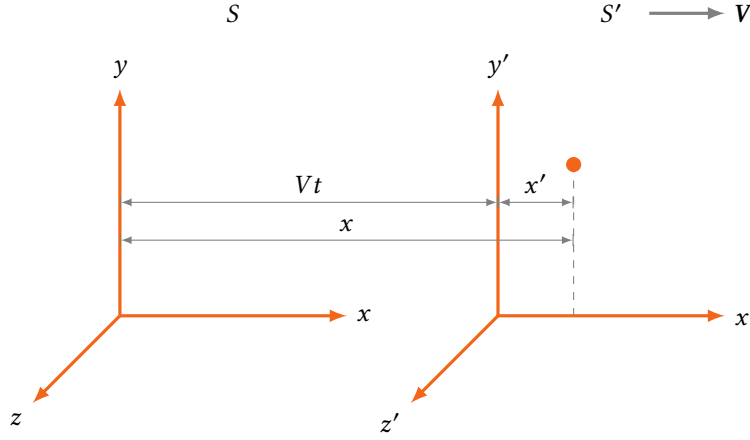


Figure 2.1: Two frames,  $S$  and  $S'$ , in the standard configuration with  $S'$  moving at speed  $V$  relative to  $S$ . Also shown is an event, here a dot, and how this relates to various lengths in the diagram.

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:

$$\dot{x}' = \dot{x} - V, \quad \dot{y}' = \dot{y}, \quad \text{and} \quad \dot{z}' = \dot{z}, \quad (2.3.3)$$

which implies

$$u'_x = u_x - V, \quad u'_y = u_y, \quad \text{and} \quad u'_z = u_z \quad (2.3.4)$$

or equivalently

$$\mathbf{u}' = \mathbf{u} - \mathbf{V} \quad (2.3.5)$$

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

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  $\mathbf{u}$  and  $\mathbf{v}$  in frame  $S$  which are  $\mathbf{u}'$  and  $\mathbf{v}'$  in frame  $S'$ . Then the relative velocity of these two objects in  $S'$  is

$$\mathbf{v}' - \mathbf{u}' = \mathbf{v} - \mathbf{V} - (\mathbf{u} - \mathbf{V}) = \mathbf{v} - \mathbf{u} \quad (2.3.6)$$

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  $F$  and a particle of constant mass,  $m$ . Then in some inertial frame,  $S$ , we have

$$F_x = m \frac{dv_x}{dt} = m \frac{d^2x}{dt^2} \quad (2.3.7)$$

and in another initial frame,  $S'$ , we have

$$F'_x = m \frac{dv'_x}{dt} = m \frac{d}{dt}(v_x - V) = m \frac{dv_x}{dt} = F_x \quad (2.3.8)$$

where we assume that  $V$  is constant. Similarly we can show that  $F_y = F'_y$  and  $F_z = F'_z$ . This shows that  $F' = F$  for inertial frames and hence  $a' = a$ . 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

$$\mathbf{r}' = \mathbf{r} - \mathbf{X}, \quad \text{and} \quad t' = t. \quad (2.4.1)$$

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

$$\mathbf{u}' = \mathbf{u} - \dot{\mathbf{X}} = \mathbf{u} - \mathbf{V} \quad (2.4.2)$$

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

$$\mathbf{V} = \mathbf{V}_0 + \mathbf{A}t \quad (2.4.3)$$

where  $\mathbf{V}_0$  is the speed of frame  $S'$  in  $S$  at  $t = 0$ . Substituting this into our equation we have

$$\mathbf{u}' = \mathbf{u} - \mathbf{V} = \mathbf{u} - \mathbf{V}_0 - \mathbf{A}t \quad (2.4.4)$$

differentiating gives us

$$\mathbf{a}' = \mathbf{a} - \mathbf{A}. \quad (2.4.5)$$

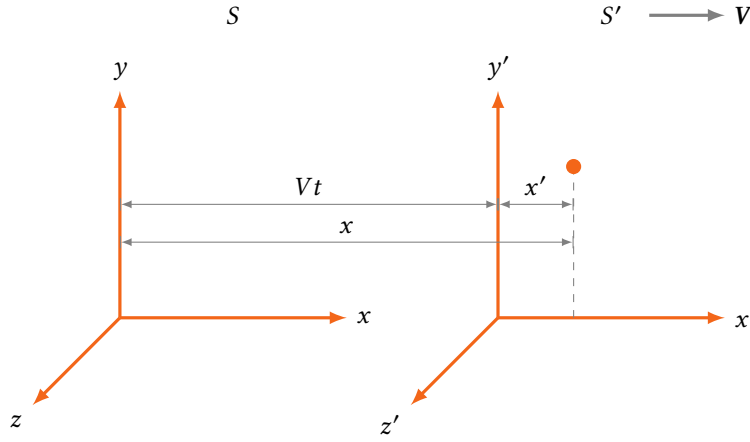


Figure 2.1: Two frames,  $S$  and  $S'$ , in the standard configuration with  $S'$  moving at speed  $V$  relative to  $S$ . Also shown is an event, here a dot, and how this relates to various lengths in the diagram.

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

$$m\mathbf{a}' = m\mathbf{a} - m\mathbf{A}. \quad (2.4.6)$$

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

$$\mathbf{F}' = m\mathbf{a} - m\mathbf{A}. \quad (2.4.7)$$

We see that the force is composed of two parts. The expected  $m\mathbf{a}$  and an extra term,  $m\mathbf{A}$ , 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,  $\mathbf{A}$ . 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

$$T \sin \vartheta = mA, \quad \text{and} \quad T \cos \vartheta - mg = 0 \quad (2.4.9)$$

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.

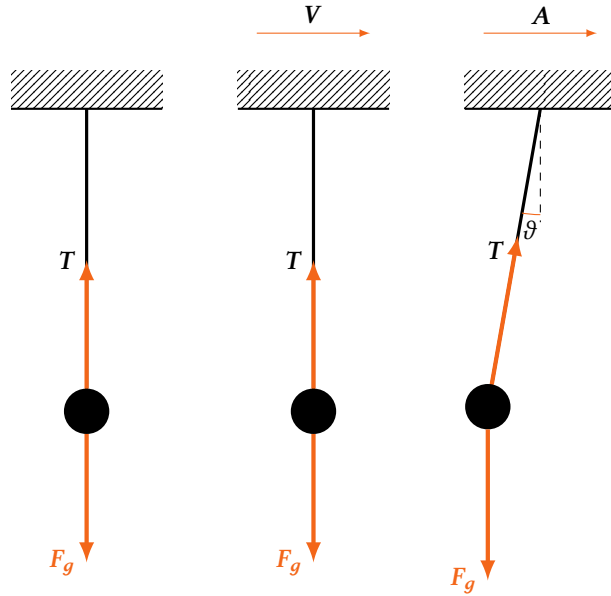


Figure 2.2: A pendulum on a train at rest (left), constant velocity (middle) and accelerating (left).

Working now in the frame of the train and accounting for the resulting fictitious force,  $mA$ , we can again resolve horizontally and vertically to get

$$T \sin \vartheta - mA, \quad \text{and} \quad T \cos \vartheta - mg = 0. \quad (2.4.10)$$

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  $\mathbf{r}_1$  and  $\mathbf{r}_2$ . We define  $\mathbf{f}$  to be the force on body one due to body two, and by N3  $-\mathbf{f}$  is the force on body two due to body one. We also consider the external forces  $\mathbf{F}_1$  and  $\mathbf{F}_2$  acting on bodies one and two respectively.

