Special Relativity

William F. Barnes

March 8, 2021

Contents

Introduction

Before the rise of Albert Einstein's popularity in 1905, physicists squarely explained natural phenomena using Newtonian mechanics, thermodynamics, and electrodynamics. There was no reason to suspect, for instance, that space itself might influence matter rather than simply hold it, or that the concept of time becomes slippery when thoroughly examined. Despite the intuitiveness of Newtonian physics that propelled us to the eighteenth century, we find that our notions of space and time are not only wrong, but must be corrected by a spectacular theory called *special relativity.*

1 Pre-Einstein Physics

1.1 Light as a Wave

A new understanding of space and time begins where electrodynamics leaves off. It was discovered that all electric and magnetic phenomenon can be contained Maxwell's four equations (using vector calculus notation):

$$\vec{\nabla} \cdot \vec{E} = \frac{\rho}{\epsilon_0} \qquad \vec{\nabla} \times \vec{E} = -\frac{\partial \vec{B}}{\partial t}$$

$$\vec{\nabla} \cdot \vec{B} = 0 \qquad \vec{\nabla} \times \vec{B} = \mu_0 \vec{J} + \mu_0 \epsilon_0 \frac{\partial \vec{E}}{\partial t}$$

In a region without electric charges, Maxwell's equations can be combined to yield two simultaneous relations:

$$\epsilon_0 \mu_0 \frac{\partial^2 \vec{E}}{\partial t^2} = \nabla^2 \vec{E} \qquad \epsilon_0 \mu_0 \frac{\partial^2 \vec{B}}{\partial t^2} = \nabla^2 \vec{B}$$

These were immediately recognized as wave equations, having sinusoidal solutions. Moreover, the propagation speed of such 'electromagnetic' waves can be interpreted as

$$c = \frac{1}{\sqrt{\epsilon_0 \mu_0}} \,.$$

Historical measurements of ϵ_0 (electric permittivity of free space) and μ_0 (magnetic permeability of free space) indicated that $c \approx 3 \times 10^8 m/s$, which was suspiciously in agreement with known measurements of the speed of visible light. Compelled by this connection, Maxwell himself asserted:

'We have strong reason to conclude that light itself - including radiant heat and other radiation, if any - is an electromagnetic disturbance in the form of waves propagated through the electromagnetic field according to electromagnetic laws.'

1.2 Aether Hypothesis

To understand light in terms of electromagnetic waves, speculation quickly arose over the medium of propagation. Light, having a faster speed than any known wave, should only move through something immensely rigid. Paradoxically though, such a medium needs to be thin enough to not be 'noticed' by ordinary matter. This ghostly inhabitant of all visible space was deemed the *Luminiferous Aether*.

Physicists reasoned the aether must be at rest with respect to *some* frame of reference, and surely Earth constantly moves with respect to such a universal rest frame. It should follow that light rays deviate away from straight lines by an 'aether wind' when observed from any Earth-based laboratory. Using measurements of the off-target deflection, Earth's absolute motion through the aether could be deduced.

1.3 Michelson-Morley Experiment

An experiment to detect the presence of ether wind was devised in 1880-1881 by German physicist A. A. Michelson, and refined in partnership with E. W. Morley by 1887 in the United States. They devised an experiment in where a light beam was split into two perpendicular components, with each ray reflected back and recombined into a pair of interfering beams. With the so-called Michelson-Morley apparatus, Earth's motion through the aether would be detected by measuring moving 'fringes' in the recombined beam. To eliminate any bias from the physical alignment of the apparatus, it was designed as free-floating on a bed of mercury, able to rotate around a center axis.

Conducting their experiment, Michelson and Morley were shocked to find that the fringe pattern did *not* change when varying *any* experimental parameters. Time of day, duration of the experiment, orientation of the apparatus - none of these changed the result. In other words, no aether wind was measured, and the aether hypothesis was in trouble.

1.4 Magnet and Ring

It's unclear whether Einstein knew of the Michelson-Morley experiment; he was motivated purely by electrodynamics. As a direct example of Faraday's law, it's well known that a permanent magnet in motion causes an electric field to arise around the magnet. Such a field can be captured by a ring-shaped conductor, oriented perpendicular to the magnet's motion, resulting an induced current in the ring. By symmetry, the same effect arises when the magnet is instead held at rest while ring moves along the axis of the magnet.

Einstein was the first to ponder the notion of a permanent magnet 'at rest', or a loop at rest for that matter, is not an absolute idea at all. For example, a magnet that lies on the floor of a uniformly-coasting train car does not 'know' it's moving, and does not induce a local electric field. A ring traveling along with the magnet will have no induced current. When viewed from a non-moving frame of reference however, such as a train station, the magnet *does* appear to induce a field around it. (You could test this by making a large-enough loop for the whole train to pass through.) It should be impossible for the magnet to simultaneously emit an electric field and to not emit an electric field - unless we discard the notion of 'absolute' motion altogether and accept that relative motion is all that matters.

This insight allowed Einstein to discard the aether hypothesis. If the stationary aether exists, then all magnets, even those 'at rest', should have an induced electric field caused by aether wind. This is experimentally not the case - instead only the relative motion between the magnet and the loop is significant.

1.5 Emission Theories of Light

Several attempts were made to modify the equations of electrodynamics (by altering the so-called retardation time) in order to correct the problems pointed out by Einstein. In the *Maxwell emission theory* of light, all electromagnetic radiation propagates at speed c from the *fixed point in aether* from which it was emitted. In a modified version named after Ritz, all electromagnetic radiation propagates at speed c with respect to a *moving* point that follows the instantaneous motion of the source at emission time.

After plenty of effort to defend any version of emission theory by Einstein and others, it was decided that such a theory isn't strong enough to fix electrodynamics, and all versions of emission theory were abandoned.

1.6 Enter Vacuum

With the aether hypotheses sufficiently deflated, physicists were left with a set of wave equations, but *no* medium to support the waves. The idea that a medium wasn't needed eventually gained traction, so the preferred medium for light propagation became vacuum. Further studies of optics showed that materials are in fact a hindrance to light wave propagation.

2 Frame of Reference

A *reference frame* (or simply 'frame') is an observer or set of observers embedded in one coordinate system that can be used to track position and motion. Common examples of reference frames may be a laboratory, a moving car, or the planet Jupiter. In any frame, events and trajectories are expressed in terms of displacement $\vec{x}$ and time t from a chosen 'origin' $\vec{x} = 0,\ t = 0$.

2.1 Inertial Frame

An *inertial* reference frame is one that 'feels' no unjustified accelerations. With no gravity or external fields present, any constant-velocity (including rest) frame is considered inertial. When forces are present, only the 'freefall' frame is inertial.

When inside an inertial reference frame, nobody can 'feel' whether the frame is accelerated or moving. Special relativity (what we're doing in this study) only applies to inertial reference frames, whereas general relativity can handle accelerated frames.

Galilean Transformation

Consider an inertial reference frame at rest with respect to an origin ($\vec{x} = 0, t = 0$). Also consider a second reference frame that moves with speed v with respect to the first frame. This 'primed' frame has coordinates represented with a prime symbol (a single quote), namely $(\vec{x})'$, t'.

Before 1905, it was natural for physicists to reconcile different events as observed from different reference frames by writing down the *Galilean transformation*, namely

$$(\vec{x})' = \vec{x} - \vec{v}t \qquad\qquad \vec{x} = (\vec{x})' + \vec{v}t\ .$$

All of classical physics encloses or assumes the Galilean transformation. Even certain proofs, such as momentum conservation, directly rely on the Galilean transformation being correct.

2.2 Catching Up to a Light Wave

It's nearly unfathomable to convince the world that the elementary equations of motion are somehow wrong, but Einstein pulled it off (with resistance). Begin by considering a light source, such as a laser, light bulb, or car headlight, at rest in a laboratory. Rays emitted from the light source, which take off at speed c, are made of oscillating E- and B-fields oriented mutually perpendicular to the propagation direction.

Next consider a train, car, or other reference frame that can move fast enough to keep pace with the speed of light. With the laboratory light source still on, what happens if we enter the car, accelerate to speed c, and observe the light rays in their 'own' reference frame?

Since light rays are made of oscillating E- and B-fields, we expect to measure the waves to 'slow down' as the car approaches speed c, and finally become stationary at exactly c. Einstein realized this 'frozen light wave' is clearly nonsense, as Maxwell's equations do not predict stationary oscillatory waves. Moreover, experiment has never turned up such a result.

Even worse, there is no evidence in theory or experiment of a light ray slowing down at all (in vacuum). The *only* observation that fits into existing electrodynamics is for the speed of light to *never* deviate from c (in vacuum), regardless of the frame in which it is measured.

Thought Experiment

To reiterate the puzzle on hand, consider three students named Alice, Bob, and Cameron. Alice stands with a light source at the beginning of a straight road, where Bob stands at the end of the road with a light detector. Making careful measurements, (A)lice and (B)ob can measure the speed of light by measuring the distance between A and B, and dividing by the travel time of the ray. Their result, of course, will come out close enough to $c \approx 3.00 \times 10^8 m/s$.

Next, suppose Cameron is equipped with a car with a similar light source mounted on the front. When the car is moving at speed v, (C)ameron activates her light source and carefully coordinates with Bob to measure the speed of light, namely the instantaneous distance CB divided by the time interval. As the Galilean transformation suggests, the students expect to measure a relative speed $c \pm v$ as an effective speed of light. Amazingly though, Bob's distance-over-time calculation still results to c.

If we instead give Bob a light source, and then mount light detectors on the front and back of the car, the speed of light can be simultaneously calculated by both Alice and Bob, using the distances AC and CB respectively. Conducting this experiment, it turns out that both measurements somehow result in c, with no correction for the velocity of the car.

Supposing finally that Cameron has an independent way to measure the speed of light from her moving car, her results will be flatly c. No matter what that car speed is, any light rays emitting from it fly off at speed c.

2.3 The Case for Special Relativity

Basic considerations from electrodynamics hint that the notion of any 'absolute' reference frame, whether it be the aether, some point in a laboratory, or a special moving point on a train - is clearly troublesome. By trying to 'catch up' to a light ray, the Galilean transformation with $v \to c$ implies a 'frozen wave', which is not predicted by Maxwell's equations. More strangely, measurements of the speed of light always come out to c, regardless of the motions of the observer and the source.

All of these observations don't sit well together, which places the integrity of classical physics at risk. Einstein knew that *something* must give way to the stubborn behavior of light. Eventually he realized that 'absolute' notions of time and space, just like the luminiferous aether, was mere assumptions to be discarded.

3 Einstein's Postulates

After years of puzzling through thought experiments, numerous attempts to modify electromagnetism, and the pursuit of many tangents, Einstein successfully distilled the whole story of special relativity into two fundamental postulates:

1. The laws of physics are equivalent in all inertial reference frames.
2. The speed of light is a universal constant in all reference frames.

The way to 'correctly' proceed is to *only* assume Einstein's postulates, and leave *everything* else up for reinterpretation - this includes the notions of time, space, even mass.

3.1 Time Dilation Effect

Imagine a 'light-clock' that is made of two inward-facing parallel mirrors, which trap a pulse of light endlessly reflecting between the mirrors at perpendicular incidence. If the separation between mirrors is L, than an observer holding the clock counts one 'tick' every c/L seconds. Defining the time between ticks as Δt, we know the light clock obeys

$$L = c\Delta t$$

at rest.

Next, consider an observer who moves with constant velocity v in a direction perpendicular to the pulse's motion. According to a moving observer, there are two perpendicular components to the pulse's displacement in one 'tick',

$$\sqrt{L^2 + (v\Delta t')^2} = c\Delta t' \; .$$

In the above, we have not assumed Δt and $\Delta t'$ are the same, hence the prime symbol. The c-variable needs no prime symbol by Einstein's second postulate. Since the motion is parallel to the set of mirrors, the same length L appears in each equation.

Eliminating L between the two above equations will relate Δt to $\Delta t'$, namely

$$\Delta t' = \frac{1}{\sqrt{1 - v^2/c^2}} \Delta t \; ,$$

which is our first evidence of the *time dilation* effect. Evidently, a moving reference frame's sense of time 'slows down' as a function of velocity with respect to a stationary frame. The rest-frame time Δt is called the *proper time*, often denoted $\Delta \tau$, always larger than $\Delta t'$.

3.2 Length Contraction Effect

The mechanism behind time dilation has a similar effect on space. Consider a reference frame in which a clock moves at constant velocity v with respect to a stationary rod of length L. The time required for the clock to pass from between endpoints of the rod is represented by Δt in the stationary frame, and $\Delta t'$ in the moving frame.

