Special Relativity

Raphael Esquivel

September 26, 2024

1 Introduction

Welcome to Fun Meeting #1 of 'Iolani Physics Club! Today we will be talking about Einstein's theory of special relativity. While the lecture might not get to the end of the contents in these notes, extra content has been provided to you as a reference. These topics will show up in exams like Physics Bowl and Physics Brawl. Some derivations which aren't particularly helpful but may be interesting are left in an appendix. If you have any questions after the meeting, feel free to email rre2601@iolani.org. Special thanks to Emir Tataroğlu for helping write problems. Solutions will be posted in the Google Drive.

2 Postulates

Definition 2.1. *An inertial frame of reference is a stationary or uniformly moving frame of reference.*

Definition 2.2. *An event is anything that happens at a specific point in space, at a specific point in time.*

An event is described by coordinates of space and time (spacetime coordinates). The separation between two events in spacetime is defined to be

$$\Delta s^2 = c^2\Delta t^2 - \Delta x^2 - \Delta y^2 - \Delta z^2 \tag{1}$$

where c is the speed of light. Spacetime is four-dimensional (three spatial dimensions, one temporal dimension), but note that it is not particularly *Euclidean* four-dimensional space. In Euclidean space, an interval would be given by the Pythagorean theorem, but there are minus signs and scaling factors in our definition. Spacetime is a different sort of space, called *Minkowski space.*[1]

Theorem 2.3. *(Galilean relativity principle) The laws of motion are the same in all inertial frames of reference.*

In his *Dialogo sopra i due massimi sistemi del mondo*, Galileo uses a ship as an example. If a ship travels at constant velocity on a smooth sea, an observer below the deck would not be able to tell whether the ship was moving or stationary. In 1904, on the basis of experimental facts, Henri Poincare generalized this principle to all natural phenomena.

[1]More information on Minkowski space can be found in the derivation for the energy-momentum relation in the appendix.

Theorem 2.4. *(Postulate of relativity) "The laws of physical phenomena must be identical for an observer at rest and for an observer undergoing uniform rectilinear motion, so we have no way and cannot have any way for determining whether we are undergoing such motion or not".*

A very important second postulate, originally proposed by Einstein, is the following:

Theorem 2.5. *(Postulate of the constancy of the velocity of light) The velocity of light is independent of the motion of the source.*

This allows us to deduce the connection between the space-time coordinates in different inertial reference frames.

3 The Lorentz Transformation

Consider two coordinate systems K and K'. System K' has its axes parallel to K but it is moving with a velocity v in the positive z direction relative to the system K. Points in space and time in the two systems are specified by (ct, x, y, z) and $(ct', x', y', z',)$ respectively, where c is again the speed of light. For convenience we suppose that a common origin of time $t = t' = 0$ is chosen at the instant when the two sets of coordinate axes exactly overlap. Now, imagine an observer in each reference frame equipped with a machine that detects the arrival time of a light signal from the origin at various points in space. If there is a light source at rest in the system K which is flashed on and off rapidly at $t = t' = 0$, then Einstein's second postulate implies that each observer will see his machine respond to a spherical shell of radiation moving outward from his origin of coordinates with velocity c. The arrival time t of the pulse at a detector located at (x, y, z) in K will satisfy

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

Similarly, in K', for a detector at (x', y', z')

$$x'^2 + y'^2 + z'^2 - c^2t'^2 = 0 \tag{3}$$

Because the transformation between the coordinate systems is linear, we can write

$$x'^2 + y'^2 + z'^2 - c^2t'^2 = \lambda^2(x^2 + y^2 + z^2 - c^2t^2) \tag{4}$$

where λ is a function of velocity, and $\lambda(0) = 1$; when the coordinate frames have zero velocity relative to each other, the recorded measurements will obviously be the same. The presence of λ allows for the possibility of an overall scale change in going from K to K'. But the shells of radiation are spheres in both systems. From the hypothesis that K' is moving parallel to the z axis of K it is evident that