### 3.1 Conservation of Momentum

Newton's second law gives us the equations of motion

$$\mathbf{F}_1 + \mathbf{f} = m_1 \ddot{\mathbf{r}}_1 \quad (3.1.1)$$

$$\mathbf{F}_2 - \mathbf{f} = m_2 \ddot{\mathbf{r}}_2 \quad (3.1.2)$$

Adding together these equations we get

$$\mathbf{F}_1 + \mathbf{F}_2 = m_1 \ddot{\mathbf{r}}_1 + m_2 \ddot{\mathbf{r}}_2 = m_1 \dot{\mathbf{v}}_1 + m_2 \dot{\mathbf{v}}_2. \quad (3.1.3)$$

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

$$m_1 \dot{\mathbf{v}}_1 + m_2 \dot{\mathbf{v}}_2 = \mathbf{0}. \quad (3.1.4)$$

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

$$m_1 \mathbf{v}_1 + m_2 \mathbf{v}_2 = \mathbf{p}_1 + \mathbf{p}_2 = \text{constant}. \quad (3.1.5)$$

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  $\mathbf{f}$ 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  $\mathbf{F}$  acting on a particle to move it through an infinitesimal displacement  $d\mathbf{r}$  is  $dW = \mathbf{F} \cdot d\mathbf{r}$ . The work done to move the particle along a path from  $A$  to  $B$  is then

$$W_{AB} = \int_A^B \mathbf{F} \cdot d\mathbf{r} \quad (3.2.1)$$

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

$$W_{AB} = \int_A^B \mathbf{F} \cdot d\mathbf{r} \quad (3.2.2)$$

$$= \int_A^B \dot{\mathbf{p}} \cdot d\mathbf{r} \quad (3.2.3)$$

$$= m \int_A^B \dot{\mathbf{v}} \cdot d\mathbf{r} \quad (3.2.4)$$

$$= m \int_A^B \dot{\mathbf{v}} \cdot \frac{d\mathbf{r}}{dt} dt \quad (3.2.5)$$

$$= m \int_A^B \dot{\mathbf{v}} \cdot \mathbf{v} dt \quad (3.2.6)$$

$$= \frac{m}{2} \int_A^B \frac{d}{dt}(v^2) dt \quad (3.2.7)$$

$$= \frac{1}{2}mv_B^2 - \frac{1}{2}mv_A^2 \quad (3.2.8)$$

$$= T_B - T_A. \quad (3.2.9)$$

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

$$\frac{d}{dt}(v^2) = \frac{d}{dt}(\mathbf{v} \cdot \mathbf{v}) = \mathbf{v} \cdot \dot{\mathbf{v}} + \dot{\mathbf{v}} \cdot \mathbf{v} = 2\dot{\mathbf{v}} \cdot \mathbf{v} \implies \dot{\mathbf{v}} \cdot \mathbf{v} = \frac{1}{2} \frac{d}{dt}(v^2). \quad (3.2.10)$$

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

$$\frac{1}{2}m_1v_{1i}^2 + \frac{1}{2}m_2v_{2i}^2 = \frac{1}{2}m_1v_{1f}^2 + \frac{1}{2}m_2v_{2f}^2 \quad (3.2.11)$$

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

$$\frac{1}{2}m_1v_{1i}^2 + \frac{1}{2}m_2v_{2i}^2 > \frac{1}{2}m_1v_{1f}^2 + \frac{1}{2}m_2v_{2f}^2. \quad (3.2.12)$$

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  $\mathbf{R}$  which is such that

$$\mathbf{R} := \frac{m_1\mathbf{r}_1 + m_2\mathbf{r}_2}{m_1 + m_2}. \quad (3.3.1)$$

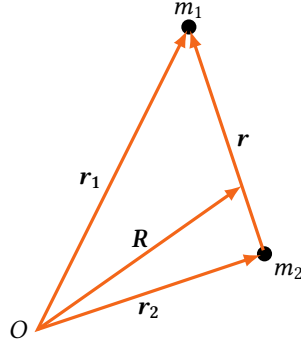


Figure 3.1: Two particles in the centre of mass frame. Drawn as if  $m_2 > m_1$  but this needn't be the case.

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

$$\mathbf{r} := \mathbf{r}_1 - \mathbf{r}_2. \quad (3.3.2)$$

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  $\mathbf{r}$  points along this line, from particle 2 to 1, we have

$$\mathbf{r}_1 = \mathbf{R} + \frac{m_2}{M}\mathbf{r}, \quad \text{and} \quad \mathbf{r}_2 = \mathbf{R} - \frac{m_1}{M}\mathbf{r}. \quad (3.3.3)$$

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

$$\mu := \frac{m_1 m_2}{m_1 + m_2}. \quad (3.3.4)$$

Using these we can work out [Equations \(3.1.1\)](#) and [\(3.1.2\)](#) in centre of mass co-ordinates in the absence of external forces. Substituting in the centre of mass and relative position vectors we get

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

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

Adding these together we get

$$\mathbf{0} = m_1 \ddot{\mathbf{R}} + \mu \ddot{\mathbf{r}} + m_2 \ddot{\mathbf{R}} - \mu \ddot{\mathbf{r}} = M \ddot{\mathbf{R}}. \quad (3.3.7)$$

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

$$\mathbf{f} = \mu \ddot{\mathbf{r}}. \quad (3.3.8)$$

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  $\mathbf{R}^* = \mathbf{0}$ .

**Notation 3.4.1 — CM Quantities** We denote quantities in the CM frame by an asterisk, \*. So, for example  $\mathbf{r}_1^*$  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

$$\mathbf{r}_1^* = \mathbf{r}_1 - \mathbf{R} = \frac{m_2}{M} \mathbf{r} = \frac{\mu}{m_1} \mathbf{r}, \quad \text{and} \quad \mathbf{r}_2^* = \mathbf{r}_2 - \mathbf{R} = -\frac{m_1}{M} \mathbf{r} = -\frac{\mu}{m_2} \mathbf{r}. \quad (3.4.2)$$

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

$$\mathbf{r}^* = \mathbf{r}_1^* - \mathbf{r}_2^* = \mathbf{r}_1 - \mathbf{r}_2 = \mathbf{r}. \quad (3.4.3)$$

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

$$m_1 \dot{\mathbf{r}}_1^* = -m_2 \dot{\mathbf{r}}_2^* = \mu \dot{\mathbf{r}} =: \mathbf{p}^*. \quad (3.4.4)$$

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  $\dot{\mathbf{R}}$  and the particles have velocities

$$\dot{\mathbf{r}}_1 = \dot{\mathbf{R}} + \dot{\mathbf{r}}_1^*, \quad \text{and} \quad \dot{\mathbf{r}}_2 = \dot{\mathbf{R}} + \dot{\mathbf{r}}_2^*. \quad (3.4.5)$$

The momenta of the particles will be

$$\mathbf{p}_1 = m_1 \dot{\mathbf{r}}_1 = m_1 \dot{\mathbf{R}} + \mathbf{p}^*, \quad \text{and} \quad \mathbf{p}_2 = m_2 \dot{\mathbf{r}}_2 = m_2 \dot{\mathbf{R}} - \mathbf{p}^* \quad (3.4.6)$$

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

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

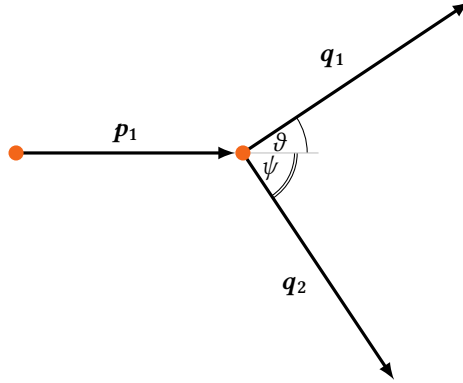


Figure 3.1: Elastic collision of particles in the LAB frame.

### 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  $\mathbf{p}_1$ . The momentum of the target particle is zero. After the collision we denote the momenta of the incoming and target particle by  $\mathbf{q}_1$  and  $\mathbf{q}_2$  respectively. We denote the angle through which the incoming and target particles are scattered by  $\vartheta$  and  $\psi$  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  $\mathbf{p}^*$  and the momentum of particle two is  $-\mathbf{p}^*$ . 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  $\mathbf{q}^*$ . See Figure 3.2.