Without assuming that lengths are consistent between reference frames, denote the *observed* length of the rod, as seen from the primed frame, as L'. Two distance equations immediately follow:

$$L = v\Delta t \qquad\qquad L' = v\Delta t'$$

The rest-frame length L is called the *proper length*, often denoted L_p. The proper time of the clock is in the primed frame, so replace $\Delta t' \to \Delta\tau$ to write

$$L_p = v\Delta t \qquad\qquad L' = v\Delta\tau \; .$$

Taking the ratio of the equations, and then substituting the time dilation formula, we end up with

$$L' = L_p\sqrt{1 - v^2/c^2} \; ,$$

known as the *length contraction* effect. It follows that L' is always less than L_p, which means objects must appear to shorten along their direction of motion.

Atmospheric Muons

Cosmic rays that interact with Earth's atmosphere have a chance of generating muon particles (somewhat like heavy electrons). Being unstable particles, muons decay in about 2.2 microseconds. Interestingly, muons generated in the high atmosphere, approximately $6000m$, can still be detected at ground level. Using Newtonian mechanics to calculate the downward speed of a detected muon, the result is at least

$$v_{\text{Newton}} = \frac{6.0 \times 10^3 \; m}{2.2 \times 10^{-6} \; s} \; ,$$

exceeding the speed of light. Moreover, muons are experimentally measured to typically move at $99\%c$, but never faster.

The time dilation effect rescues the picture. If the ratio v/c is great enough, the observed muon lifetime is stretched out by a factor of $1/\sqrt{1 - v^2/c^2}$. In the reference frame of the muon, the lifetime is still 2.2 microseconds. Due to the length contraction effect, the muon 'observes' the ground to be much closer than it would appear if at rest (by a factor of $\sqrt{1 - v^2/c^2}$).

4 Spacetime

As a direct consequence of Einstein's postulates, we simultaneously discover the time dilation effect and the length contraction effect, both depending on the quantity $\sqrt{1 - v^2/c^2}$. In a single stroke, Einstein's postulates have implied that time and space are mutually flexible, which taken together form a four-dimensional manifold known as *spacetime*.

4.1 Gamma Factor

This quantity $\sqrt{1 - v^2/c^2}$ is so common, its reciprocal has a special name called the *gamma factor*:

$$\gamma\left(v\right) = \frac{1}{\sqrt{1 - v^2/c^2}}$$

The gamma factor becomes infinite as $v \rightarrow c$, and limits to 1 when $v = 0$. When the ratio v/c is much less than one, $\gamma\left(v\right)$ can be expanded:

$$\gamma = \frac{1}{\sqrt{1 - v^2/c^2}} \approx 1 + \frac{v^2}{2c^2} + \frac{3v^4}{8c^4} + \frac{5v^6}{16c^6} + \cdots$$

Replacing $\sqrt{1 - v^2/c^2}$ with $\gamma\left(v\right)$, the time dilation and length contraction equations simplify to

$$\Delta t' = \gamma \Delta \tau \qquad\qquad L' = \frac{L_p}{\gamma} \,,$$

where $\Delta\tau$ is the proper time and L_p is the proper length of an object at rest. Primed variables are measured in a moving frame at speed v.

4.2 Lorentz Transformation

The success of Newtonian mechanics in so many scenarios validates the Galilean transformation $x' = x - vt$, however such a form is blind to the effects of time dilation and length contraction. To proceed we must apply a correction involving γ that captures each, namely

$$x' = \gamma\left(x - vt\right) \qquad\qquad x = \gamma\left(x' + vt'\right) \,,$$

known as the *Lorentz transformation.* Using only the Lorentz transformation as a starting point, we may re-derive the gamma factor, and also cook up general versions of the time dilation and length contraction relations. We'll do all calculations in one dimension for starters.

To begin, consider two reference frames that overlap at $x = x' = 0$ and $t = t' = 0$, where the primed frame moves at constant velocity v with respect to a stationary frame. At $t = t' = 0$, a pulse of light emerges from $x = x' = 0$ and moves along the $+x$ direction. In each respective frame, the position x of the pulse is given by

$$x = ct \qquad\qquad x' = ct' \,.$$

Substituting each of these into the pair of Lorentz transformation equations, we get

$$ct' = \gamma\,(ct - vt) \qquad\qquad ct = \gamma\,(ct' + vt')\ ,$$

where eliminating the ratio t'/t yields the anticipated $\gamma = 1/\sqrt{1 - v^2/c^2}$.

The next task is to pursue a time dilation equation by solving for t' in terms of all unprimed variables (and similarly for t with primed variables). Using the identity

$$1 - \gamma^2 = -\frac{\gamma^2 v^2}{c^2}\ ,$$

the two Lorentz transformation equations yield simultaneous equations for time dilation,

$$t' = \gamma\left(t - \frac{vx}{c^2}\right) \qquad\qquad t = \gamma\left(t' + \frac{vx'}{c^2}\right)\ .$$

Amazingly, the equations that keep track of time have a spatial component.

Finally, we recover length contraction by considering a stationary rod of length L as observed from a moving reference frame of velocity v with respect to the rod. From the moving reference frame, the rod's length must be measured using *simultaneous* observations of its endpoints. Supposing both measurements occur a the same time $t'_1 = t'_2$, and the rod's endpoints are observed at x'_1 and x'_2, respectively, the time dilation equation gives

$$\begin{aligned}\Delta t = t_2 - t_1 &= \gamma\left(\cancel{t'_2 - t'_1} + \frac{v}{c^2}\,(x'_2 - x'_1)\right)\\ \Delta t &= \gamma\frac{v}{c^2}\Delta x'\ .\end{aligned}$$

Meanwhile, the Lorentz transformation tells us $\Delta x' = \gamma\,(\Delta x - v\Delta t)$. Eliminating Δt and solving for $\Delta x' = L'$, we find

$$\Delta x' = \frac{1}{\gamma}\Delta x \qquad\qquad \rightarrow \qquad\qquad L' = \frac{L_p}{\gamma}\ .$$

4.3 Lorentz Boost

To a stationary observer, any reference frame moving with constant velocity is said to have a *Lorentz boost* in the direction of motion. For a single boost along the x-direction, the full Lorentz transformation reads:

$$\begin{aligned} ct' &= \gamma\left(ct - \frac{v}{c}x\right)\\ x' &= \gamma\,(x - ct)\\ y' &= y\\ z' &= z \end{aligned}$$

To keep the notation tight, define the ratio v/c as β such that $0 \leq \beta < 1$. In a matrix form, a Lorentz boost along the x-direction looks like:

$$\begin{bmatrix} ct' \\ x' \\ y' \\ z' \end{bmatrix} = \begin{bmatrix} \gamma & -\gamma\beta & 0 & 0 \\ -\gamma\beta & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} ct \\ x \\ y \\ z \end{bmatrix}$$

Similar matrices exist for isolated boosts in the y- and z- directions.

3D Lorentz Boost (Advanced)

To handle a general boost in mixed directions, we begin by noting that any position vector can be resolved into parallel-to-v and perpendicular-to-v components, namely

$$\vec{r} = \vec{r}_{\parallel v} + \vec{r}_{\perp v} \; .$$

Next, note that only the parallel-to-v component of the motion is subject to the Lorentz boost. The one-dimensional transformation and time dilation equations generalize to:

$$x' = \gamma\left(x - vt\right) \quad \rightarrow \quad (\vec{r})' = \vec{r}_{\perp v} + \gamma\left(\vec{r}_{\parallel v} - \vec{v}t\right)$$

$$t' = \gamma\left(t - \frac{vx}{c^2}\right) \quad \rightarrow \quad t' = \gamma\left(t - \frac{\vec{r}\cdot\vec{v}}{c^2}\right)$$

Replace all instances of $\vec{r}_{\parallel v}$ and $\vec{r}_{\perp v}$ using

$$\vec{r}_{\parallel v} = \frac{\vec{r}\cdot\vec{v}}{v} \qquad \text{and} \qquad \vec{r}_{\perp v} = \vec{r} - \vec{r}_{\parallel v}$$

to arrive at

$$ct' = \gamma\left(ct - \vec{r}\cdot\vec{\beta}\right) \qquad (\vec{r})' = \vec{r} + \left(\frac{\vec{r}\cdot\vec{v}}{v^2}\left(\gamma - 1\right) - \gamma t\right)\vec{v} \, ,$$

where $\vec{\beta} = \vec{v}/c$. In block matrix form, these are

$$\begin{bmatrix} ct' \\ \vec{r}' \end{bmatrix} = \begin{bmatrix} \gamma & -\gamma\vec{\beta}^T \\ -\gamma\vec{\beta} & I + \left(\gamma - 1\right)\vec{\beta}\vec{\beta}^T/\beta^2 \end{bmatrix} \begin{bmatrix} ct \\ \vec{r} \end{bmatrix} ,$$

where $\vec{\beta}^T$ is the transpose of $\vec{\beta}$ and I is the identity matrix. In terms of x, y, z components, we finally have

$$\begin{bmatrix} ct' \\ x' \\ y' \\ z' \end{bmatrix} = \begin{bmatrix} \gamma & -\gamma\beta_x & -\gamma\beta_y & -\gamma\beta_z \\ -\gamma\beta_x & 1 + \left(\gamma - 1\right)\beta_x^2/\beta^2 & \left(\gamma - 1\right)\beta_x\beta_y/\beta^2 & \left(\gamma - 1\right)\beta_x\beta_z/\beta^2 \\ -\gamma\beta_y & \left(\gamma - 1\right)\beta_y\beta_x/\beta^2 & 1 + \left(\gamma - 1\right)\beta_y^2/\beta^2 & \left(\gamma - 1\right)\beta_y\beta_z/\beta^2 \\ -\gamma\beta_z & \left(\gamma - 1\right)\beta_z\beta_x/\beta^2 & \left(\gamma - 1\right)\beta_z\beta_y/\beta^2 & 1 + \left(\gamma - 1\right)\beta_z^2/\beta^2 \end{bmatrix} \begin{bmatrix} ct \\ x \\ y \\ z \end{bmatrix} .$$

4.4 Simultaneous Events

An *event* is a single location in spacetime (x, t) as measured in a given reference frame. A pair of events 1 and 2 generate two copies of the Lorentz transformation

$$x_1' = \gamma\left(x_1 - vt_1\right) \qquad x_1 = \gamma\left(x_1' + vt_1'\right)$$

$$x_2' = \gamma\left(x_2 - vt_2\right) \qquad x_2 = \gamma\left(x_2' + vt_2'\right) \; ,$$

and similarly for the time dilation equations,

$$t_1' = \gamma\left(t_1 - vx_1/c^2\right) \qquad t_1 = \gamma\left(t_1' + vx_1'/c^2\right)$$

$$t_2' = \gamma\left(t_2 - vx_2/c^2\right) \qquad t_2 = \gamma\left(t_2' + vx_2'/c^2\right) \; .$$

Subtracting the like pairs, we arrive at the equations of simultaneity:

$$\Delta x' = \gamma\left(\Delta x - v\Delta t\right) \qquad \Delta x = \gamma\left(\Delta x' + v\Delta t'\right)$$

$$\Delta t' = \gamma\left(\Delta t - \frac{v\Delta x}{c^2}\right) \qquad \Delta t = \gamma\left(\Delta t' + \frac{v\Delta x'}{c^2}\right)$$

Simultaneity in Time

In light of the above relations, the notion of 'simultaneous' events between boosted reference frames becomes ambiguous. Consider an observer in a stationary frame where two light sources separated by L simultaneously emit a pulse simultaneously at $t = 0$. The two pulses comprise a pair of events with $\Delta x = L$ and $\Delta t = 0$. A boosted observer will find that $\Delta t'$ reduces to

$$\Delta t' = -\gamma \frac{v}{c^2} L \, ,$$

implying the pulses are not simultaneous when observed in a moving frame.

Simultaneity in Space

For another example, consider two events in a stationary frame that occur at the same location but are separated in time by τ. In a boosted frame, the events do not occur in the same location, but instead obey

$$\Delta x' = -\gamma v \tau \, .$$

Notice that for $\gamma \approx 1$, the Newtonian model $\Delta t' \to 0$ and $\Delta x' \to -v\tau$ emerges.