$$\begin{aligned} x' &= \lambda x \\ y' &= \lambda y \end{aligned} \tag{5}$$

Then the most general linear connection between z', t' and z, t is

$$\begin{aligned} z' &= \lambda(a_1 z + a_2 t) \\ t' &= \lambda(b_1 t + b_2 z) \end{aligned} \tag{6}$$

where we have factored out a λ for convenience. Because the position of the origin of K' is specified by $z = vt$, solving for the constants shows $a_2 = -va_1$. Consider that $v = 0$, the constants become $a_1 = 1$, $a_2 = 0$, $b_1 = 1$, $b_2 = 0$. Choosing the signs to agree with our case at $v = 0$, we do a little algebra to find that

$$\begin{aligned} a_1 = b_1 &= \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}} \\ b_2 &= -\frac{v}{c^2} a_1 \end{aligned} \tag{7}$$

This quantity is so important in special relativity that we give it a name:

Definition 3.1. *The Lorentz factor γ is*

$$\gamma = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}} \tag{8}$$

with v the relative velocity between two frames.

The ratio v/c is often written β. The remaining problem is solving for λ. Considering a third reference frame K'' moving a velocity $-v$ parallel to the z axis to K', the coordinates can be obtained in terms of (x', y', z', t') by a change $v \to -v$. But K'' is the same as K, leading to the requirement that

$$\lambda(v)\lambda(-v) = 1 \tag{9}$$

$\lambda(v)$ must be independent of sign because it represents a scale change in the transverse direction. Therefore $\lambda(v) = 1$ for *all* v. We finally write down the Lorentz transformation.

Theorem 3.2. *The Lorentz transformation connects the coordinates in K' to K:*

$$\begin{aligned} z' &= \gamma(z - vt) \\ t' &= \gamma\left(t - \frac{v}{c^2} z\right) \\ x' &= x \\ y' &= y \end{aligned} \tag{10}$$

Another way of expressing the Lorentz transformation in matrix form is

$$\begin{pmatrix} ct' \\ x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} \gamma & 0 & 0 & -\gamma v/c \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -\gamma v/c & 0 & 0 & \gamma \end{pmatrix} \begin{pmatrix} ct \\ x \\ y \\ z \end{pmatrix} \tag{11}$$

Remember that this Lorentz transformation is a a coordinate transformation only between inertial frames moving in the z direction relative to each other. There are, of course, more "general" Lorentz transformation matrices in three dimensions. We will not consider these.

A crucial concept in physics is that of *invariance*. We speak often of quantities that are *invariant*–that is, remain the same–under a certain transformation. In this context, we consider *Lorentz invariance*, or the property of being left invariant under the Lorentz transformation. Mass, for example, is Lorentz-invariant. Another very important Lorentz-invariant is the following.

Theorem 3.3. *(Invariance of the interval) The spacetime interval is Lorentz invariant.*

This result follows from our example from which we derived the Lorentz transformation. The importance of this is that observers travelling in different inertial frames measuring two events in spacetime will not agree on the time and distance separating the events, but they will agree on the spacetime separation of events; that is, they will agree that

$$c^2\Delta t^2 - \Delta x^2 - \Delta y^2 - \Delta z^2 = c^2\Delta t'^2 - \Delta x'^2 - \Delta y'^2 - \Delta z'^2 \tag{12}$$

or $\Delta s^2 = \Delta s'^2$.

Theorem 3.4. *(Length contraction)* A moving object's length is shorter than its proper length, the length as measured in the object's own rest frame.

$$L' = \frac{L}{\gamma} \tag{13}$$

Theorem 3.5. *(Time dilation)* A moving clock is observed to run slower than a clock in the observer's rest frame. The time elapsed on the moving clock $\Delta t'$ is related to the time on the rest clock Δt by

$$\Delta t' = \gamma \Delta t \tag{14}$$

These are results that can be derived from the Lorentz transformation. Proofs of the theorems are left as an exercise.