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

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

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

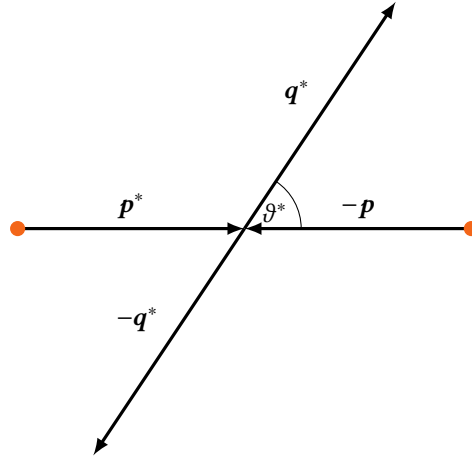


Figure 3.2: Elastic collision of particles in the CM frame.

Similarly after the collision we have

$$\mathbf{q}_1 = \mathbf{q}^* + m_1 \mathbf{R}, \quad (3.5.3)$$

$$\mathbf{q}_2 = -\mathbf{q}^* + m_2 \dot{\mathbf{R}}. \quad (3.5.4)$$

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

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

$$\mathbf{p}_2 = \mathbf{q}^* + \mathbf{p}^* \frac{m_1}{m_2}, \quad (3.5.6)$$

$$\mathbf{q}_2 = \mathbf{p}^* - \mathbf{q}^*. \quad (3.5.7)$$

Conservation of momentum also gives us  $\mathbf{p}_1 = \mathbf{q}_1 + \mathbf{q}_2$ . 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

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

This last relation in the special case of equal masses gives

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

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

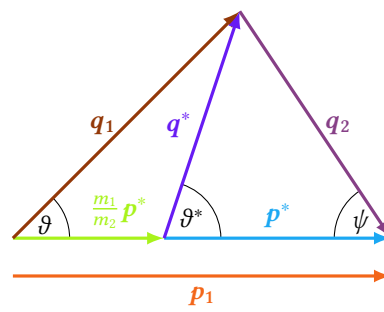


Figure 3.3: A vector diagram showing the relationships between the various momenta and angles in the LAB and CM frames.

# 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,  $\vartheta$ , 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,  $\varphi$ .

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  $\varphi$  and  $\varphi + d\varphi$ . If we change  $b$  slightly, say by  $db$ , then this will change the scattering angle,  $\vartheta$ , by a small amount,  $d\vartheta$ . We

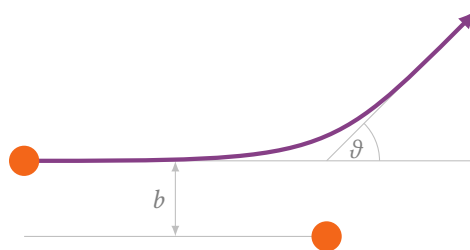


Figure 4.1: Scattering process showing the impact parameter,  $b$ , and deflection angle,  $\vartheta$ .

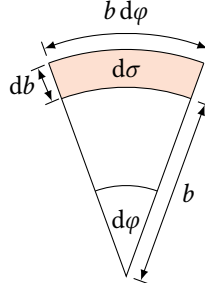


Figure 4.2: The cross section,  $d\sigma$ . Seen as if looking along the beam direction.

can interpret these small intervals of polar coordinates as defining an area,  $d\sigma$ . See Figure 4.2. Changing  $b$  also changes the **scattering rate**,  $w$ , by some small amount,  $dw$ , given by

$$dw = f d\sigma = f b d\varphi |db|. \quad (4.1.1)$$

Note that we take the absolute value of  $db$  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,  $d\Omega$ . We normalise this out by dividing by  $f$  and  $d\Omega$ .

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

$$dA = (L \sin \vartheta d\vartheta)(L d\varphi) \quad (4.1.2)$$

which gives

$$d\Omega = \sin \vartheta d\vartheta d\varphi. \quad (4.1.3)$$

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

$$\int_{S^2} \int_0^\pi \sin \vartheta d\vartheta \int_0^{2\pi} d\varphi = 4\pi. \quad (4.1.4)$$

Here  $S^2$  is the unit sphere.

It follows that the number of particles scattered in the direction defined by  $\vartheta$  and  $\varphi$  per second, per unit flux, per unit angular size is

$$\frac{dw}{f d\Omega} = \frac{d\sigma}{d\Omega} = \frac{b}{\sin \vartheta} \left| \frac{db}{d\vartheta} \right|. \quad (4.1.5)$$

We call the quantity  $d\sigma/d\Omega$  the **differential cross section**. It depends only on the relationship between the impact parameter,  $b$ , and the scattering angle,  $\vartheta$ . 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

$$\sigma := \int \frac{d\sigma}{d\Omega} d\Omega. \quad (4.1.6)$$

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,  $d\sigma^*/d\Omega^*$ , we need to find a relationship between  $b$  and  $\vartheta^*$ . 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  $\alpha^*$  to  $\vartheta^*$  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

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

Hence

$$\frac{d\sigma^*}{d\Omega^*} = \frac{b}{\sin \vartheta^*} \left| \frac{db}{d\vartheta^*} \right| = R^2. \quad (4.3.2)$$

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

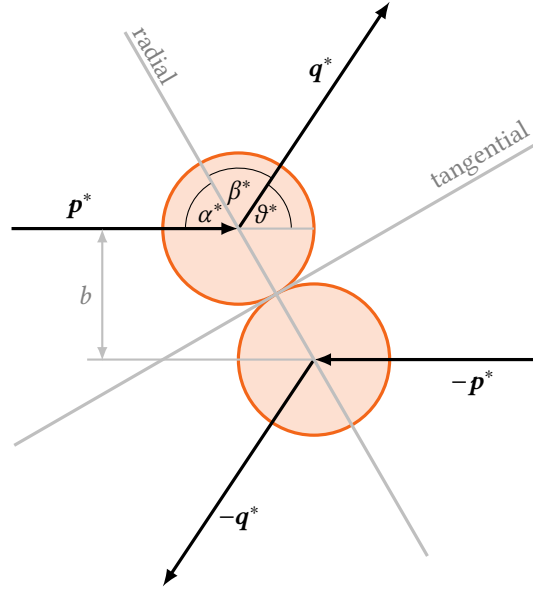


Figure 4.1: Two hard spheres colliding.

section is

$$\sigma^* = \int_{S^2} \frac{d\sigma^*}{d\Omega^*} d\Omega^* = R^2 \int_{S^2} d\Omega^2 = 4\pi R^2. \quad (4.3.3)$$

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 were 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. Its 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

$$\Delta p = 2mv \sin \frac{\vartheta}{2}. \quad (4.4.1)$$

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

$$\Delta p = \int_{-\infty}^{\infty} F \cos \Phi dt = \int_{-\infty}^{\infty} \frac{qQ}{4\pi\epsilon_0 r^2} \cos \Phi dt. \quad (4.4.2)$$



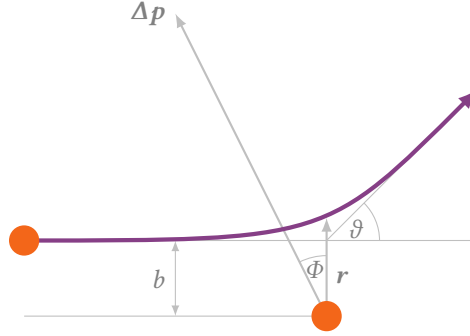


Figure 4.1: Rutherford scattering.

Here  $q$  and  $Q$  are the charge of the alpha particle and nucleus,  $r$  is the distance between them and  $\Phi$  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

$$L = mvb = mr^2 \frac{d\Phi}{dt} \implies \frac{d\Phi}{dt} = \frac{vb}{r^2} \implies dt = d\Phi \frac{r^2}{vb}. \quad (4.4.3)$$

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

$$2mv \sin \frac{\vartheta}{2} = \int_{-\Phi_{\max}}^{\Phi_{\max}} \frac{qQ}{4\pi\epsilon_0 r^2} \cos \Phi \frac{r^2}{vb} d\Phi = \frac{qQ}{4\pi\epsilon_0 vb} [\sin \Phi]_{-\Phi_{\max}}^{\Phi_{\max}} = \frac{2qQ}{4\pi\epsilon_0 vb} \cos \frac{\vartheta}{2}. \quad (4.4.4)$$

From this we have

$$b = \frac{qQ}{4\pi\epsilon_0 mv^2} \cot \frac{\vartheta}{2}. \quad (4.4.5)$$

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

$$r_c = a + \sqrt{a^2 + b^2}. \quad (4.4.6)$$

It then follows that

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

This is strongly dependent on the energy of the incoming particle and the scattering angle. It also goes to infinity as  $\vartheta \rightarrow 0$ . 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 were 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  $1 \times 10^{-14}$  m.