4.5 Addition of Velocities

Consider a stationary frame along with a boosted frame with velocity v along the x-direction. A particle moves along x in both frames has speeds u_x and u'_x, respectively. According to Galileo, the speeds would relate by $u_x = u'_x + v$, however the Lorentz transformation tells us

$$u_x = \frac{\Delta x}{\Delta t} = \frac{\gamma \left(\Delta x' + v \Delta t'\right)}{\gamma \left(\Delta t' + v \Delta x'/c^2\right)} = \frac{u'_x + v}{1 + v u'_x / c^2} \qquad\qquad u'_x = \frac{u_x - v}{1 - v u_x / c^2} \, .$$

Despite the boost being along x, the particle is still subject to time dilation if its motion is along the y-direction

$$u_y = \frac{\Delta y'}{\Delta t} = \frac{\Delta y'}{\gamma \left(\Delta t' + v \Delta x'/c^2\right)} = \frac{u'_y}{\gamma \left(1 + v u'_x / c^2\right)} \qquad\qquad u'_y = \frac{u_y}{\gamma \left(1 - v u_x / c^2\right)} \, ,$$

and similarly for z.

Rapidity

Since velocities don't stack linearly via $u_x = u'_x + v$, it's natural to wonder just *which* quantity does, which we tackle by writing

$$\Phi = \phi' + \phi$$

such that

$$u_x = c f\left(\phi' + \phi\right) \qquad\qquad u'_x = c f\left(\phi'\right) \qquad\qquad v = c f\left(\phi\right) \, ,$$

where ϕ is called the *rapidity*. In terms of rapidity, the addition-of-velocities formula becomes

$$f\left(\phi' + \phi\right) = \frac{f\left(\phi'\right) + f\left(\phi\right)}{1 + f\left(\phi'\right) f\left(\phi\right)} \, .$$

By observation (luck, actually) notice that one suitable function that stacks the same way as $f\left(\right)$ is the hyperbolic tangent, specifically

$$\tanh\left(\phi' + \phi\right) = \frac{\tanh\left(\phi'\right) + \tanh\left(\phi\right)}{1 + \tanh\left(\phi'\right)\tanh\left(\phi\right)} \, ,$$

so we finally interpret

$$u_x = c\tanh\left(\phi' + \phi\right) \qquad u'_x = c\tanh\left(\phi'\right) \qquad v = c\tanh\left(\phi\right) \, .$$

Starting with $v = c\tanh\left(\phi\right)$, it's straightforward to show that

$$\cosh\phi = \gamma \qquad \sinh\phi = \frac{\gamma v}{c} \, .$$

In the same notation, the Lorentz transformation for a single boost in the x direction appears as

$$\begin{bmatrix} ct' \\ x' \\ y' \\ z' \end{bmatrix} = \begin{bmatrix} \cosh\phi & -\sinh\phi & 0 & 0 \\ -\sinh\phi & \cosh\phi & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} ct \\ x \\ y \\ z \end{bmatrix} \, .$$

To handle two Lorentz boosts, simply replace $\phi \rightarrow \Phi = \phi' + \phi$. The addition-of-velocities equation is contained in the ϕ-notation.

4.6 Spacetime Interval

Observers in different inertial reference frames don't generally agree on the coordinates, duration, or simultaneity of events. It's natural to wonder if such observers agree on *anything* at all. As it turns out, there do exist quantities, called *invariants*, that are the same in all reference frames.

Starting with the simultaneity relations for Δx and Δt, square each and subtract off the Δt-equation:

$$\begin{aligned} \Delta x^2 &= \gamma^2\left(\Delta x'\right)^2 + \gamma^2 v^2\left(\Delta t'\right)^2 + 2\gamma^2 v\left(\Delta t'\right)\left(\Delta x'\right) \\ c^2\Delta t^2 &= \gamma^2 c^2\left(\Delta t'\right)^2 + \frac{\gamma^2 v^2}{c^2}\left(\Delta x'\right)^2 + 2\gamma^2 v\left(\Delta t'\right)\left(\Delta x'\right) \\ \Delta x^2 - \left(c\Delta t\right)^2 &= \left(\Delta x'\right)^2\gamma^2\left(1 - \frac{v^2}{c^2}\right) + \left(c\Delta t'\right)^2\gamma^2\left(\frac{v^2}{c^2} - 1\right) + 0 \\ \Delta x^2 - \left(c\Delta t\right)^2 &= \left(\Delta x'\right)^2 - \left(c\Delta t'\right)^2 \end{aligned}$$

Evidently, the form $\Delta x^2 - c^2\Delta t^2$ is the same in both the primed and un-primed frame, and is therefore invariant. This quantity is called the *spacetime interval*, denoted ΔS^2 such that

$$\Delta S^2 = -c^2\Delta t^2 + \Delta\vec{x}\cdot\Delta\vec{x} \, ,$$

where the variable x generalizes to the three-dimensional position vector $\vec{x}$. In the limit that $\Delta x = 0$, the spacetime interval readily reduces to $-c^2\Delta\tau^2$, but this is in fact true for any Δx:

$$\Delta S^2 = -c^2\Delta t^2 + \Delta x^2 = -c^2\Delta t^2\left(1 - \frac{1}{c^2}\frac{\Delta x^2}{\Delta t^2}\right) = \Delta S^2 = -c^2\frac{\Delta t^2}{\gamma^2}$$

$$\Delta S^2 = -c^2\Delta\tau^2$$

4.7 Causality

The spacetime interval

$$\Delta S^2 = -c^2 \Delta t^2 + \Delta x^2 = -c^2 \Delta \tau^2$$

is defined for any pair of events (x_1, t_1), (x_2, t_2), and the overall sign of ΔS^2 indicates the causal relation between them.

Events that are sufficiently separated by distance such that the first event could not possibly have caused the second event, where $\Delta x^2 > c^2 \Delta t^2$, are called *spacelike* events:

$$\Delta S^2 > 0 \qquad\qquad \text{Spacelike}$$

If the pair of events instead obey $\Delta x^2 < c^2 \Delta t^2$, the earlier event has a potential influence on the later event, and the pair is called *timelike*:

$$\Delta S^2 < 0 \qquad\qquad \text{Timelike}$$

Finally, the special case $\Delta x^2 = c^2 \Delta t^2$ is only satisfied by $\Delta x / \Delta t = c$, meaning the events can only be connected by a light ray extending from the first event, arriving at the location of the second event as it occurs. These are *lightlike* events:

$$\Delta S^2 = 0 \qquad\qquad \text{Lightlike}$$

As a corollary to the fact that $\Delta S^2 = -c^2 \Delta \tau^2 = 0$ for lightlike events, it follows that the rest frame of a light pulse experiences no passage of time. The proper time is always zero in the frame of light rays.

4.8 Spacetime Diagrams

The Lorentz transformation and time dilation equations, namely

$$x' = \gamma (x - vt) \qquad\qquad t' = \gamma \left(t - \frac{vx}{c^2} \right) ,$$

can be visualized by plotting both the primed and unprimed coordinates on the same graph, called a *spacetime diagram* or *Minkowski diagram*. For simplicity we assume the origin of each system coincides at $(0, 0)$.

Axes

Assigning x and ct to the horizontal and vertical axes (as one would do with x and y), we begin by finding equations for the axes x' and ct'. Setting $t' = 0$ to find the x'-axis and vice versa, we write

$$ct = \frac{v}{c} x \qquad\qquad x' \text{ axis}$$

$$x = \frac{v}{c} ct \qquad\qquad ct' \text{ axis} .$$

The x'- and ct' axes are not perpendicular, however their slopes are reciprocal. Note that the 45-degree line $x = ct$ depicts the path of a light ray in both the primed and the unprimed frame. In the limit that $v \to c$, the x'- and ct'-axes converge on $x = ct$.

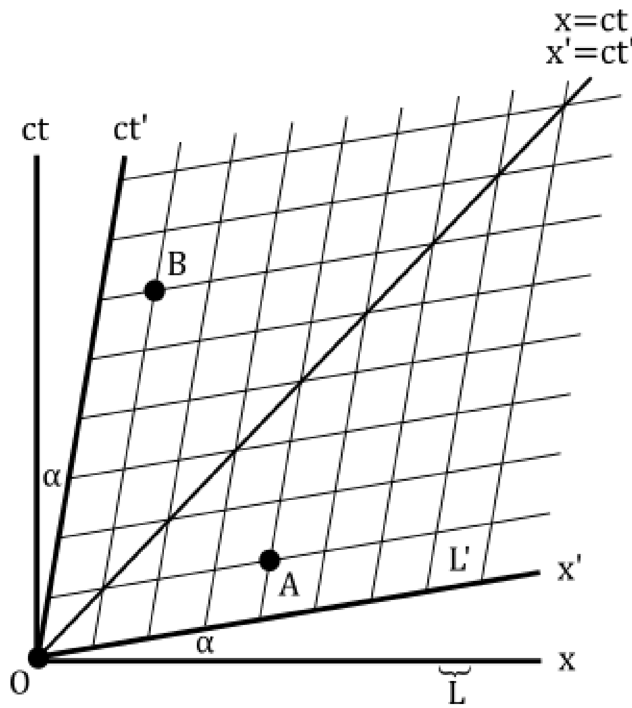

Figure 1: Spacetime diagram. Events OA are spacelike; events OB are timelike. Any event occurring on the like $x/t = x'/t' = c$ is lightlike with event O.

The angle α formed between the ct- and ct'-axes is the same as the angle between x- and x'axes, and is given by

$$\tan\alpha = \frac{v}{c}\,.$$

Scale

To compare the sense of 'unit length' between the two systems, consider a pair of events that are simultaneous in the unprimed reference frame, separated by length $\Delta x = L$. According to the boosted frame, there are two components to track:

$$\Delta x' = \gamma\,(\Delta x - 0) = \gamma L \qquad\qquad c\Delta t' = -\gamma\frac{v}{c}L$$

Calculating $(L')^2 = (\Delta x')^2 + (c\Delta t')^2$, we find

$$L' = L\sqrt{\frac{1 + v^2/c^2}{1 - v^2/c^2}}\,.$$

The same result is attained by considering a different set of events such that $c\Delta t = L$ and $\Delta x = 0$. Either way, the unit length $L' > L$ is the same on each primed axis, and is always larger in the boosted frame.

Light Cone

A given event E in spacetime can only be influenced by prior events that are timelike or lightlike with E. By the same token, the influence of E extends only to future timelike or spacelike events.

For a given event E, the family of all possible past and future timelike events comprise the event's *light cone*. The light cone reduces to a point at E, and 'opens up' along the lines $x = \pm ct$ in the rest frame.

Worldline

The continuous path in spacetime traced out by a physical object is called a *worldline*. An object at rest in its reference frame traces a straight line parallel to its respective *ct*-axis, and any motion introduces a non-parallel component along $\pm x$. The slope of a worldline is

$$m = \frac{c\Delta t}{\Delta x} = \frac{c}{v} \, ,$$

which never deviates more than 45 degrees from the rest path.

Problem 1

The trajectory of a particle moving at speed $c/2$ may be parameterized by the wordline:

$$t = 2\lambda \qquad\qquad x = c\lambda$$

Determine the proper time parameter τ in terms of λ.

Solution 1

Calculate the spacetime interval

$$\Delta S^2 = -c^2 \Delta t^2 + \Delta x^2 = -c^2 2^2 \Delta\lambda^2 + c^2 \Delta\lambda^2 = c^2 \Delta\lambda^2 \left(-4 + 1\right) \, ,$$

where comparison to $\Delta S^2 = -c^2 \Delta\tau^2$ implies:

$$\Delta\tau^2 = 3\Delta\lambda^2 \qquad \rightarrow \qquad \tau = \sqrt{3}\lambda$$

4.9 Minkowski Space

Our study of inertial reference frames has shown that time and space are inseparably woven together in a velocity-dependent 'spacetime fabric'. The four-dimensional manifold on which inertial frames exist is called *Minkowski space*, also known as *flat space*. There are in fact many 'named' spaces, but the majority of them involve curvature, gravity, a strange number of dimensions, or other theoretical toys.

4.10 Paradoxes

Special relativity is a uniquely counter-intuitive theory, and tends to generate nonsense if applied in a sloppy manner. Here we shall open and shut two such cases, called the *ladder paradox* and the *twin paradox*.

Ladder Paradox

Consider a barn, garage, or similar building with 'front' and back' doors, separated by length D, that can open or close instantly. Consider also a ladder of length L, oriented horizontally, that is able to move relativistically toward the front door of the barn. Things get interesting when we insist that $L > D$, meaning the ladder would be too large to fit in the barn at rest.

If the ladder has sufficient velocity, its length appears contracted according to observers in the barn, so there does exist (at least) an instant in which the entire ladder fits within the barn. On the other hand, if we ask an observer on the ladder what happens, the length of the barn is contracted, and it would seem that the ladder is far too long to fit inside the barn.