Some terminology is worth noting. We observe things in the *laboratory frame*. The frame of reference that moves *with* a particle (that is, the frame in which the particle is stationary) is called its *rest frame*. For example, if you're sitting next to a shopping bag in a moving car, people outside will see you and the bag moving, but since you're in the particle's rest frame, the bag is stationary. As we've seen, each inertial frame has its own coordinate of time, which varies from observer to observer. The time measured by an observer using by their own clock is called *proper time*. The length of an object as measured in its rest frame is called *proper length*.

Exercise 3.6. *A muon is a subatomic particle that lives for (on average)* $\Delta\tau = 2.2\mu s$ *in its own rest frame before decomposing. If a muon is travelling with a speed* $v = 0.995c$ *relative to an observer on earth, what is its lifetime?*

Exercise 3.7. *Chad is born on the same day that his mom turns 40 years old. Chad stays on earth while his mom then goes on a trip far away, moving at* $v = 0.866c$ *for 40 years in her own reference frame, and stops at a new planet extremely fast. Ignore the process of stopping. When the mother has stopped at said planet, what is the new age difference between them? Assume the earth is at rest.*

Exercise 3.8. *Chad's mom measures her heart rate at 70 bpm. What will her heart rate be as measured by an observer on Earth?*

4 Mechanics

The relationships between various invariant physical quantities can be described using equations. if these equations have the same form in all inertial frames, they are called *form-invariant.*

Consider the velocity of a particle moving in four-dimensional spacetime. A clock in its rest frame will measure the particle's proper time τ.

Definition 4.1. *Four-velocity is the rate of change of a particle's position in spacetime with respect to proper time. It is a form-invariant vector defined as*

$$\vec{u} = \left(c\frac{dt}{d\tau}, \frac{dx}{d\tau}, \frac{dy}{d\tau}, \frac{dz}{d\tau} \right) \tag{15}$$

These types of vectors are generally referred to as *four-vectors.* Consider the position of this particle in spacetime, as measured in some arbitrary reference frame in which the particle is moving at a velocity v. We will denote the position by (ct, x, y, z). This is a four-vector which we will denote by (x^0, x^1, x^2, x^3), where $x^0 = ct$, $x^1 = x$, $x^2 = y$, and so on.[2] The definition of four-velocity then becomes

$$u = \left(\frac{\mathrm{d}x^0}{\mathrm{d}\tau}, \frac{\mathrm{d}x^1}{\mathrm{d}\tau}, \frac{\mathrm{d}x^2}{\mathrm{d}\tau}, \frac{\mathrm{d}x^3}{\mathrm{d}\tau} \right) \tag{16}$$

The μ'th component of u is $u^\mu = \frac{\mathrm{d}x^\mu}{\mathrm{d}\tau}$.

Using our result of time dilation, we write the time coordinate of our particle in terms of proper time (the time in the rest frame of the particle) as $x^0 = ct = c\gamma\tau$. When we take the derivative with respect to τ to obtain the four-velocity u, we just have $c\gamma$. As for the spatial components of u, we use the chain rule:

$$u^\mu = \frac{\mathrm{d}x^\mu}{\mathrm{d}\tau} = \frac{\mathrm{d}x^\mu}{\mathrm{d}x^0}\frac{\mathrm{d}x^0}{\mathrm{d}\tau} = \frac{\mathrm{d}x^\mu}{\mathrm{d}x^0}c\gamma = \frac{\mathrm{d}x^\mu}{\mathrm{d}(ct)}c\gamma = \frac{\mathrm{d}x^\mu}{c\mathrm{d}t}c\gamma = \frac{\mathrm{d}x^\mu}{\mathrm{d}t}\gamma \tag{17}$$

Since $\frac{\mathrm{d}x^\mu}{\mathrm{d}t}$ is just the particle's ordinary spatial velocity, which we express by a vector $\mathbf{v}$, we can reexpress the particle's four-velocity as

$$u = \gamma(c, \mathbf{v}) \tag{18}$$

Another quantity which we much define relativistically is momentum. We use proper time τ instead of coordinate time t.