Subsequent experiments showed deviation from these predictions at higher energies, suggesting that the alpha particles were 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  $m_1/m_2 \rightarrow 0$  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,  $\vartheta$ , and solid angle,  $\Omega$ , but not the azimuthal angle,  $\varphi$ , or  $d\sigma$ . Therefore we have

$$\frac{d\sigma}{d\Omega} = \frac{d\sigma}{d\Omega^*} \frac{d\Omega^*}{d\Omega} \quad (4.5.1)$$

$$= \frac{d\sigma^*}{d\Omega^*} \frac{d\Omega^*}{d\Omega} \quad (4.5.2)$$

$$= \frac{d\sigma^*}{d\Omega^*} \frac{\sin \vartheta^* d\varphi d\vartheta^*}{\sin \vartheta d\varphi d\vartheta} \quad (4.5.3)$$

$$= \frac{d\sigma^*}{d\Omega^*} \frac{\sin \vartheta^* d\vartheta^*}{\sin \vartheta d\vartheta} \quad (4.5.4)$$

$$= \frac{d\sigma^*}{d\Omega^*} \frac{d(\cos \vartheta^*)}{d(\cos \vartheta)}. \quad (4.5.5)$$

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,  $3 \times 10^8 \text{ m s}^{-1}$ .

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 ether 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  $c = 2.998 \times 10^8 \text{ m s}^{-1}$  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  $c \rightarrow \infty$ .

### 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  $B$  and 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,  $\Delta t$ , 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  $\ell$  is the distance between the observer and  $A$ . However, the distance to  $B$  from the observer is also  $\ell$  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  $\ell$  is not what the “stationary” observer measures but very slightly less. Call it  $\ell'$ . The train observer deduces that the journey time for the two photons where

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

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**<sup>1</sup> 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.

<sup>1</sup>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.

$$x' = \gamma(v)[x - \beta ct], \quad (5.4.1)$$

$$y' = y, \quad (5.4.2)$$

$$z' = z, \quad (5.4.3)$$

$$ct' = \gamma(v)[ct - \beta x]. \quad (5.4.4)$$

Here

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

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

The speed of light appears as a limiting factor in the Lorentz transformation. If  $v > c$  then  $\gamma$  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  $\gamma$  is a function of  $v^2$  we have that  $\gamma(v) = \gamma(-v)$  and so

$$x = \gamma(v)[x' + \beta ct'], \quad (5.4.6)$$

$$y = y', \quad (5.4.7)$$

$$z = z', \quad (5.4.8)$$

$$ct = \gamma(v)[ct' + \beta x']. \quad (5.4.9)$$

## 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  $\hat{y}$  and  $\hat{z}$  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.

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  $\alpha$ , and satisfies

$$\tan \alpha = \frac{v}{c} = \beta. \quad (5.5.1)$$

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,  $\alpha$ , 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.

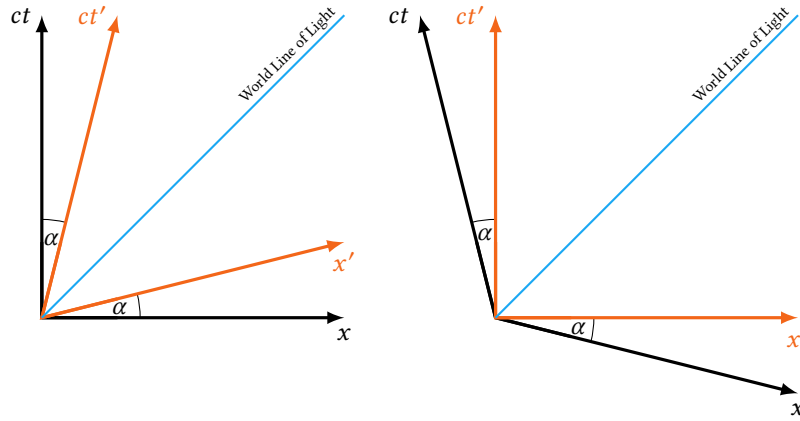


Figure 5.1: Minkowski diagrams showing two frames,  $S$  and  $S'$ , with axes  $(x, ct)$  and  $(x', ct')$  respectively. The world line of light (or any other light speed particle) is shown. The left hand side shows frame  $S$  with orthogonal axes and the right frame  $S'$ . The angle between axes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

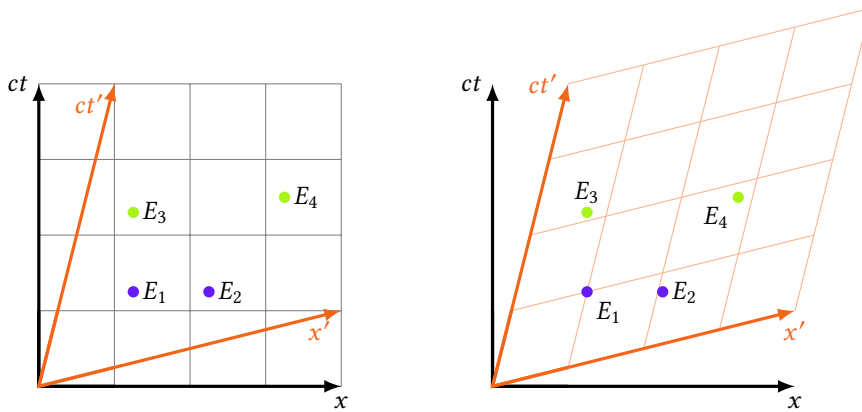


Figure 5.1: Minkowski diagram demonstrating how the order of events can differ between frames.



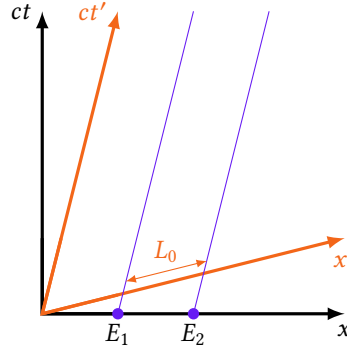


Figure 5.2: The setup for demonstrating length contraction.

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 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

$$L_0 = x'_2 - x'_1 \quad (5.6.1)$$

$$= \gamma[x_2 - \beta ct_2] - \gamma[x_1 - \beta ct_1] \quad (5.6.2)$$

$$= \gamma[x_2 - x_1] + \gamma\beta c[t_2 - t_1] \quad (5.6.3)$$

$$= \gamma[x_2 - x_1] \quad (5.6.4)$$

$$= \gamma L. \quad (5.6.5)$$

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

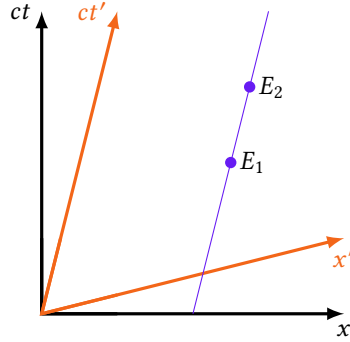


Figure 5.3: The setup for demonstrating time dilation.