The so-called 'ladder paradox' addresses the question of whether or not the ladder truly fits within the barn. To illustrate further, suppose an observer at rest in the barn operates a switch to simultaneously close the front and back doors during a moment when the ladder is entirely enclosed, but then re-opens the rear door to avoid a collision. If we imagine the frame of the ladder however, it might seem that the front door closes down on the ladder.

Whether or not the ladder contacts a door cannot be a matter of reference frame - certain facts must be invariant. In order to avoid a paradox, we must note that, in the reference frame of the ladder, *the doors do not operate simultaneously.* Instead, the rear door shuts and opens *before* the ladder has fully entered the barn. The front door shuts only after the ladder passes the doorway, by which time the front of the ladder has moved well-outside the back door. The ladder in fact touches no door, and there is no paradox.

Trapping the Ladder

Let us explore what would happen if the back door remains closed, i.e. stopping the ladder while inside the barn. According to an observer in the barn, the ladder becomes fully enclosed so that the front door shuts. On hitting the back wall, the velocity instantaneously drops to zero and all length contraction vanishes. The ladder has no choice but to snap or explode between the two doors.

In the frame of the ladder however, contact with the rear door takes place with part of the ladder still not having reached the front door, so how can the ladder still be trapped within the barn and explode between the doors? The resolution involves noting that *the ladder does not instantaneously decelerate* as it does in the frame of the barn. The deceleration occurs sequentially along its length such that the front of the ladder stops first, whereas the back of the ladder hasn't 'received the news' to decelerate until after it's already inside the barn. Clearly an object won't hold together under this circumstance, so the ladder blows up in both reference frames, as it should.

Twin Paradox

Imagine two identical twins, one of whom enters a relativistically-enabled rocket ship, travels the solar system at high speed, and returns to Earth. The Earth-bound twin, who considers himself at rest, would conclude that time dilation caused the traveling twin to age more slowly - he returns to Earth younger than he 'should' be. On the other hand however, the traveling twin could insist that Earth was moving through space with respect to his rocket, and the Earth-bound twin should be the younger of the two. Clearly, both twins can't be correct, hence the 'twin paradox'.

The resolution to the twin paradox is simple: the traveling twin must turn around at some point in his journey, so he moves with *two* inertial frames. The Earth-bound twin remains in the same frame throughout the experiment, thus the situation is not symmetric. The traveling twin unambiguously ages less than the Earth-bound twin.

Example

Suppose a twin named Albert always stays on Earth. A second twin, named Barney, travels 5 light-years straight out into space at speed $c/2$ and returns back to Earth at speed $c/4$. Determine (i) how much Albert has aged during Barney's trip, and (ii) how much Barney has aged during his own trip.

According to Albert, the duration of the trip is the sum of two times:

$$\Delta t_A = \frac{5\,ly}{c/2} + \frac{5\,ly}{c/4} = 10\,yr + 20\,yr = 30\,yr$$

According to Barney, the two legs of the trip obey the Lorentz transformation

$$\Delta t_1' = \frac{10\,yr - 5\,ly/2c}{\sqrt{1-1/4}} = 5\sqrt{3}\,yr \qquad\qquad \Delta t_2' = \frac{20\,yr - 5\,ly/4c}{\sqrt{1-1/16}} = 5\sqrt{15}\,yr\ ,$$

where we find

$$\Delta t_B' = \Delta t_1' + \Delta t_2' = 5\left(\sqrt{3}+\sqrt{15}\right)\ yr \approx 28.03\,yr\ .$$

The traveling twin is nearly two years younger than the stationary twin after the trip.

5 Relativistic Mechanics

With the notions of time and space being far less rigid than originally thought, the entire body of classical mechanics must be adjusted.

5.1 Force and Momentum

To derive a relativistic version of Newton's second law, begin by considering an inertial reference frame through which a particle of mass m accelerates in one dimension with speed u. A second inertial frame, having constant speed v with respect to the first, follows behind the particle such that the particle's instantaneous non-relativistic speed is u'.

According to the primed reference frame, the force F' on the particle is

$$F' = m\frac{du'}{dt'} = m\frac{du'}{dt}\frac{dt}{dt'} = m\frac{d}{dt}\left(\frac{u-v}{1-vu/c^2}\right)\frac{d}{dt'}\left(\gamma t' + \gamma v x'/c^2\right) \ ,$$

where the Lorentz transformation and velocity addition formulas have been used. After handling the derivative terms, each being straightforwardly evaluated because v and γ are constant, we let the primed reference frame 'catch up' to the particle by setting

$$u \to v \qquad\qquad u' \to 0 \qquad\qquad F' \to F \ ,$$

reducing the force equation to

$$F = \frac{m}{\left(1-v^2/c^2\right)^{3/2}}\frac{dv}{dt} = m\frac{d}{dt}\left(\frac{v}{\sqrt{1-v^2/c^2}}\right) = m\frac{d}{dt}\left(\gamma v\right) \ .$$

Remarkably, the only adjustment to Newton's second law is a gamma factor tacked onto the velocity. Simultaneously this defines the relativistic momentum p in terms of γv:

$$\vec{F} = \frac{d\vec{p}}{dt} \qquad\qquad \vec{p} = \gamma m \vec{v}$$

Expanding γ at low speeds, the relativistic momentum $p = \gamma m v$ reduces to the Newtonian momentum plus correction terms:

$$p = \gamma m v \approx mv + \frac{mv^3}{2c^2} + \frac{3mv^5}{8c^4} + \frac{5mv^7}{16c^6} + \cdots$$

Uniformly-Accelerated Particle

While special relativity deals strictly with inertial reference frames, we can still integrate the above to solve for the motion of a uniformly-accelerated particle of mass m from rest at $t = 0$:

$$F = \frac{dp}{dt} = m\frac{d}{dt}\left(\gamma v\right) = m\gamma^3\frac{dv}{dt} = \frac{m}{\left(1-v^2/c^2\right)^{3/2}}\frac{dv}{dt}$$

$$\int_0^t \frac{F}{m}\,dt' = \int_0^v \frac{dv'}{\left(1-v'^2/c^2\right)^{3/2}} \qquad\qquad \frac{Ft}{m} = \frac{v}{\sqrt{1-v^2/c^2}}$$

$$v(t) = \frac{c}{\sqrt{1 + (mc/Ft)^2}} \qquad x(t) = \frac{c}{F}\sqrt{(Ft)^2 + (mc)^2} - \frac{mc^2}{F}$$

Identifying

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

a relativistic version of the impulse-momentum theorem can be written for uniform acceleration:

$$Ft = \gamma m v = p$$

5.2 Kinetic Energy and Rest Energy

The kinetic energy of a relativistically-moving particle has no fundamental reason to equal $mv^2/2$ or $p^2/2m$. By definition, the energy of a particle is the work integral of the force vector projected onto the displacement vector, namely

$$E = \int_{x_i}^{x_f} \vec{F} \cdot d\vec{x} = m \int_{t_i}^{t_f} \frac{d}{dt}(\gamma \vec{v}) \cdot \vec{v}\, dt = m \int_{t_i}^{t_f} \left(\frac{d\gamma}{dt}\, \vec{v} + \gamma \frac{d\vec{v}}{dt} \right) \cdot \vec{v}\, dt .$$

As a side problem, note that the quantity $\vec{v} \cdot d\vec{v}/dt$ readily reduces to $v\, dv/dt$ for one-dimensional motion, but this is also true for three dimensions:

$$\begin{aligned}
v \frac{dv}{dt} &= \sqrt{v_x^2 + v_y^2 + v_z^2}\, \frac{d}{dt}\left(\sqrt{v_x^2 + v_y^2 + v_z^2}\right) \\
v \frac{dv}{dt} &= \cancel{\sqrt{v_x^2 + v_y^2 + v_z^2}}\, \frac{1}{\cancel{\sqrt{v_x^2 + v_y^2 + v_z^2}}} \left(v_x \frac{dv_x}{dt} + v_y \frac{dv_y}{dt} + v_z \frac{dv_z}{dt} \right) \\
v \frac{dv}{dt} &= \vec{v} \cdot \frac{d\vec{v}}{dt}
\end{aligned}$$

The energy integral becomes

$$E = m \int_{t_i}^{t_f} \left(\frac{d\gamma}{dt}\, v + \gamma \frac{dv}{dt} \right) v\, dt = m \int_{t_i}^{t_f} \frac{d}{dt}(\gamma v)\, v\, dt ,$$

which can be integrated by parts via

$$\int U\, dV = UV\Big| - \int V\, dU ,$$

where

$$U = v \qquad dV = \frac{d}{dt}(\gamma v)\, dt \qquad dU = dv \qquad V = \gamma v ,$$

to arrive at

$$E = m\gamma v^2\Big|_{v_i}^{v_f} - m \int_{v_i}^{v_f} \frac{v}{\sqrt{1 - v^2/c^2}}\, dv .$$

Assuming the particle starts from rest and accelerates to velocity v, the energy simplifies to

$$E = m\gamma v^2 + \frac{mc^2}{\gamma} = \gamma mc^2 \ ,$$

which is a curious formula for kinetic energy, as E is nonzero when $v = 0$. Indeed, a particle at rest with $\gamma = 1$ still has an energy equal to mc^2. Einstein realized the novelty of this so-called *rest energy* E_0, and coined the most famous equation in the world, namely

$$E_0 = mc^2 \ .$$

The true kinetic energy subtracts away the rest energy from E, so we must have

$$T = mc^2 \left(\gamma - 1\right) \ ,$$

which for low speeds reproduces the Newtonian result plus correction terms:

$$T \approx \frac{mv^2}{2} \left(1 + \frac{3v^2}{4c^2} + \frac{5v^4}{8c^4} + \cdots\right)$$

Acceleration and Energy

In Newtonian physics, the energy of a particle changes at a rate given by

$$\frac{dE}{dt} = \frac{d}{dt} \int \vec{F} \cdot d\vec{x} = \frac{d}{dt} \int \left(\frac{d\vec{p}}{dt} \cdot \vec{v}\right) \ dt = \frac{d\vec{p}}{dt} \cdot \vec{v} \ .$$

In the relativistic regime, we start with modified equations but get the same result:

$$\frac{d}{dt} E^2 = \frac{d}{dt} \left(p^2c^2 + m^2c^4\right) \qquad \rightarrow \qquad 2E\frac{dE}{dt} = 2c^2 \, \vec{p} \cdot \frac{d\vec{p}}{dt} \qquad \rightarrow \qquad \frac{dE}{dt} = \frac{d\vec{p}}{dt} \cdot \vec{v}$$

5.3 Energy-Momentum Relation

The momentum equation $\vec{p} = m\gamma\vec{v}$ and the energy equation $E = m\gamma c^2$ can be combined to eliminate all factors of $\vec{v}$ and γ. First square the momentum equation, and noting that

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

we find

$$p^2 = -m^2c^2 + \gamma^2m^2c^2 \ ,$$

where substituting $E^2 = \gamma^2m^2c^4$ gives the *energy-momentum relation*:

$$E^2 = p^2c^2 + m^2c^4$$

5.4 Massless Particles

So far, we've built several equations for the energy and momentum of massive particles, namely

$$E = \gamma m c^2 \qquad \vec{p} = \gamma m \vec{v} \qquad E^2 = p^2c^2 + m^2c^4 \, .$$

It would appear that we reach a dead end by exploring $m = 0$, however if we simultaneously let

$$m \to 0 \qquad v \to c \, ,$$

a curious limit emerges. The quantity $m\gamma$ becomes ambiguous, so for now let us denote

$$E_\gamma = \lim_{m\to 0, v\to c} m\gamma c^2$$

as an unknown (but not constant). In terms of E_γ, the momentum vector becomes

$$\vec{p} = \lim_{v\to c} \left(\frac{E_\gamma}{c^2}\right) \vec{v} = \left(\frac{E_\gamma}{c}\right) \hat{k} \, ,$$

where the unit vector $\hat{k}$ denotes the direction of motion. Note this result is consistent with the energy-momentum relation in the limit $m \to 0$, telling us

$$E_\gamma = pc \, .$$

The equations for energy and momentum remain intact if we accept simultaneously that massless particles can exist, but *only* travel at speed c. Of course, the only thing in nature that travels at speed c is light itself, thus any massless particle is regarded as a small packet of radiation called a *photon*.