Definition 4.2. *Relativistic momentum* $\mathbf{p}$ *is defined*

$$\mathbf{p} = m\left(\frac{dx}{d\tau}, \frac{dy}{d\tau}, \frac{dz}{d\tau} \right) \tag{19}$$

where m is rest mass.

In terms of coordinate time, we write

$$\mathbf{p} = m\gamma\left(\frac{\mathrm{d}x}{\mathrm{d}t}, \frac{\mathrm{d}y}{\mathrm{d}t}, \frac{\mathrm{d}z}{\mathrm{d}t} \right) = m\gamma\mathbf{v} \tag{20}$$

We might also be interested in how energy works in relativity. The derivations of the following (important) results are long and are left in the appendix.

[2]Note that we are dealing with *indices*, not exponents. While this notation is a possibly a bit confusing, it is standard in special relativity.

Definition 4.3. The relativistic kinetic energy of a moving object is

$$E_k = (\gamma - 1)mc^2 \tag{21}$$

Theorem 4.4. *(Mass–energy equivalence)* Energy and mass are related by the famous

$$E = \gamma mc^2 \tag{22}$$

Notice that $E = E_k + mc^2$. This mc^2 term is Einstein's important discovery: that a particle with mass m can be converted to pure energy, given by $E = mc^2$.

Let us introduce yet another quantity.

Definition 4.5. *Four-momentum is given by*

$$p = mu \tag{23}$$

where u is four-velocity.

Using $u = (c\gamma, \gamma\mathbf{v})$, we have that $p = (mc\gamma, m\gamma\mathbf{v})$; by mass-energy equivalence, this becomes

$$p = (E/c, \mathbf{p}) \tag{24}$$

Corollary 4.6. *(Energy-momentum relation)* The total energy of a moving object is

$$E = \sqrt{p^2c^2 + m^2c^4} \tag{25}$$

Theorem 4.7. *(Planck relation) The energy of a photon is*

$$E = hf \tag{26}$$

where h is Planck's constant.

Corollary 4.8. *(deBroglie relation)* The momentum of a photon is

$$p = \frac{h}{\lambda} \tag{27}$$

with h the Planck constant.

Exercise 4.9. *(Physics Bowl 2023 #50) An electron (e^-) is moving and has a kinetic energy of 1.00 MeV. It makes a head-on collision with a positron (e^+) initially at rest. In the collision the two particles annihilate each other and are replaced by two photons (γ) of equal energy, each traveling at angles θ to the electron's original direction of motion. The reaction is $e^- + e^+ \to 2\gamma$. The mass of an electron is equal to the mass of a positron, and is 0.511 MeV/c^2. Determine the angle of emission θ of each photon. (Hint: Relativistic momentum is conserved during the collision).*

Exercise 4.10. *Chad has a meter stick of mass 10 kg. Frustrated that his mom left him, he throws the stick in a direction parallel to its length, at a velocity v. The meter stick moves at a constant velocity, and in doing so has an energy of $30c^2$ Joules in Bob's reference frame.*

(i) What is v in terms of c?

(ii) How long does Chad observe the meter stick to be?

(iii) What is the meter stick's energy as observed in its own frame?

5 Appendix

For those that are curious, we will derive some of the relations from the mechanics section. We will make assumption of prerequisite knowledge, although everything you need to derive these is covered either in a high school physics class or has been covered in last year's physics club meetings. You do not need to know these for competitions.

5.1 Kinetic Energy

First, we derive the relativistic kinetic energy expression. The work done over a path $\mathcal{C}$ from a point x_1 to x_2 is[3]

$$W = \int_{\mathcal{C}} F dx \tag{28}$$

Of course, Newton's second law tells us $F = ma = m\frac{\mathrm{d}v}{\mathrm{d}t}$, so

$$\begin{aligned} W &= \int_{\mathcal{C}} m\frac{\mathrm{d}v}{\mathrm{d}t} dx \\ &= \int_{\mathcal{C}} m\frac{\mathrm{d}v}{\mathrm{d}x}\frac{\mathrm{d}x}{\mathrm{d}t} dx \\ &= \int_{\mathcal{C}} mv dv \\ &= \int_{\mathcal{C}} v dp \end{aligned} \tag{29}$$