of the proper length as

$$L = \frac{L_0}{\gamma}. \quad (5.6.6)$$

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

$$ct_2 - ct_1 = \gamma[ct'_2 + \beta x'_2] - \gamma[ct'_1 + \beta x'_1] \quad (5.6.7)$$

$$= \gamma c[t'_2 - t'_1] + \gamma\beta[x'_2 - x'_1] \quad (5.6.8)$$

$$= \gamma c[t'_2 - t'_1] \quad (5.6.9)$$

$$\implies t_2 - t_1 = \gamma[t'_2 - t'_1]. \quad (5.6.10)$$

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  $\tau$  passes in some frame then observers in a different frame will see the time  $t$  pass with

$$t = \gamma\tau \quad (5.6.11)$$

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

$$k := \frac{T}{T_0}. \quad (5.7.1)$$

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

$$k' := \frac{T_1}{T}. \quad (5.7.2)$$

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

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

We then find that

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

Trig tells us that

$$PE' = E_1P \tan \alpha = E_1P\beta \quad (5.7.5)$$

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,

$$(k^2 + 1)\beta = k^2 - 1 \quad (5.7.6)$$

which gives us

$$k = \sqrt{\frac{1 + \beta}{1 - \beta}}. \quad (5.7.7)$$

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

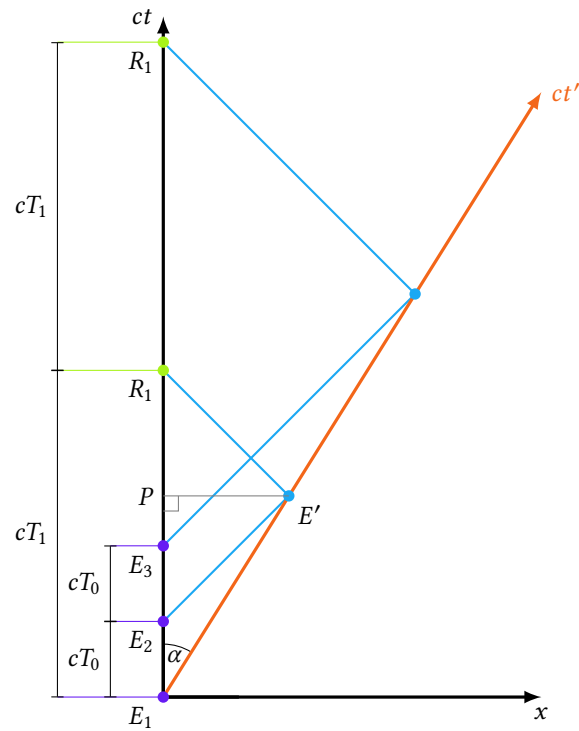


Figure 5.1: Two observers shine lights at each other.

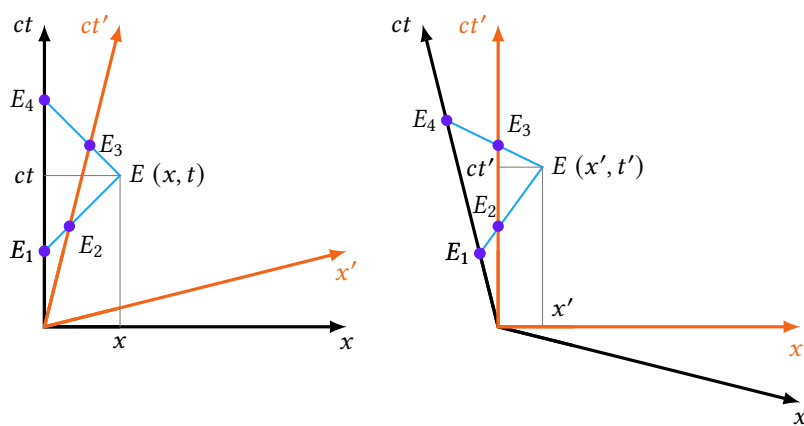


Figure 5.2: Reflection from an arbitrary point in a moving frame,

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

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

We can rewrite this as

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

Subtracting the first from the second we get

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

Using the definition of  $k$  we have

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

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

Hence we have

$$x' = \gamma(x - \beta ct). \quad (5.7.13)$$

Similarly if we add the two expressions then we instead get

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

which gives

$$ct' \gamma (ct - \beta x). \quad (5.7.15)$$

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  $dt$  passing between the endpoints of this world-line segment. For the observer the proper time  $d\tau$  passes between the endpoints. These are related by the time dilation formula [Equation \(5.6.11\)](#):

$$d\tau = \frac{dt}{\gamma(u)}. \quad (6.1.1)$$

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:

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

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

$$t = \int_{\text{start}}^{\text{end}} dt \quad (6.1.3)$$

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  $\mathbf{u} = (u_x, u_y, u_z)$  in frame  $S$  and  $\mathbf{u}' = (u'_x, u'_y, u'_z)$  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  $\mathbf{u}$  and  $\mathbf{u}'$ . To do so we write the Lorentz transformation in infinitesimal

form:

$$dx' = \gamma(v)[dx - \beta c dt], \quad (6.2.1)$$

$$dy' = dy, \quad (6.2.2)$$

$$dz' = dz, \quad (6.2.3)$$

$$cdt' = \gamma(v)[cdt - \beta dx]. \quad (6.2.4)$$

Taking the ratios of these quantities we find that

$$u'_x = \frac{dx'}{dt'} \quad (6.2.5)$$

$$= \frac{\gamma(v)[dx - \beta c dt]}{\gamma(v)[cdt - \beta dx]/c} \quad (6.2.6)$$

$$= \frac{dx - \beta c dt}{dt - \beta dx/c} \quad (6.2.7)$$

$$= \frac{\frac{dx}{dt} - \beta c}{1 - \frac{\beta}{c} \frac{dx}{dt}} \quad (6.2.8)$$

$$= \frac{u_x - \beta c}{1 - \frac{\beta}{c} u_x} \quad (6.2.9)$$

$$= \frac{u_x - v}{1 - \frac{vu_x}{c^2}}. \quad (6.2.10)$$

Similarly we find that

$$u'_y = \frac{dy'}{dt'} = \frac{u_y}{\gamma(v) \left[1 - \frac{vu_x}{c^2}\right]}, \quad (6.2.11)$$

$$u'_z = \frac{dz'}{dt'} = \frac{u_z}{\gamma(v) \left[1 - \frac{vu_x}{c^2}\right]}. \quad (6.2.12)$$

Similarly using the inverse Lorentz transformations we have

$$u_x = \frac{u'_x + v}{1 + \frac{vu'_x}{c^2}}, \quad (6.2.13)$$

$$u_y = \frac{u'_y}{\gamma(v) \left[1 + \frac{vu'_x}{c^2}\right]}, \quad (6.2.14)$$

$$u_z = \frac{u'_z}{\gamma(v) \left[1 + \frac{vu'_x}{c^2}\right]}. \quad (6.2.15)$$

Consider the special case of  $\mathbf{u} = \mathbf{0}$ . 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  $vu_x/c^2 \approx 0$ .

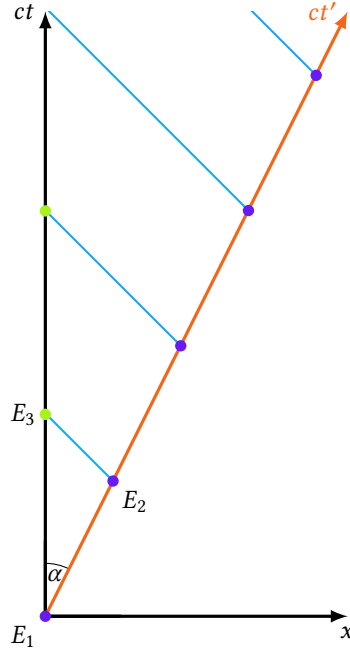


Figure 6.1: Device emitting electromagnetic waves.

### 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

$$ct_3 = ct_2 + \beta ct_2 = ct_2(1 + \beta). \quad (6.3.1)$$

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

$$ct_2 = \gamma(ct'_2 + \beta x'_2). \quad (6.3.2)$$

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

$$ct_3 = ct'_2 \gamma(1 + \beta) \quad (6.3.3)$$

$$= ct'_2 \frac{1 + \beta}{\sqrt{1 - \beta^2}} \quad (6.3.4)$$

$$= ct'_2 \frac{1 + \beta}{\sqrt{(1 - \beta)(1 + \beta)}} \quad (6.3.5)$$

$$= ct'_2 \sqrt{\frac{1 + \beta}{1 - \beta}}. \quad (6.3.6)$$



Hence the frequencies are related by

$$\nu = \nu' \sqrt{\frac{1 - \beta}{1 + \beta}}. \quad (6.3.7)$$

This is called the **relativistic longitudinal Doppler** formula. When  $\beta$  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,  $\mathbf{u}$ , into two components,  $u_r$ , radially along the line of sight, and  $u_t$ , transverse to the line of sight. Defining the angle  $\vartheta$  to satisfy  $\tan \vartheta = u_t/u_r$  and noting that the speed satisfies  $u^2 = u_r^2 + u_t^2$  we have

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

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:

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

## 6.4 High Speed Travel

Consider some long journey, from  $A$  to  $B$ , which are a distance  $\ell$  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,  $\ell$  is a proper length. Travelling at speed  $u$  the journey will be covered in time  $\ell/u$ . Therefore  $\ell/c$  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

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

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  $\gamma$  to  $\ell/\gamma$ . 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  $\ell$  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  $a = du/dt$ . In some other inertial frame  $S'$  (i.e. not the frame of the particle) what is the corresponding acceleration  $a' = du'/dt'$ ?