Photons

Being a wave, a monochromatic photon has a wavelength λ and a frequency f related by

$$c = \lambda f \, ,$$

where using $E = pc$, we have

$$\frac{E_\gamma}{f} = p\lambda \, ,$$

and by construction, both sides of the equation are equal to a constant. Historically, we denote the so-called *Planck's constant* as h, giving simultaneous relations:

$$E_\gamma = hf \qquad p = \frac{h}{\lambda} \, .$$

Since wave equations (especially those describing light) have sinusoidal solutions, it's natural to define the angular frequency ω and the wavenumber k such that

$$\omega = 2\pi f \qquad k = \frac{2\pi}{\lambda} \, .$$

Planck's constant can also be modified by a factor 2π to give the 'h-bar' form:

$$\hbar = \frac{h}{2\pi}$$

In terms of $\hbar$, ω, and k, the energy and momentum simplify nicely:

$$E = \hbar\omega \qquad\qquad \vec{p} = \left(\frac{E_\gamma}{ck}\right)\vec{k} = \left(\frac{hf}{c}\frac{\lambda}{2\pi}\right)\vec{k} = \hbar\,\vec{k}$$

5.5 Conservation of Mass-Energy

The mass m of a particle or body has been slipped into force, momentum, and energy calculations under the assumption that mass is constant. This assumption is correct, provided that m refers to the mass as observed in the rest-frame of the body. Like c, the *rest mass* of a body is an invariant of motion.

For bodies in motion however, mass and energy become intermixed. As we'll see, the results previously derived predict that mass and energy are not individually conserved in relativistic dynamics, but are instead conserved as a pair.

Relativistic Mass (Optional)

The total energy of a body in motion is the rest energy plus the kinetic energy, conveniently condensed in $E = \gamma mc^2$. The quantity

$$m_\gamma = \gamma m$$

is called the *relativistic mass*, which is clearly not an invariant. Instead, the relativistic mass increases as v increases, meaning that moving bodies appear to be 'heavier' then they should be.

Since we already understand a particle's energy as the sum of mc^2 plus a kinetic term, the notion of relativistic mass is redundant and often misleading. It's included here for the sake of completeness, however the m_γ-notation will be mostly avoided.

Collision of Identical Masses

Consider an observer at rest who prepares two identical bodies (balls of clay perhaps), each of mass m, to move toward each other in opposing directions. If each body has the same initial kinetic energy T, the total energy of the system is

$$E = 2 \times \left(mc^2 + T\right) \ ,$$

and the total linear momentum is zero.

If the objects collide and merge together (with no energy loss from heat, elasticity, etc.), the resulting body emerges at rest, having total energy $E = Mc^2$, where M is the rest mass. Eliminating E and solving for M gives

$$M = 2m + \frac{2T}{c^2} \ .$$

Evidently, the kinetic energy is absorbed into the resultant body's rest mass. Eliminating T, the resultant mass is nicely expressed in terms of m and γ, namely

$$M = 2\gamma m \, .$$

Decay of a Motionless Particle

Suppose a particle of mass M at rest spontaneously decays into two particles having mass m_j, momentum p_j, and energy E_j, where $j = 1, 2$. The momentum-energy relation applies before and after the decay, so we begin with

$$Mc^2 = \sqrt{p_1^2 c^2 + m_1^2 c^4} + \sqrt{p_2^2 c^2 + m_2^2 c^4} \qquad \vec{p}_1 + \vec{p}_2 = 0 \, .$$

Square both sides and replace remaining square root terms with E_j. Also note that $p_1^2 + p_2^2 = -2\vec{p}_1 \cdot \vec{p}_2$ to arrive at

$$M^2 = m_1^2 + m_2^2 + 2\left(\frac{E_1 E_2}{c^4} - \frac{\vec{p}_1 \cdot \vec{p}_2}{c^2}\right) \, .$$

The above can be independently solved for each E_j, resulting in:

$$E_1 = c^2 \, \frac{M^2 + m_1^2 - m_2^2}{2M} \qquad E_2 = c^2 \, \frac{M^2 + m_2^2 - m_1^2}{2M}$$

5.6 Momentum-Energy Lorentz Transformation

The Lorentz transformation

$$x' = \gamma \left(x - vt\right) \qquad t' = \gamma \left(t - vx/c^2\right)$$

$$x = \gamma \left(x' + vt'\right) \qquad t = \gamma \left(t' + vx'/c^2\right)$$

can be recast in terms of energies E, E' and momenta p, p'. To begin, consider two reference frames in relative motion at velocity v, where a particle moving in one dimension is observed in each frame to have velocity u, u', respectively.

In the non-boosted frame, the energy and momentum read

$$E = \frac{mc^2}{\sqrt{1 - u^2/c^2}} \qquad p = \frac{mu}{\sqrt{1 - u^2/c^2}} \, .$$

The boosted energy E' of the particle is

$$E' = \gamma' mc^2 = \frac{mc^2}{\sqrt{1 - \left(u'\right)^2/c^2}} = mc^2 \Big/ \sqrt{1 - \frac{1}{c^2}\left(\frac{u - v}{1 - vu/c^2}\right)^2} \, ,$$

where the velocity addition formula has been used. Simplifying inside the square root, E' becomes

$$E' = \frac{mc^2\left(1 - vu/c^2\right)}{\sqrt{\left(1 - v^2/c^2\right)\left(1 - u^2/c^2\right)}} = \gamma\left(v\right)\gamma\left(u\right) mc^2 \left(1 - vu/c^2\right) = \gamma\left(E - vp\right) \, .$$

Repeating the exercise for $p' = \gamma' m u'$ and making similar substitutions, we find

$$p' = \gamma' m u' = \frac{m u'}{\sqrt{1 - (u')^2 / c^2}} = \gamma(v)\,\gamma(u)\,m\,(u - v) = \gamma\left(p - vE/c^2\right) \ .$$

The inverse relations come from solving for for E and p in terms of unprimed variables:

$$E = \gamma\left(E' + vp'\right) \qquad\qquad p = \gamma\left(p' + vE'/c^2\right)$$

For pairs of events, we may also write the delta-version of the above:

$$\Delta E' = \gamma\left(\Delta E - v\Delta p\right) \qquad\qquad \Delta E = \gamma\left(\Delta E' + v\Delta p'\right)$$

$$\Delta p' = \gamma\left(\Delta p - \frac{v\Delta E}{c^2}\right) \qquad\qquad \Delta p = \gamma\left(\Delta p' + \frac{v\Delta E'}{c^2}\right)$$

Minkowski Invariant

Consider a photon moving in one dimension with constant energy $E_\gamma = hf$ and momentum $p = h/\lambda$. Integrating E_γ with respect to time, we have

$$E_\gamma \int dt = \frac{h}{\lambda} \int c\, dt \qquad \rightarrow \qquad E_\gamma t = px + \text{constant} \ .$$

Evidently, the quantity $px - E_\gamma t$ is a constant of motion for photons. Naturally, we may wonder if $px - Et$ is constant for massive particles. Calculating this out, we find

$$\begin{aligned}
px - Et &= \gamma\left(p' + \frac{vE'}{c^2}\right)\gamma\left(x' + vt'\right) - \gamma\left(E' + vp'\right)\gamma\left(t' + \frac{vx'}{c^2}\right) \\
px - Et &= \gamma^2 p'x'\left(1 - \frac{v^2}{c^2}\right) + \gamma^2 E't'\left(\frac{v^2}{c^2} - 1\right) \\
px - Et &= p'x' - E't' \ .
\end{aligned}$$

Generalizing to three dimensions, we find that $\vec{p} \cdot \vec{x} - Et$ is not only constant, but is yet another spacetime invariant that we'll call the *Minkowski invariant*. For the case of photons, the momentum is $p = h/\lambda = \hbar k$, and the energy is $E_\gamma = hf = \hbar\omega$, thus the Minkowski invariant takes the form

$$\vec{k} \cdot \vec{x} - \omega t = (\vec{k})' \cdot (\vec{x})' - \omega' t' \ .$$

5.7 Relativistic Doppler Effect

Our study has assumed that the speed of light is the same in all reference frames, however this should not mean that light carries the same energy in all frames, a phenomenon called the *Doppler effect*. To derive this, start with the Minkowski invariant and replace all substitute $E = hf$ and $p = h/\lambda = hf/c$ to write

$$px - Et = p'x' - E't' \qquad \rightarrow \qquad \frac{xf}{c} - ft = \frac{x'f'}{c} - f't' \ .$$

Replace x' and t' using the Lorentz transformation and simplify:

$$
\begin{aligned}
\frac{xf}{c} - ft &= f'\frac{\gamma\,(x+vt)}{c} - f'\gamma\left(t + \frac{vx}{c^2}\right) \\
f\cancel{\left(\frac{x}{c} - t\right)} &= \gamma f' \cancel{\left(\frac{x}{c} - t\right)} \left(1 - \frac{v}{c}\right) \\
f &= f' \, \frac{1 - v/c}{\sqrt{1 - v/c}\sqrt{1 + v/c}} \\
f' &= f \, \frac{\sqrt{1 + v/c}}{\sqrt{1 - v/c}}
\end{aligned}
$$

We see that the observed frequency of light is increased when the source and observer are moving toward each other, an effect called *blueshift*. The other case when source and observer are receding is called *redshift*.

3D Doppler Effect

Consider a light ray propagating in space with momentum vector $\vec{p} = h\vec{k}$ according to an inertial reference frame. A second reference frame having velocity v with respect to the first will observe the same light ray to have a primed momentum $(\vec{p})' = h(\vec{k})'$. As a sinusoidal wave, the electric field components E and E' respectively, are

$$
E = E_0 \cos\left(\vec{k} \cdot \vec{x} - \omega t\right) \qquad\qquad E' = E_0' \cos\left(\vec{k'} \cdot \vec{x'} - \omega' t'\right) \; .
$$

Without loss of generality, assume all relative motion takes place in two dimensions. That is, let the frame boost be along x, and let the vectors $\vec{k}$, $(\vec{k})'$ live somewhere in the xy-plane. Each vector makes an angle θ, θ' with respect to the horizontal x- and x'-axes.

To proceed, take each electric field equation and bloom out the dot product to get

$$
\begin{aligned}
E &= E_0 \cos\left(2\pi\left(\frac{x\cos\theta + y\sin\theta}{\lambda} - ft\right)\right) \\
E' &= E_0' \cos\left(2\pi\left(\frac{x'\cos\theta' + y'\sin\theta'}{\lambda} - f't'\right)\right) \; .
\end{aligned}
$$

The relevant Lorentz transformation relations are

$$
x' = \gamma\,(x - vt) \qquad\qquad t' = \gamma\left(t - vx/c^2\right) \qquad\qquad y' = y \; ,
$$

telling us

$$
E' = E_0' \cos\left(2\pi\left(\frac{\gamma x}{\lambda'}\left(\cos\theta' + \frac{v}{c}\right) + \frac{y\sin\theta'}{\lambda'} - \gamma f' t\left(1 + \frac{v}{c}\cos\theta'\right)\right)\right) \; .
$$

Since the cosine argument in E, E' is the Minkowski invariant for photons, we set each argument equal to pick out three simultaneous relations:

$$
\frac{\cos\theta}{\lambda} = \frac{\gamma}{\lambda'}\left(\cos\theta' + \frac{v}{c}\right) \qquad\qquad \frac{\sin\theta}{\lambda} = \frac{\sin\theta'}{\lambda'} \qquad\qquad f = \gamma f'\left(1 + \frac{v}{c}\cos\theta\right)
$$

The third equation is the most revealing. Consider the case $\theta = 0$, corresponding to source receding away from the observer (or vice versa). This gives

$$f'_{\text{recede}} = f\,\frac{\sqrt{1 - v/c}}{\sqrt{1 + v/c}}\,,$$

where f is considered the 'proper' frequency. On the other hand, setting $\theta = \pi$ predicts the opposite signs when the source and observer are approaching:

$$f'_{\text{approach}} = f\,\frac{\sqrt{1 + v/c}}{\sqrt{1 - v/c}}$$

6 Index Notation

To begin developing new tools and notation, recall the spacetime invariant we named the 'interval':

$$\Delta S^2 = -c^2\Delta t^2 + \Delta x^2 + \Delta y^2 + \Delta z^2 = -c^2\Delta t^2 + \Delta\vec{x}\cdot\Delta\vec{x}$$

The spacetime interval almost has the appearance of a four-dimensional vector $(c\Delta t, \Delta\vec{x})$ projected onto itself if it weren't for the minus sign on the Δt^2-term.