Substituting our expression for relativistic momentum and integrating by parts,

$$\begin{aligned} W &= \int_{\mathcal{C}} v d\left(\frac{mv}{\sqrt{1-(v/c)^2}}\right) \\ &= \left[\frac{mv^2}{\sqrt{1-(v/c)^2}}\right]_{v_0}^{v_1} - \int_{\mathcal{C}} \frac{mv}{\sqrt{1-(v/c)^2}} dv \\ &= \left[\frac{mv^2}{\sqrt{1-(v/c)^2}} + mc^2\sqrt{1-(v/c)^2}\right]_{v_0}^{v_1} \\ &= \left[\frac{mv^2}{\sqrt{1-(v/c)^2}} + \frac{mc^2(1-(v/c)^2)}{\sqrt{1-(v/c)^2}}\right]_{v_0}^{v_1} \\ &= \left[\frac{mc^2}{\sqrt{1-(v/c)^2}}\right]_{v_0}^{v_1} \end{aligned} \tag{30}$$

[3]For a conservative force, this path doesn't matter, so you may have seen people simply integrate from x_1 to x_2.

The work is the change in kinetic energy. If we let $v_0 = 0$, intuitively our initial kinetic energy will be zero, and this expression will simply be our kinetic energy:

$$\begin{aligned} E_k &= \left[\frac{mc^2}{\sqrt{1-(v/c)^2}}\right]_0^{v_1} \\ &= \left(\frac{mc^2}{\sqrt{1-(v_1/c)^2}} - \frac{mc^2}{\sqrt{1-(0/c)^2}}\right) \quad = (\gamma - 1)mc^2 \end{aligned} \tag{31}$$

5.2 Mass-Energy Equivalence

If you use the previous expression of kinetic energy and define the Lagrangian as kinetic minus potential energy, you'll find that the equations of motion are wrong. We can derive it by d'Alembert's principle, which is that for a system of n particles,

$$\sum_i (F_i - \dot{p}_i)\delta x_i = 0 \tag{32}$$

where p_i is relativistic momentum of the i'th particle, δx_i is a virtual displacement, and a dot designates differentiation with respect to coordinate time. Focusing on the momentum term,

$$\begin{aligned} -\dot{p}\delta x &= -\frac{\mathrm{d}\gamma mv}{\mathrm{d}t}\delta x \\ &= -\frac{\mathrm{d}}{\mathrm{d}t}(\gamma mv\delta x) + \gamma mv\frac{\mathrm{d}(\delta x)}{\mathrm{d}t} \end{aligned} \tag{33}$$

The first term should vanish because the variation δx is 0 on the boundary. In the second term, ordiary and variational derivatives commute, so

$$-\dot{p}\delta x = \gamma mv\delta v \tag{34}$$

Notice that

$$\delta\left(\frac{1}{\gamma}\right) = -\frac{1}{c^2}\gamma v\delta v \tag{35}$$

which is a result you can derive from differentiating $\frac{1}{\gamma}$ with respect to v. This means that

$$\delta\left(-\frac{mc^2}{\gamma}\right) = \gamma mv\delta v \tag{36}$$

which is precisely our result for $-\dot{p}\delta x$. We can integrate the d'Alembert principle with respect to time and rearrange terms to find

$$\int \sum_i (F_i - \dot{p}_i)\delta x_i = \delta \int \left(-\sum_i \frac{m_i c^2}{\gamma_i} - V\right) dt = 0 \tag{37}$$

The right hand side looks just like the principle of least action

$$\delta S = \delta \int L dt = 0 \tag{38}$$

So we find that

$$L = -\frac{mc^2}{\gamma} - V \tag{39}$$

Consider a particle that is not under the influence of any potential energy ($V = 0$). You can find that

$$p = \frac{\partial L}{\partial v} = mv\gamma \tag{40}$$

This aligns with our definition of relativistic momentum; furthermore, by the relation between the Lagrangian and Hamiltonian,

$$E = pv - L = \gamma m \left(v^2 + \frac{c^2}{\gamma^2} \right) = \gamma mc^2 \tag{41}$$

When $v = 0$, our $\gamma = 1$; the Lagrangian reduces to $L = -mc^2$, and our Hamiltonian/energy becomes mc^2. That is, when the particle is at rest, it still has an energy $E = mc^2$ which is endowed by its mass.