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):

$$\frac{du'}{dt'} = \frac{1}{\left(1 - \frac{uv}{c^2}\right)} \left[ \left(1 - \frac{uv}{c^2}\right) - \frac{(u-v)(-v)}{c^2} \right] \frac{du}{dt'} \quad (6.6.1)$$

$$= \frac{1 - \frac{v^2}{c^2}}{\left(1 - \frac{uv}{c^2}\right)} \frac{du}{dt'}. \quad (6.6.2)$$

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

$$\frac{du'}{dt'} = \gamma(u)^2 \frac{du}{dt'}. \quad (6.6.3)$$

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

$$\frac{du'}{dt'} = \gamma(u)^3 \frac{du}{dt}. \quad (6.6.4)$$

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

$$a_0 = \gamma^3 a. \quad (6.6.5)$$

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

$$\gamma(u'')^3 a'' = \gamma(u)^3 a \quad (6.6.6)$$

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

$$\frac{d}{dt}[\gamma(u)u] = \frac{d\gamma}{dt}u + \gamma \frac{du}{dt}. \quad (6.6.7)$$

For this we need to find  $d\gamma/dt$ :

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

$$= \frac{u}{c^2} \frac{du}{dt} \left( 1 - \frac{u^2}{c^2} \right)^{-3/2} \quad (6.6.9)$$

$$= \frac{u}{c^2} \frac{du}{dt} \gamma^3. \quad (6.6.10)$$

Hence,

$$\frac{d}{dt}[\gamma(u)u] = \frac{u^2}{c^2} \frac{du}{dt} \gamma^3 + \frac{du}{dt} \gamma. \quad (6.6.11)$$

With some algebra expanding the  $\gamma$  factors we can show that this is equivalent to

$$\frac{d}{dt}[\gamma(u)u] = \gamma^3 \frac{du}{dt}. \quad (6.6.12)$$

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

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

Rearranging this we find that

$$\frac{u}{c} = \frac{a_0 t / c}{\sqrt{1 + \frac{a_0^2 t^2}{c^2}}} = \frac{ct}{\sqrt{c^2 + a_0^2 t^2}} \quad (6.6.14)$$

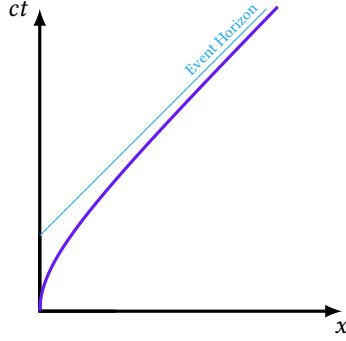


Figure 6.1: The world line of a particle undergoing constant acceleration is a hyperbola. The asymptote is called the event horizon.

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

The proper time experienced by an observer undergoing constant acceleration is

$$c\tau = c \int_{\text{start}}^{\text{end}} \frac{dt}{\gamma} \quad (6.6.15)$$

$$= c \int_{\text{start}}^{\text{end}} \sqrt{1 - \frac{u^2}{c^2}} dt \quad (6.6.16)$$

$$= c\rho \int_{\text{start}}^{\text{end}} \frac{1}{\sqrt{\rho^2 + c^2 t^2}} dt \quad (6.6.17)$$

$$= \rho \left[ \operatorname{arsinh} \left( \frac{ct}{\rho} \right) \right]_{\text{start}}^{\text{end}}. \quad (6.6.18)$$

Integrating  $u = dx/dt$  we get

$$x = \int \frac{ct}{\sqrt{\rho^2 + c^2 t^2}} c dt \quad (6.6.19)$$

$$= \sqrt{\rho^2 + c^2 t^2} + C \quad (6.6.20)$$

$$= \sqrt{\rho^2 + c^2 t^2} + x_0 - \rho \quad (6.6.21)$$

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

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

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  $\rho$ , 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

$$x - x_0 = \frac{c^2 t^2}{2\rho} = \frac{1}{2} a_0 t^2, \quad (6.6.23)$$

which is a familiar classical result.

### 6.6.1.2 Highly Relativistic Regime

Rearranging the result of Equation (6.6.22) we have

$$ct = \sqrt{(x - x_0 + \rho)^2 - \rho^2} \quad (6.6.24)$$

$$= \sqrt{(x - x_0)^2 + 2\rho(x - x_0)} \quad (6.6.25)$$

$$= (x - x_0) \sqrt{1 + \frac{2\rho}{x - x_0}}. \quad (6.6.26)$$

For any  $\rho$  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

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

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,  $\ell$ , 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

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

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

$$c\tau_1 = \rho \left[ \operatorname{arsinh} \left( \frac{ct}{\rho} \right) \right]_0^{t_1} = \rho \operatorname{arsinh} \left( \frac{v/c}{\sqrt{1 - \frac{v^2}{c^2}}} \right). \quad (6.7.2)$$

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

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

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

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

Hence the age difference of the twins is

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

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

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

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

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  $\vartheta$  anti-clockwise about the  $z$ -axis then the same point has the coordinates  $(x', y', z')$ , which are given by

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

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<sup>1</sup>. We can show this by considering  $\mathbf{r} \cdot \mathbf{r}$ :

<sup>1</sup>or “invariant under the action of  $O(3)$ ” if we're feeling fancy

$$\mathbf{r} \cdot \mathbf{r} = x'^2 + y'^2 + z'^2 \quad (7.1.2)$$

$$= (x \cos \vartheta + y \sin \vartheta)^2 + (-x \sin \vartheta + y \cos \vartheta)^2 + z^2 \quad (7.1.3)$$

$$= x^2 \cos^2 \vartheta + 2xy \cos \vartheta \sin \vartheta + y^2 \sin^2 \vartheta + x^2 \sin^2 \vartheta \quad (7.1.4)$$

$$- 2xy \sin \vartheta \cos \vartheta + y^2 \cos^2 \vartheta + z^2 \quad (7.1.5)$$

$$= x^2 (\cos^2 \vartheta + \sin^2 \vartheta) + y^2 (\cos^2 \vartheta + \sin^2 \vartheta) + z^2 \quad (7.1.6)$$

$$= x^2 + y^2 + z^2. \quad (7.1.7)$$

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  $\mathbf{r}_2 - \mathbf{r}_1$  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:

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

Here  $\mu$  (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  $x^\mu$ ,  $x_0$ , 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

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

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

$$\cosh^2 u - \sinh^2 u = 1. \quad (7.2.3)$$

Rearranging this gives us

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

Hence

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



Using these relations we have

$$x' = \gamma x - \beta \gamma ct = \cosh(\omega)x - \sinh(\omega)ct \quad (7.2.6)$$

$$ct' = \gamma ct - \beta \gamma x = \cosh(\omega)ct - \sinh(\omega)x \quad (7.2.7)$$

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

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

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

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

This allows us to compactly write the Lorentz transformation of  $x^\mu$  as

$$x'^\mu = \Lambda^\mu_\nu x^\nu. \quad (7.2.10)$$

**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 ( $\mu, \nu, \sigma$ , etc.) take the values 0, 1, 2, 3.

For example,

$$\mathbf{a} \cdot \mathbf{b} = a_i b^i = a_1 b^1 + a_2 b^2 + a_3 b^3. \quad (7.2.12)$$

$$a \cdot b = a_\mu b^\mu = a_0 b^0 + a_1 b^1 + a_2 b^2 + a_3 b^3 = a_0 b^0 + \mathbf{a} \cdot \mathbf{b}. \quad (7.2.13)$$