6.1 Four-Vectors

To understand ΔS^2 as a kind of dot product, let us propose there are two manifestations of the vector $(c\Delta t, \Delta\vec{x})$, denoted Δq^μ and Δq_μ, respectively, where $\mu = 0, 1, 2, 3$. Next, we write

$$\Delta q^\mu \Delta q_\mu = -c^2\Delta t^2 + \Delta\vec{x}\cdot\Delta\vec{x}\ ,$$

where clearly one of the Δq-terms contains the minus sign, resolved by setting

$$\Delta q^\mu = (c\Delta t, \Delta\vec{x}) \qquad\qquad \Delta q_\mu = (-c\Delta t, \Delta\vec{x})\ .$$

The terms Δq^μ and Δq_μ are known as *four-vectors*. The high-index version is called *contravariant*, and the low-index version is called *covariant*.

Position Four-Vector

Removing the delta symbols (by integration if simply erasing them seems dirty), we have the *position four-vector* (in both forms):

$$q^\mu = (ct, \vec{x}) \qquad\qquad q_\mu = (-ct, \vec{x})$$

Velocity Four-Vector

The derivative of the position vector q^μ with respect to the proper time τ gives the *velocity four-vector*:

$$U^\mu = \frac{d}{d\tau}q^\mu = \left(c\frac{dt}{d\tau}, \frac{d\vec{x}}{dt}\frac{dt}{d\tau}\right) = (\gamma c, \gamma\vec{v})$$

Momentum Four-Vector

Multiply the velocity four-vector by the rest mass m of a particle to get the *momentum four-vector*:

$$P^\mu = mU^\mu = (\gamma mc, \gamma m\vec{v}) = \left(\frac{E}{c}, \vec{p}\right)$$

Force Four-Vector

A τ-derivative of the momentum four-vector gives the *force four-vector*:

$$F^\mu = \frac{d}{d\tau}P^\mu = \frac{dt}{d\tau}\left(\frac{1}{c}\frac{dE}{dt}, \frac{d\vec{p}}{dt}\right) = \gamma\left(\frac{\vec{F}\cdot\vec{v}}{c}, \vec{F}\right)$$

Photon Four-Vector

For massless particles, the momentum four-vector P^μ is renamed to K^μ, abbreviated γ. The form of K^μ is the same as its predecessor with $\vec{p}$ replaced by $\hbar\vec{k}$:

$$K^\mu = \gamma = \left(\frac{E_\gamma}{c}, \hbar\vec{k}\right)$$

6.2 Contraction

The notion of 'projection' or 'dot product' in the case of four-vectors is replaced by *contraction*, which entails summing the products of similar components of a contravariant four-vector with a covariant four-vector. One already-familiar contraction is

$$\Delta q^\mu \Delta q_\mu = \Delta q^0 \Delta q_0 + \Delta q^1 \Delta q_1 + \Delta q^2 \Delta q_2 + \Delta q^3 \Delta q_3 \ ,$$

simplifying to

$$\Delta S^2 = -c^2 \Delta\tau^2 \ .$$

The contraction of two four-vectors always results in a scalar.

Norm

The contraction of a four-vector onto itself is called the *norm* of the four-vector. The norm of the velocity four-vector is easily calculated,

$$U^\mu U_\mu = -\gamma^2 c^2 + \gamma^2 v^2 = -c^2\gamma^2\left(1 - v^2/c^2\right) = -c^2 \ ,$$

as is the norm of the four-momentum:

$$P^\mu P_\mu = -m^2 c^2$$

The advantage of four-vectors is to tightly contain the equations of kinematics. For instance, bloom out the velocity four-vector to recover the gamma factor

$$-c^2\Delta\tau^2 = \Delta q^\mu \Delta q_\mu = -c^2\Delta t^2 + \Delta x^2 \qquad \rightarrow \qquad \frac{dt}{d\tau} = \frac{1}{\sqrt{1 - v^2/c^2}} \ ,$$

and repeat with the momentum four-vector to recover the energy formula:

$$-m^2c^2 = P^\mu P_\mu = -\frac{E^2}{c^2} + p^2 \qquad \rightarrow \qquad E^2 = p^2c^2 + m^2c^4$$

6.3 Invariants

Invariant quantities are those that are the same in all inertial reference frames. The contractions $U^\mu U_\mu = -c^2$ and $P^\mu P_\mu = -m^2c^2$ always yield invariant quantities. The first of these we've encountered is the spacetime interval:

$$\Delta S^2 = -c^2\Delta\tau^2$$

The contraction between the four-force and the four-velocity is also invariant:

$$F^\mu U_\mu = mU_\mu \frac{d}{d\tau} U^\mu = \frac{m}{2} \frac{d}{d\tau} \left(U^\mu U_\mu \right) = \frac{m}{2} \frac{d}{d\tau} \left(-c^2 \right) = 0$$

We may also try $q^\mu P_\mu$, and the Minkowski invariant $-Et + \vec{p} \cdot \vec{x}$ happens to fall right out:

$$q^\mu P_\mu = -Et + \vec{p} \cdot \vec{x}$$

Applying this to a particle in one dimension, rearrange the right side of the above to get

$$q^\mu P_\mu = -\gamma mc^2 t - \gamma mvx = mc^2 \gamma \left(-t + \frac{vx}{c^2} \right) = -mc^2 t' \ ,$$

and take a τ-derivative of each side:

$$-mc^2 + q^\mu F_\mu = -mc^2 \frac{dt'}{d\tau}$$

Solving for $q^\mu F_\mu$, we have

$$q^\mu F_\mu = mc^2 \left(1 - \frac{dt'}{d\tau} \right) \ .$$

On the other hand, $q^\mu F_\mu$ can be calculated directly from four-vectors

$$q^\mu F_\mu = -\gamma t F v + \gamma F x = F\gamma \left(x - vt \right) = Fx' \ ,$$

from which we deduce:

$$mc^2 \left(1 - \frac{dt'}{d\tau} \right) = Fx'$$

6.4 Metric Tensor

One question naturally implied by four-vectors concerns whether there exists an object or operation that converts a contravariant four-vector to a covariant one, or vice-versa. It turns out that contraction with the *metric tensor* does just this. A *tensor* is the full generalization of a four-vector: there can be any number of components, and any number of indices in the up- or down- position. One particularly special tensor is called the *flat space metric*, also known as the *Minkowski space metric*, denoted $\eta_{\mu\nu}$, defined as

$$\eta_{\mu\nu} = \begin{bmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \ .$$

Note $\eta_{\mu\nu}$ is *not* a matrix.

Raising and Lowering Indices

To see $\eta_{\mu\nu}$ at work, consider the contraction $\eta_{\mu\nu}A^\mu$, where A^μ is any four-vector. Note the μ-index occurs twice, both in the up- and down- positions. When any index is repeated, contraction triggers a sum over that index. This means

$$\eta_{\mu\nu}A^\mu = \sum_{\mu=0}^{3} \eta_{\mu\nu}A^\mu \ ,$$

however the summation symbol is almost always ignored, a shortcut called the *Einstein summation convention*. Carrying out the calculation, we find:

$$\begin{aligned}
\eta_{\mu\nu}A^\mu &= \eta_{00}A^0 + \eta_{11}A^1 + \eta_{22}A^2 + \eta_{33}A^3 \\
\eta_{\mu\nu}A^\mu &= -A^0 + A^1 + A^2 + A^3 \\
\eta_{\mu\nu}A^\mu &= A_\nu
\end{aligned}$$

Similarly, the *inverse metric tensor*, has both indices in the up-position, namely $\eta^{\mu\nu}$. In flat space (but not generally for curved space), the inverse metric components are the same as the ordinary metric:

$$\eta^{\mu\nu} = \eta_{\mu\nu}$$

By identical arguments, it follows that the inverse metric can raise the index on a four-vector, namely

$$\eta^{\mu\nu}A_\nu = A^\mu \ .$$

Note that the definition of the inverse metric tensor tells us the contraction between an up- and down-index yields a Kronecker delta function:

$$\eta_{\nu\rho}\eta^{\rho\mu} = \delta^\mu_\nu \ ,$$

where δ^ν_μ resolves to 1 if $\mu = \rho$, and equals zero otherwise.

Proper Time Differential

The differential version of the spacetime interval can be easily expressed in terms of the metric and two contravariant vectors:

$$dq^\mu dq_\mu = \eta_{\mu\nu}dq^\mu dq^\nu = dS^2 = -c^2 d\tau^2$$

Proper Time Integral

Using $dS^2 = -c^2 d\tau^2$, we can isolate $d\tau$ and set up an integral:

$$\tau = \int d\tau = \frac{1}{c}\int \sqrt{-\eta_{\mu\nu}dq^\mu dq_\nu}$$

6.5 Lorentz Transformation Tensor

A boosted reference frame is characterized by its observed velocity v, which when re-parameterized in terms of rapidity, namely $v = c\tanh\phi$, the Lorentz transformation appears as

$$\begin{bmatrix} ct' \\ x' \\ y' \\ z' \end{bmatrix} = \begin{bmatrix} \cosh\phi & -\sinh\phi & 0 & 0 \\ -\sinh\phi & \cosh\phi & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} ct \\ x \\ y \\ z \end{bmatrix} .$$

For multiply-boosted frames, recall that velocities don't stack linearly, however the rapidity is indeed linear, as in $\Phi = \phi' + \phi$.

The above is written in a matrix-like form, but this is merely a convenience. In the preferred four-vector and tensor language, the Lorentz transformation is written

$$(q')^{\mu} = \Lambda^{\mu}_{\nu} q^{\nu} ,$$

where q^{μ} is the position four-vector (ct, x, y, z), and Λ^{μ}_{ν} is the *Lorentz transformation tensor*. The above also applies in differential form, namely

$$(dq')^{\mu} = \Lambda^{\mu}_{\nu} dq^{\nu} .$$

To (re-)derive the restriction on Λ, begin with the differential spacetime interval $(dS')^2 = \eta_{\mu\nu} (dq')^{\mu} (dq')^{\nu}$ and replace each dq'-term with the un-boosted (unprimed) version, giving

$$(dS')^2 = \eta_{\mu\nu} \Lambda^{\mu}_{\alpha} \Lambda^{\nu}_{\beta} dq^{\alpha} dq^{\beta} ,$$

and since $(dS')^2$ is an invariant, we replace the left side with $dS^2 = \eta_{\alpha\beta} dq^{\alpha} dq^{\beta}$ to get:

$$\begin{aligned} \eta_{\alpha\beta} dq^{\alpha} dq^{\beta} &= \eta_{\mu\nu} \Lambda^{\mu}_{\alpha} \Lambda^{\nu}_{\beta} dq^{\alpha} dq^{\beta} \\ \eta_{\alpha\beta} &= \eta_{\mu\nu} \Lambda^{\mu}_{\alpha} \Lambda^{\nu}_{\beta} \end{aligned}$$

To proceed, take various choices of α and β and write the surviving terms. Two interesting cases are

$$\begin{aligned} \alpha = 0, \beta = 0 : & \qquad 1 = \left(\Lambda^{0}_{0}\right)^2 - \left(\Lambda^{1}_{0}\right)^2 \\ \alpha = 1, \beta = 1 : & \qquad 1 = \left(\Lambda^{1}_{1}\right)^2 - \left(\Lambda^{0}_{1}\right)^2 , \end{aligned}$$

hinting that the Λ-components ring like hyperbolic functions. Indeed, by identifying

$$\Lambda^{0}_{0} = \Lambda^{1}_{1} = \cosh\phi \qquad\qquad \Lambda^{1}_{0} = \Lambda^{0}_{1} = -\sinh\phi ,$$

we recover the structure previously derived.

Multiple Boosts

By studying velocity addition under two consecutive boosts, we found

$$\begin{bmatrix} ct'' \\ x'' \\ y'' \\ z'' \end{bmatrix} = \begin{bmatrix} \cosh(\phi' + \phi) & -\sinh(\phi' + \phi) & 0 & 0 \\ -\sinh(\phi' + \phi) & \cosh(\phi' + \phi) & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} ct \\ x \\ y \\ z \end{bmatrix} ,$$

which, for $\Phi = \phi' + \phi$, is written

$$(q'')^{\mu} = \Lambda\,(\Phi)^{\mu}_{\nu}\, q^{\nu} \ .$$

Meanwhile, an equivalent expression for the doubly-boosted position four-vector is

$$(q'')^{\mu} = \Lambda\,(\phi')^{\mu}_{\rho}\, \Lambda\,(\phi)^{\rho}_{\nu}\, q^{\nu} \ .$$

By comparing each q''-equation, we see that Lorentz transformation tensors stack together under two boosts, namely

$$\Lambda\,(\Phi)^{\mu}_{\nu} = \Lambda\,(\phi')^{\mu}_{\rho}\, \Lambda\,(\phi)^{\rho}_{\nu} \ ,$$

which readily generalizes for more than two boosts. This is readily verified by utilizing the hyperbolic trigonometric identities:

$$\cosh\,(a \pm b) = \cosh a \cosh b \pm \sinh a \sinh b$$

$$\sinh\,(a \pm b) = \sinh a \cosh b \pm \cosh a \sinh b$$

Inverse Lorentz Transformation

The *inverse Lorentz transformation tensor*, in index notation, is given by

$$q^{\mu} = \left(\Lambda^{-1}\right)^{\mu}_{\nu} (q')^{\nu} \ .$$

Substituting for q' gives

$$q^{\mu} = \left(\Lambda^{-1}\right)^{\mu}_{\nu} \Lambda^{\nu}_{\rho} q^{\rho} \ .$$

The combination $(\Lambda^{-1})^{\mu}_{\nu} \Lambda^{\nu}_{\rho}$ must resolve to a Kronecker delta function with μ as the up-index and ρ as the down-index:

$$\left(\Lambda^{-1}\right)^{\mu}_{\nu} \Lambda^{\nu}_{\rho} = \delta^{\mu}_{\rho}$$

General Transformations

The Lorentz transformation applies to uniformly-boosted reference frames, i.e. inertial frames. Special cases of the Lorentz transformation also include four-space translations and three-space rotations of the position vector, a 10-member collection called the *Poincare group*.