5.3 Energy-Momentum Relation

To derive this expression, one must consider the *metric.* Consider two points in three-dimensional space. The distance between them, by the Pythagorean theorem, is

$$\Delta\ell^2 = \Delta x^2 + \Delta y^2 + \Delta z^2 \tag{42}$$

If we take an infinitely small step in each coordinate direction (dx, dy, dz), we find that the *line element* $d\ell$ can be found using

$$d\ell^2 = dx^2 + dy^2 + dz^2 \tag{43}$$

This allows us to define a *metric tensor.* In this case, we have a 3 x 3 matrix

$$[g_{ij}] = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{44}$$

We call the metric for Euclidean space the *Euclidean metric.* The line element is given by

$$d\ell^2 = \sum_{i=1}^{3} \sum_{j=1}^{3} g_{ij} dx^i dx^j \tag{45}$$

where g_{ij} is the element of the i'th row and j'th column of $[g_{ij}]$, and dx^i is the i'th coordinate. Having multiple Σ's can be a bit cumbersome at times, so we employ the Einstein summation convention, named after the legend himself; whenever we have a repeated index, we sum over that index. So we rewrite our last equation as

$$d\ell^2 = g_{ij} dx^i dx^j \tag{46}$$

This takes a little getting used to, but it's widely used in physics and mathematics. Anyways, notice that we can define an inner product of position vectors r_1 and r_2 by

$$\langle r_1, r_2 \rangle = g_{ij} r_1^i r_2^j \tag{47}$$

This fact will become important later on. In the meantime, recall that we defined an interval in spacetime to be given by

$$\Delta s^2 = c^2 \Delta t^2 - \Delta x^2 - \Delta y^2 - \Delta z^2 \tag{48}$$

The line element for this becomes

$$ds^2 = c^2dt^2 - dx^2 - dy^2 - dz^2 \tag{49}$$

A metric $[\eta_{ij}]$ must satisfy

$$ds^2 = \eta_{ij}dx^i dj^i \tag{50}$$

The metric tensor for spacetime is given by the 4 x 4 matrix

$$[\eta_{ij}] = \begin{pmatrix} c & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \tag{51}$$

This metric for spacetime is called the *Minkowski metric.*[4] The dot product of four-velocity becomes

$$\begin{aligned} \langle u, u \rangle = \eta_{ij}u^i, U^j &= \gamma^2c^2 - \gamma^2\langle \mathbf{v}, \mathbf{v} \rangle \\ &= \gamma^2(c^2 - v^2) \qquad = c^2 \end{aligned} \tag{52}$$

So for four momentum,

$$\eta_{ij}p^ip^j = m^2\eta_{ij}u^iu^j = m^2c^2 \tag{53}$$

because $p = mu$; but if we used $p = (E/c, \mathbf{p})$ to compute the dot product, we'd find that

$$\eta_{ij}p^ip^j = \frac{E^2}{c^2} - \langle \mathbf{p}, \mathbf{p} \rangle = \frac{E^2}{c^2} - p^2 \tag{54}$$

Therefore,

$$\frac{E^2}{c^2} - p^2 = m^2c^2 \tag{55}$$

or

$$E^2 = p^2c^2 + m^2c^4 \tag{56}$$

From algebra class, you may notice that taking the square root on both sides yields *two* solutions, wherein one has *negative energy.* This caused a bit of a conundrum for P. A. M. Dirac, which led to great discoveries in particle physics.

[4]Often times, you will find people set $c = 1$ for convenience. We will not use this convention here so as to prevent confusion.