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:

$$x'^\mu = \Lambda^\mu_\nu x^\nu = \Lambda^\mu_0 x^0 + \Lambda^\mu_1 x^1 + \Lambda^\mu_2 x^2 + \Lambda^\mu_3 x^3 = \Lambda^\mu_0 ct + \Lambda^\mu_1 x + \Lambda^\mu_2 y + \Lambda^\mu_3 z. \quad (7.2.14)$$

Now consider specific values of  $\mu$ :

$$x'^0 = ct' = \Lambda^0_0 ct + \Lambda^0_1 x + \Lambda^0_2 y + \Lambda^0_3 z = \gamma ct' - \beta \gamma x + 0y + 0z = \gamma ct - \beta \gamma x$$

$$x'^1 = x' = \Lambda^1_0 ct + \Lambda^1_1 x + \Lambda^1_2 y + \Lambda^1_3 z = -\beta \gamma ct' + \gamma x + 0y + 0z = \gamma x - \beta \gamma ct$$

$$x'^2 = y' = \Lambda^2_0 ct + \Lambda^2_1 x + \Lambda^2_2 y + \Lambda^2_3 z = 0ct' + 0x + 1y + 0z = y$$

$$x'^3 = z' = \Lambda^3_0 ct + \Lambda^3_1 x + \Lambda^3_2 y + \Lambda^3_3 z = 0ct' + 0x + 0y + 1z = z$$

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,  $d\tau$ :

$$c d\tau = \frac{c dt}{\gamma} \quad (7.2.15)$$

$$= c dt \sqrt{1 - \frac{u^2}{c^2}} \quad (7.2.16)$$

$$= c dt \sqrt{1 - \frac{1}{c^2} \left( \frac{dx}{dt} \right)^2 - \frac{1}{c^2} \left( \frac{dy}{dt} \right)^2 - \frac{1}{c^2} \left( \frac{dz}{dt} \right)^2} \quad (7.2.17)$$

$$= \sqrt{c^2 dt^2 - dx^2 - dy^2 - dz^2}. \quad (7.2.18)$$

It follows that

$$c^2 d\tau^2 = c^2 dt^2 - dx^2 - dy^2 - dz^2, \quad (7.2.19)$$

or, in a more condensed notation

$$ds^2 := c^2 d\tau^2 = g_{\mu\nu} dx^\mu dx^\nu. \quad (7.2.20)$$

Here  $dx^\mu = (c dt, dx, dy, dz)$ , and  $g$  is the **metric tensor**, specifically the **Minkowski metric**, defined as

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

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

The metric tensor plays the role of the Kronecker delta,  $\delta$ , in the normal dot product,  $\mathbf{a} \cdot \mathbf{b} = \delta_{ij} a^i b^j = a_i b^i$ , 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  $ds^2$  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

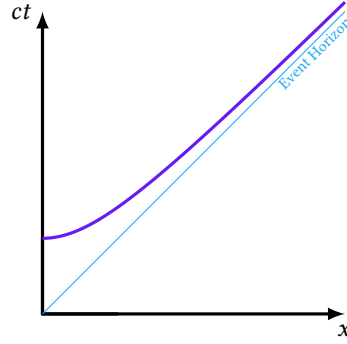


Figure 7.1: A line of constant space-time interval,  $c^2t^2 - x^2 = \text{constant}$ . Notice the gradient is less than 1 and so this represents an object moving faster than the speed of light.

fairly easy to show for frames in the standard configuration:

$$s'^2 = g_{\mu\nu}x'^{\mu}x'^{\nu} \quad (7.2.22)$$

$$= (ct')^2 - x'^2 - y'^2 - z'^2 \quad (7.2.23)$$

$$= \gamma^2(ct - \beta x)^2 - \gamma^2(x - \beta ct)^2 - y^2 - z^2 \quad (7.2.24)$$

$$= \gamma^2(ct)^2 - \gamma^2\beta c t x + \gamma^2\beta^2 x^2 - \gamma^2 x^2 + \gamma^2\beta x c t - \gamma^2\beta^2(ct)^2 - y^2 - z^2 \quad (7.2.25)$$

$$= (\gamma^2 - \gamma^2\beta^2)(ct)^2 - (\gamma^2 - \gamma^2\beta^2)x^2 - y^2 - z^2 \quad (7.2.26)$$

$$= (ct)^2 - x^2 - y^2 - z^2 \quad (7.2.27)$$

$$= s^2. \quad (7.2.28)$$

Here we have used

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

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  $ct > |\mathbf{x}|$ .

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

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

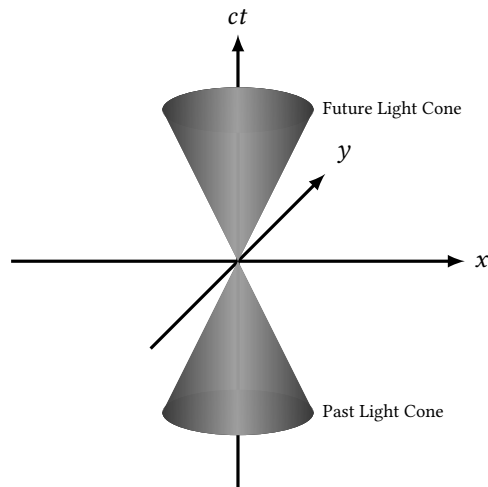


Figure 7.2: Light cones drawn in two spatial dimensions and time.

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  $45^\circ$ , 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,  $\mathbf{r} = (x, y, z)$ , and a time,  $t$ , the position four-vector is

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

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  $\mathbf{r}$  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

$$|\mathbf{r}|^2 = \mathbf{r} \cdot \mathbf{r} = x_i x^i = \delta_{ij} x^i x^j = x^2 + y^2 + z^2, \quad (8.1.2)$$

or in matrix form:

$$|\mathbf{r}|^2 = \mathbf{r} \cdot \mathbf{r} = \mathbf{r}^\top \mathbf{r} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}^\top \begin{pmatrix} x \\ y \\ z \end{pmatrix} = (x \quad y \quad z) \begin{pmatrix} x \\ y \\ z \end{pmatrix} = x^2 + y^2 + z^2 \quad (8.1.3)$$

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:

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

or in matrix form

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

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

$$g = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \quad (8.1.6)$$

Looking at these we can make the identification

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

Formally  $x^\mu$  is a **contravariant** four-vector, and  $x_\mu$  is a **covariant** four-vector. We can think of the act of raising and lowering indices by transposing and multiplying by the metric<sup>1</sup>.

**R** The same notation,  $x^\mu$ , 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.

<sup>1</sup>if the metric is not orthogonal and symmetric then sometimes you will need to multiply by the inverse of the metric

## 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  $a^\mu$  and  $b^\mu$  is defined as

$$a \cdot b = a_\mu b^\mu = g_{\mu\nu} a^\nu b^\mu = a^0 b^0 - a^1 b^1 - a^2 b^2 - a^3 b^3 = a^0 b^0 - \mathbf{a} \cdot \mathbf{b}. \quad (8.2.2)$$

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

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

*Proof.* Let  $a^\mu$  and  $b^\mu$  be four-vectors in some frame  $S$ . The same four-vectors in frame  $S'$  have the components

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

and similar for  $b^\mu$ .