For non-inertial frames, namely those that contain gravity, forces, etc., the Lorentz transformation is replaced with machinery from differential geometry. For the simple case of one-index vectors V^{μ} and V_{μ}, the transformation law reads

$$V^{\mu'} = \frac{\partial q^{\mu'}}{\partial q^{\mu}} V^{\mu} \qquad\qquad V_{\mu'} = \frac{\partial q^{\mu}}{\partial q^{\mu'}} V_{\mu} \ .$$

6.6 Uniform Acceleration

Recall that a particle of mass m initially at rest while subject to a constant force F has equations of motion:

$$v(t) = \frac{c}{\sqrt{1 + (mc/Ft)^2}} \qquad\qquad x(t) = \frac{c}{F}\sqrt{(Ft)^2 + (mc)^2} - \frac{mc^2}{F}$$

Of course, the combination F/m is interpreted as the uniform acceleration $\tilde{a}$.

Note that $v(t)$ does not have uniform slope with respect to t, however the combination γv must. The calculation is made quite simple by substituting

$$\gamma(t) = \cosh(\phi(t)) \qquad\qquad v(t) = c\tanh(\phi(t))$$

to write

$$\frac{d}{dt}(\gamma v) = \frac{d}{dt}(\cosh\phi \cdot c\tanh\phi) = c\frac{d}{dt}\sinh\phi = c\cosh\phi\frac{d}{dt}\phi(t) \ ,$$

or

$$\frac{d}{dt}(\gamma v) = c\cancel{\gamma}\frac{d}{d\tau}\phi(\tau)\frac{d\cancel{\tau}}{\cancel{dt}} = c\frac{d}{d\tau}\phi(\tau) \ .$$

Evidently, the uniform acceleration $\tilde{a}$ is proportional to the proper-time derivative of the rapidity:

$$\tilde{a} = \frac{d}{dt}(\gamma v) = c\frac{d}{d\tau}\phi(\tau)$$

To see this in terms of four-vectors, start with the velocity

$$U^\mu = (\gamma c, \gamma v) = c(\cosh\phi, \sinh\phi) \ ,$$

which has a derivative A^μ with respect to proper time:

$$A^\mu = \frac{d}{d\tau}U^\mu = c\frac{d}{d\tau}\phi(\tau)(\sinh\phi, \cosh\phi)$$

The norm of the acceleration vector, namely $A^\mu A_\mu$, is straightforwardly calculated:

$$A^\mu A_\mu = \left(c\frac{d}{d\tau}\phi(\tau)\right)^2\left(-\sinh^2\phi + \cosh^2\phi\right) = \left(c\frac{d}{d\tau}\phi(\tau)\right)^2 = \tilde{a}^2$$

The same connection emerges using the Lorentz transformation, where the boosted acceleration components $(A')^\mu$ are

$$(A')^\mu = \Lambda^\mu_\nu A^\nu$$

such that

$$(A')^0 = \Lambda^0_0 A^0 + \Lambda^0_1 A^1 = 0 \qquad\qquad (A')^1 = \Lambda^1_0 A^0 + \Lambda^1_1 A^1 = \tilde{a} \, ,$$

so the final four-vector is

$$(A')^\mu = (0, \tilde{a}) \ .$$

The ϕ-equation can be integrated with respect to τ to get

$$\phi(\tau) = \frac{\tilde{a}}{c}\tau + \phi_0 \,,$$

where the integration constant ϕ_0 is set to zero. With this known, the four-velocity reads

$$U^\mu = c\left(\cosh\left(\frac{\tilde{a}\tau}{c}\right), \sinh\left(\frac{\tilde{a}\tau}{c}\right)\right) \,.$$

Integrating with respect to τ, the position four-vector emerges:

$$q^\mu = \frac{c^2}{\tilde{a}}\left(\sinh\left(\frac{\tilde{a}\tau}{c}\right), \cosh\left(\frac{\tilde{a}\tau}{c}\right)\right) + q_0^\mu$$

After some algebra, it turns out that q^μ is equivalent to the pair of equations for $v(t)$ and $x(t)$ above.

6.7 Affine Parameter

The position four-vector $q^\mu = (ct, \vec{x})$ traces a worldline in Minkowski space. The components of q^μ are often expressed in terms of a time parameter t or τ, but not strictly so. A general parameterization λ is called *affine* if $\lambda \propto \tau$.

Suppose the position four-vector q^μ depends on a parameter λ to trace a worldline. In the small neighborhood around some $\tilde{\lambda}$ along the worldline, we have

$$dq^\mu(\lambda) = q^\mu\left(\tilde{\lambda} + d\lambda\right) - q^\mu\left(\tilde{\lambda}\right) = \left(\frac{d}{d\lambda}q^\mu(\lambda)\right)\bigg|_{\tilde{\lambda}} d\lambda \,,$$

where the derivative on the right is interpreted as the *tangent four-vector*. It follows that

$$dS^2 = dq^\mu dq_\mu = \eta_{\mu\nu}dq^\mu dq^\nu = -\left(\frac{d}{d\lambda}q^0(\lambda)\right)^2 (d\lambda)^2 + \left(\frac{d}{d\lambda}\vec{q}(\lambda)\right)^2 (d\lambda)^2 \,,$$

equivalent to, by the chain rule,

$$dS^2 = -\left(\frac{dt}{d\lambda}\right)^2 (d\lambda)^2 \left(c^2 - \left(\frac{d\vec{q}}{dt}\right)^2\right) = -c^2\left(\frac{dt}{d\lambda}\right)^2 (d\lambda)^2\left(1 - \frac{v^2}{c^2}\right) \,.$$

Using the chain rule again, the above simplifies to

$$dS^2 = -c^2\cancel{\left(\frac{dt}{d\tau}\right)^2} (d\lambda)^2 \left(\frac{d\tau}{d\lambda}\right)^2 \cancel{\gamma^{-2}} = -c^2\left(\frac{d\tau}{d\lambda}\right)^2 (d\lambda)^2$$

From here, we see that λ must be an affine parameter in order to recover $dS^2 = -c^2 d\tau^2$. That is, we must have $\tau = \alpha\lambda + \alpha_0$ such that

$$dS^2 = -c^2(\alpha)^2\frac{d\tau^2}{\alpha^2} = -c^2 d\tau^2 \,.$$

Problem 2

Consider a worldline $x^\mu(\lambda)$ that goes across the event $E = (0, 0, 1m, 0)$, parameterized as

$$x^0(\lambda) = \lambda \qquad\qquad x^1(\lambda) = \lambda/2$$

$$x^2(\lambda) = (1m)\sqrt{1 - (\lambda/1m)^2} \qquad\qquad x^3(\lambda) = 0\,.$$

(i) What value of λ corresponds to event E? (ii) Compute the components of the *tangent vector* $dx^\mu/d\lambda$ to the worldline $x^\mu(\lambda)$ at event E. (iii) Compute the velocity dx^j/dt of a particle that goes along the worldline $x^\mu(\lambda)$ at event E.

Solution 2

$$\lambda = 0 \qquad\qquad \left.\frac{dx^\mu}{d\lambda}\right|_E = (1m, 0.5m, 0, 0) \qquad\qquad \left.\frac{dx^j}{dt}\right|_E = \left(\frac{c}{2}, 0, 0\right)$$

Problem 3

Suppose the worldline in the previous problem is expressed in a frame that is boosted with speed v along the x^1 direction. Compute the nonzero components of the tangent vector $(dx^\mu)'/d\lambda$ in the primed reference frame.

Solution 3

$$\frac{(dx^0)'}{d\lambda} = \frac{d}{d\lambda}\left(\frac{\lambda\,(1 - v/2c)}{\sqrt{1 - v^2/c^2}}\right) = \gamma\left(1 - \frac{v}{2c}\right) \qquad\qquad \frac{(dx^1)'}{d\lambda} = \gamma\left(\frac{1}{2} - \frac{v}{c}\right)$$

Problem 4

Consider a worldline described by

$$q^0(\lambda) = a\lambda \qquad\qquad q^1(\lambda) = \frac{a}{2}\cos\lambda$$

$$q^2(\lambda) = \frac{a}{2}\sin\lambda \qquad\qquad q^3(\lambda) = 0\,,$$

where a is a real constant and λ is a parameter of motion. Compute the proper time between $\lambda = 0$ and $\lambda = 2\pi$.

Solution 4

$$\tau = \int \frac{1}{c}\sqrt{-\eta_{\mu\nu}dq^\mu dq^\nu} = \frac{a}{c}\int_0^{2\pi}\sqrt{1 - \frac{1}{4}}\,d\lambda = \frac{a\sqrt{3}\pi}{c}$$

7 Relativistic Collisions

We've seen that energy and mass become interchangeable when particles collide or decay. It takes little to conclude that conservation of mass and conservation of energy as separate quantities is an illusion of Newtonian mechanics, where the 'true' conservation law takes mass and energy as a pair. Meanwhile, momentum is always conserved in the 'ordinary' sense. Using four-vector notation, relativistic collisions and decays are fully characterized by the momentum four-vector, which is conserved across the event:

$$P^\mu_{\text{initial}} = P^\mu_{\text{final}}$$

7.1 Compton Scattering

Consider a particle of mass m at rest in a laboratory reference frame. A prepared photon of wavelength λ_i transfers energy and momentum to the particle if a collision takes place. As a result, a new photon with increased wavelength λ_f emerges at angle θ with respect to the incident axis, and the particle carries the photon's missing momentum away at angle ϕ with respect to the incident axis at speed v.

Setting up four-vectors for the 'initial' and 'final' states of the photon and the particle, we have

$$\begin{aligned}
\gamma_i &= \left(E_i/c, \hbar\vec{k}_i\right) && \text{Initial photon state} \\
M &= (mc, 0) && \text{Initial particle state} \\
\gamma_f &= \left(E_f/c, \hbar\vec{k}_f\right) && \text{Final photon state} \\
P &= (E_p/c, \vec{p}) && \text{Final particle state ,}
\end{aligned}$$

where the conservation equation $P^\mu_{\text{initial}} = P^\mu_{\text{final}}$ tells us

$$\gamma_i + M = \gamma_f + P \; .$$

Compton Scattering Formula

A common equation from modern physics can be derived by isolating P, squaring both sides, and simplifying:

$$(\gamma_i - \gamma_f + M)^2 = P^2$$

$$\gamma_i^2 + \gamma_f^2 + M^2 - 2\gamma_i\gamma_f + 2\gamma_i M - 2\gamma_f M = P^2$$

$$\cancel{\gamma_i^2} + \cancel{\gamma_f^2} - \cancel{m^2c^2} - 2\left(\frac{E_i}{c}, \hbar\vec{k}_i\right)\left(\frac{E_f}{c}, \hbar\vec{k}_f\right) + 2\,(mc, 0)\left(\frac{E_i}{c}, \hbar\vec{k}_i\right) - 2\,(mc, 0)\left(\frac{E_f}{c}, \hbar\vec{k}_f\right) = -\cancel{m^2c^2}$$

$$-\left(-\frac{E_iE_f}{c^2} + \hbar^2\vec{k}_i\cdot\vec{k}_f\right) + (-E_i m) - (-E_f m) = 0$$

Note note that

$$\hbar^2\vec{k}_i\cdot\vec{k}_f = \hbar^2 k_i k_f \cos\theta = \frac{E_iE_f}{c^2}\cos\theta \; ,$$

which brings about the famed *Compton scattering formula*:

$$\frac{1}{E_f} - \frac{1}{E_i} = \frac{1-\cos\theta}{mc^2}$$

In terms of wavelength, the Compton scattering formula reads

$$\lambda_f - \lambda_i = \frac{h}{mc}\left(1-\cos\theta\right) \ ,$$

where h/mc is known as the *Compton wavelength* of the particle.