Hence,

$$a \cdot b = a'_\mu b'^\mu \quad (8.2.5)$$

$$= g_{\mu\nu} a'^\nu b'^\mu \quad (8.2.6)$$

$$= a'^0 b'^0 - a'^1 b'^1 - a'^2 b'^2 - a'^3 b'^3 \quad (8.2.7)$$

$$= \gamma^2 (a^0 - \beta a^1)(b^0 - \beta b^1) - \gamma^2 (a^1 - \beta a^0)(b^1 - \beta b^0) - a^2 b^2 - a^3 b^3 \quad (8.2.8)$$

$$= \gamma^2 [a^0 b^0 (1 - \beta^2) - a^1 b^1 (\beta^2 - 1)] - a^2 b^2 - a^3 b^3 \quad (8.2.9)$$

$$= a^0 b^0 - a^1 b^1 - a^2 b^2 - a^3 b^3. \quad (8.2.10)$$

Here we have used

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

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

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

*Proof.* Consider the scalar product of  $a^\mu$  and  $b^\mu$ . In index notation we have

$$a \cdot b = a'_\mu b'^\mu \quad (8.2.13)$$

$$= g_{\mu\nu} a'^\nu b'^\mu \quad (8.2.14)$$

$$= g_{\mu\nu} \Lambda^\nu_\sigma a^\sigma \Lambda^\mu_\rho b^\rho \quad (8.2.15)$$

$$= \Lambda_{\mu\sigma} a^\sigma \Lambda^\mu_\rho b^\rho \quad (8.2.16)$$

$$= g_{\sigma\kappa} \Lambda_\mu^\kappa a^\sigma \Lambda^\mu_\rho b^\rho \quad (8.2.17)$$

$$= g_{\sigma\kappa} \Lambda_\mu^\kappa \Lambda^\mu_\rho a^\sigma b^\rho \quad (8.2.18)$$

$$= g_{\sigma\kappa} (\Lambda^\top \Lambda)^\kappa_\rho a^\sigma b^\rho \quad (8.2.19)$$

By the invariance of scalar products, [Theorem 8.2.3](#), we know that this must be equal to

$$a \cdot b = a_\rho b^\rho = g_{\sigma\rho} a^\sigma b^\rho. \quad (8.2.20)$$

So we can identify

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

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  $\kappa$  for the index  $\rho$  when  $\kappa$  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  $\Lambda^\top = \Lambda^{-1}$ .  $\square$

### 8.3 Four-Velocity

We have already seen how the three-velocity,  $\mathbf{u} = \dot{\mathbf{r}}$  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

$$ds^2 = c^2 dt^2 - dx^2 - dy^2 - dz^2 = c^2 d\tau^2. \quad (8.3.1)$$

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

This means that

$$u^\mu := \frac{dx^\mu}{d\tau} \quad (8.3.2)$$

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

$$u^\mu = \frac{d}{d\tau}(ct, \mathbf{r}) = \left( c \frac{dt}{d\tau}, \frac{d\mathbf{r}}{d\tau} \right) = \left( \gamma c, \gamma \frac{d\mathbf{r}}{dt} \right) = (\gamma c, \gamma \mathbf{u}). \quad (8.3.3)$$

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

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

$$u'^0 = \gamma(v)[u^0 - \beta u^1] \quad (8.3.4)$$

$$u'^1 = \gamma(v)[u^1 - \beta u^0] \quad (8.3.5)$$

$$u'^2 = u^2 \quad (8.3.6)$$

$$u'^3 = u^3. \quad (8.3.7)$$

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  $\mathbf{u} = \mathbf{0}$  and  $\gamma = 1$ :

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

Notice that we are using  $\gamma$  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

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

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:

$$A^\mu := \frac{d^2 x^\mu}{d\tau^2} = \frac{du^\mu}{d\tau}. \quad (8.4.1)$$

This has components

$$A^\mu = \frac{d}{d\tau}(\gamma c, \gamma \mathbf{u}) = \gamma \frac{d}{dt}(\gamma c, \gamma \mathbf{u}). \quad (8.4.2)$$

Using [Equation \(6.6.10\)](#) we have

$$\frac{d\gamma}{dt} = \frac{u}{c^2} \frac{du}{dt} \gamma^3 = \gamma^3 \frac{\mathbf{u} \cdot \mathbf{a}}{c^2} \quad (8.4.3)$$



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

$$\frac{d}{dt}(\gamma \mathbf{u}) = \gamma \mathbf{a} + \gamma^3 \frac{\mathbf{u} \cdot \mathbf{a}}{c^2} \mathbf{u}. \quad (8.4.4)$$

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

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

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  $\mathbf{a}_0$  is the proper acceleration. In this frame the four-acceleration and four-velocity are

$$A_{\text{ICMF}}^\mu = (0, \mathbf{a}_0), \quad \text{and} \quad u_{\text{ICMF}}^\mu = (c, \mathbf{0}). \quad (8.4.6)$$

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

$$A_{\text{ICMF}} \cdot u_{\text{ICMF}} = 0. \quad (8.4.7)$$

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

$$A_\mu A^\mu = -\mathbf{a}_0 \cdot \mathbf{a}_0 = -a_0^2 < 0. \quad (8.4.8)$$

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  $m\mathbf{v}$ , 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

$$p^\mu = mu^\mu = m \frac{dx^\mu}{d\tau}, \quad (8.5.1)$$

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

$$p^\mu = (\gamma mc, \gamma m\mathbf{u}) \quad (8.5.2)$$

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

$$p_\mu p^\mu = m^2 c^2 \quad (8.5.3)$$

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

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

where  $\mathbf{p} := \gamma m\mathbf{u}$  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  $\gamma$ ,  $\mathbf{p} = \gamma m \mathbf{u}$ . In the non-relativistic limit we have  $\gamma \rightarrow 1$  and so recover the Newtonian momentum,  $m \mathbf{u}$ . This relativistic three-momentum is conserved in the absence of forces.

Some sources will speak of a “relativistic mass”, which is given by  $\gamma m$ , 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

$$p'^{\mu} = \Lambda^{\mu}_{\nu} p^{\nu}, \quad (8.5.5)$$

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

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

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

$$P'^2 = P^2, \quad (8.5.8)$$

$$P'^3 = P^3. \quad (8.5.9)$$

This should be no surprise at this point.

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

$$\frac{E'}{c} = \gamma \frac{E}{c} - \gamma \beta P_{\parallel} \quad (8.5.10)$$

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

$$\mathbf{P}'_{\perp} = \mathbf{P}_{\perp}. \quad (8.5.12)$$

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

---

# Acronyms

---

## C

CM: centre of mass [12](#)

## G

GR: general relativity [1](#)

## I

ICMF: instantaneously co-moving inertial frame [37](#)

## N

N1: Newton's first law [3](#)

N2: Newton's second law [3](#)

N3: Newton's third law [3](#)

NG: Newton's law of gravitation [3](#)

## S

SR: special relativity [1](#)

---

# Index

---

## C

causality, 47  
centre of mass frame, 12  
centre of momentum frame, 12  
CM frame, *see* centre of mass frame  
contravariant, 49  
covariant, 49  
cross section, 17

## D

differential cross section, 17  
Doppler effect, 35

## E

Einstein summation convention, 44  
elastic collision, 10, 13  
event, 4  
event horizon, 40

## F

fictitious force, 7  
four-acceleration, 51  
four-momentum, 52  
four-position, 48  
four-vector, 43, 48  
four-velocity, 51

## G

Galilean transformation, 4  
gravitational constant, 4

## I

impact parameter, 16  
incident flux, 16  
inelastic collisions, 10  
inertial force, 7  
inertial frame, 4  
instantaneous rest frame, 37

instantaneously co-moving inertial frame,  
*see* instantaneous rest frame  
invariant, 5, 42  
isotropic, 18

## K

$k$ -calculus, 29  
kinetic energy, 10

## L

LAB frame, 12  
length contraction, 29  
light-like, 46  
Lorentz boost, 53  
Lorentz transformation, 25

## M

metric tensor, 45  
Minkowski diagram, 26  
Minkowski metric, 45

## N

Newton's first law, 3  
Newton's law of gravitation, 3  
Newton's second law, 3  
Newton's third law, 3  
Newtonian principle of relativity, 6

## P

principle of relativity, 24  
proper acceleration, 37  
proper length, 28  
proper time, 29

## R

rapidity, 43  
red shift, 36  
reduced mass, 11  
relativistic energy, 52

relativistic longitudinal Doppler, 36  
relativistic momentum, 52  
relativistic transverse Doppler, 36

**S**

scalar product, 49  
scattering rate, 17  
solid angle, 17  
space-like, 46  
standard configuration, 4

**T**

time dilation, 29  
time-like, 46  
total cross section, 17  
twins paradox, 37

**U**

universal time, 4

**W**

work done, 9  
world line, 26

**Z**

zero momentum frame, 12