Energy Conservation

Starting again with $\gamma_i + M = \gamma_f + P$, solve for $\gamma_i - \gamma_f$ and square both sides:

$$\begin{aligned}
\left(\gamma_i - \gamma_f\right)^2 &= \left(P - M\right)^2 \\
\cancel{\gamma_i^2} + \cancel{\gamma_f^2} - 2\gamma_i\gamma_f &= -2m^2c^2 - 2\left(mc, 0\right)\left(\frac{E_p}{c}, \vec{p}\right) \\
2m\left(E_i - E_f\right) &= -2m^2c^2 - 2\left(-E_p m\right) \\
E_i + mc^2 &= E_f + E_p
\end{aligned}$$

This is nothing more than the energy conservation statement, which could have been more easily derived by multiplying the conservation equation through by the four-vector $M = (mc, 0)$.

The total energy carried away by the particle is $E_p = E_i + mc^2 - E_f$. The kinetic energy subtracts away the rest energy namely $T = E_p - mc^2$, indicating $T = E_i - E_f$, as we'd expect.

Momentum Conservation

Squaring the conservation equation without any rearranging gives a useful identity:

$$\begin{aligned}
\left(\gamma_i + M\right)^2 &= \left(\gamma_f + P\right)^2 \\
\cancel{\gamma_i^2} + 2\left(mc, 0\right)\left(\frac{E_i}{c}, \hbar\vec{k}_i\right) - \cancel{m^2c^2} &= \cancel{\gamma_f^2} + 2\left(\frac{E_f}{c}, \hbar\vec{k}_f\right)\left(\frac{E_p}{c}, \vec{p}\right) - \cancel{m^2c^2} \\
\gamma_i M &= \gamma_f P
\end{aligned}$$

Next, solve the conservation equation for M and square both sides:

$$\begin{aligned}
\left(\gamma_f - \gamma_i + P\right)^2 &= M^2 \\
\cancel{\gamma_f^2} + \cancel{\gamma_i^2} - \cancel{m^2c^2} - 2\gamma_i\gamma_f + 2\gamma_f P - 2\gamma_i P &= -\cancel{m^2c^2} \\
-\gamma_i\gamma_f + \gamma_i M - \gamma_i P &= 0
\end{aligned}$$

In th above, the identity $\gamma_i M = \gamma_f P$ as been used. Simplifying further:

$$\begin{aligned}
\cancel{E_i}E_f\left(1-\cos\theta\right) - \cancel{E_i}mc^2 - \left(-\cancel{E_i}E_p + \cancel{E_i}pc\cos\phi\right) &= 0 \\
E_f - mc^2 + E_p &= E_f\cos\theta + pc\cos\phi \\
E_i &= E_f\cos\theta + pc\cos\phi
\end{aligned}$$

Divide the result by c to reiterate the conservation of momentum statement along the incident axis:

$$\hbar k_i = \hbar k_f \cos\theta + p\cos\phi$$

It takes little imagination to see that the perpendicular component of momentum conservation is

$$0 = \hbar k_f \sin\theta - p\sin\phi \ ,$$

such that the whole vector equation is

$$\hbar\vec{k}_i = \hbar\vec{k}_f + \vec{p} \ .$$

Note that this result is more easily attained by multiplying the conservation equation through by a dimensionless 'unit' four-vector $(0, 1)$.

7.2 Pion Production

Consider a single 'target' proton of mass m_p that is stationary with respect to a laboratory reference frame. An incoming photon of energy E_γ collides with the proton and is absorbed. A π^0 particle is created in the collision and carries away momentum $\vec{p}_\pi$ at angle ϕ with respect to the incident axis. The proton also recoils with momentum $\vec{p}$ at an angle θ.

In four-vector notation, this collision, known as 'photo-π^0 production', may be written as

$$\gamma + M = P + \pi \ ,$$

where the individual four-vectors are:

$\gamma = \left(E_\gamma/c, \hbar\vec{k}_\gamma\right)$	Initial photon state
$M = (m_p c, 0)$	Initial proton state
$\pi = (E_\pi/c, \vec{p}_\pi)$	Final pion state
$P = (E_p/c, \vec{p}_p)$	Final proton state

Energy and Momentum Conservation

The four-vectors written above imply conservation of energy and conservation of momentum statements:

$$E_\gamma + m_p c^2 = E_p + E_\pi \qquad \hbar k_\gamma = p_p\cos\theta + p_\pi\cos\phi \qquad 0 = p_p\sin\theta - p_\pi\sin\phi$$

Pion Momentum

The momentum p_π carried off my the π^0 particle can be isolated by squaring both sides of $\gamma - \pi = P - M$:

$$(\gamma - \pi)^2 = (P - M)^2$$

$$\cancel{\gamma^2} + \pi^2 - 2\gamma\pi = P^2 + M^2 - 2PM$$

$$\frac{m\pi^2 c^2}{2} - \left(\frac{E_\gamma E_\pi}{c^2} + \hbar k_\gamma p_\pi \cos\theta\right) = -m_p^2 c^2 + m_p\left(E_\gamma + m_p c^2 - E_\pi\right)$$

Let

$$A = \frac{m_\pi^2 c^2}{2} + m_p E_\gamma \qquad B = \hbar k_\gamma \cos\theta \qquad C = \frac{E_\gamma}{c^2} + m_p$$

to write

$$\begin{aligned} A + p_\pi B &= E_\pi C \\ A^2 + 2ABp_\pi + p_\pi^2 B^2 &= \left(p_\pi^2 c^2 + m_\pi^2 c^4\right) C^2 \\ p_\pi^2 \left(B^2 - C^2 c^2\right) + p_\pi \left(2B\right) + \left(A^2 - C^2 m_\pi^2 c^4\right) &= 0 \\ X p_\pi^2 + Y p_\pi + Z &= 0\,, \end{aligned}$$

where

$$X = B^2 - C^2 c^2 \qquad Y = 2B \qquad Z = A^2 - C^2 m_\pi^2 c^4\,.$$

Finally, p_π is delivered by he quadratic formula:

$$p_\pi = -\frac{Y}{2X} \pm \frac{\sqrt{Y^2 - 4XZ}}{2X}$$

7.3 Antiproton Production

At low speeds, two colliding protons will exchange energy and momentum and go off in new directions. At very high speeds however, there is a chance that the collision of two protons results in *four* products: three ordinary protons, and one antiproton. We track this reaction with four-vectors

$$P_1 + P_2 = P_A + P_B + P_C + \bar{P}\,,$$

where the indices 1, 2, track the initial particles, and A-C are the products. The antiproton is indicated by a bar above the P-symbol. Assume that P_2 is a stationary 'target' proton with respect to a laboratory reference, where P_1 initially carries energy E_0 and momentum $\vec{p}_0$.

The question on hand is: what is the minimum energy E_0 required for the reaction to take place? To proceed we will exploit the invariance of squared four-vectors by evaluating the above equation in *two* different reference frames. In the laboratory reference frame, the left-side four-vectors are:

$$\begin{aligned} P_1 &= \left(E_0/c, \vec{p}_0\right) && \text{Incident proton} \\ P_2 &= \left(m_p c, 0\right) && \text{Target proton} \end{aligned}$$

So the product $\left(P_1 + P_2\right)^2$ reads

$$\begin{aligned} \left(P_1 + P_2\right)^2 &= P_1^2 + P_2^2 + 2P_1 P_2 \\ &= -2m_p^2 c^2 - 2E_0 m_p\,. \end{aligned}$$

In the laboratory frame, the products carry away the original momentum, and it would be difficult to discern a minimal value for E_0 without additional insight. Thus, we also examine the collision in the center-of-mass frame of the products. For minimal E_0, each product emerges *at rest* in this frame. Each product has an identical four-vector $P = (m_p c, 0)$, and the square of the conservation equation gives

$$\left(P_A + P_B + P_C + \bar{P}\right)^2 = -4^2 m_p^2 c^2 \,.$$

Since the square of the total momentum four-vector is an invariant, we equate the two above results to solve for E_0:

$$\begin{aligned} \left(P_1 + P_2\right)^2 &= \left(P_A + P_B + P_C + \bar{P}\right)^2 \\ -2m_p^2 c^2 - 2E_0 m_p &= -16 m_p^2 c^2 \\ E_0 &= 7 m_p c^2 \end{aligned}$$

Evidently, the total incident proton energy must be $7m_p c^2$, where subtracting the rest mass we find $T_0 = 6m_p c^2$.

Of course, a more efficient laboratory would have the initial protons have equal and opposite velocities, such that the collision takes place in the center-of-mass frame. In such a case, the above becomes

$$\begin{aligned} \left(P_1 + P_2\right)^2 &= \left(P_A + P_B + P_C + \bar{P}\right)^2 \\ -2m_p^2 c^2 - 2E_0^2/c^2 - 2\left(E_0^2/c^2 - m_p^2 c^2\right) &= -16 m_p^2 c^2 \\ E_0 &= 2 m_p c^2 \,. \end{aligned}$$

Each proton needs to carry the same E_0 initially, so the apparatus needs to supply $T_0 = 2E_0 - 2m_p c^2 = 2m_p c^2$.

8 Variational Methods

8.1 Euler-Lagrange Equation

Classical mechanics can be re-derived using the calculus of variations and the principle of least action. Begin by considering any function $F(q, \dot{q}, t)$, depending on position, velocity ($dq/dt = \dot{q}$), and a time parameter.

Defining the action

$$S = \int_{t_i}^{t_f} F \, dt \, ,$$

it follows that variations in S are given by

$$\delta S = \int_{t_i}^{t_f} \left(\frac{\partial F}{\partial q} \delta q + \frac{\partial F}{\partial \dot{q}} \delta \dot{q} \right) dt$$

By constraining δq and $\delta \dot{q}$ to be zero on the boundaries t_i, t_f, the above can be integrated by parts to get

$$\delta S = \int_{t_i}^{t_f} \left(\frac{\partial F}{\partial q} - \frac{d}{dt} \left(\frac{\partial F}{\partial \dot{q}} \right) \right) dt \, ,$$

where taking the limit $\delta S \to 0$ tells us the parenthesized quantity is also zero, delivering the Euler-Lagrange equation

$$\frac{\partial F}{\partial q} - \frac{d}{dt} \left(\frac{\partial F}{\partial \dot{q}} \right) = 0$$

which needless to mention readily generalizes to more dimensions.

8.2 Classical Lagrangian Mechanics

By setting the general function F to equal the kinetic energy minus the potential energy of a particle of mass m, we write the *classical Lagrangian*:

$$L = \frac{1}{2} m \dot{q}^2 - U(q)$$

Immediately, the Euler-Lagrange equation produces Newton's second law:

$$\begin{aligned} \frac{\partial F}{\partial q} &= \frac{d}{dt} \left(\frac{\partial F}{\partial \dot{q}} \right) \\ -\frac{\partial U}{\partial q} &= \frac{d}{dt} \frac{\partial}{\partial \dot{q}} \left(\frac{1}{2} m \dot{q}^2 \right) = m \frac{d}{dt} (\dot{q}) = m \frac{d^2 q}{dt^2} \, . \end{aligned}$$

An expression for the momentum $p = m\dot{q}$ is a partial derivative of the Lagrangian:

$$p = m\dot{q} = \frac{\partial L}{\partial \dot{q}}$$

8.3 Relativistic Lagrangian

To derive a version of Lagrangian mechanics that accommodates special relativity, it should help to redefine the action as an integral over an invariant quantity, namely the square of the spacetime interval, also known as proper time τ. Anticipating the dimensionality of the result (energy units), we begin with

$$S = -mc^2 \int_{\tau_i}^{\tau_f} d\tau \ .$$

Substituting $\gamma = dt/d\tau$, we jump back to a non-rest-frame casting

$$S = -mc^2 \int_{t_i}^{t_f} \sqrt{1 - \dot{q}^2/c^2} \, dt$$

where $\dot{q}$ is the relative frame speed. Inserting a potential term $-U(q)$ as done in the classical case, the action to minimize is

$$S = \int_{t_i}^{t_f} \left(-mc^2 \sqrt{1 - \dot{q}^2/c^2} - U(q) \right) dt \ ,$$

where comparison to $S = \int L \, dt$ tells us the relativistic Lagrangian is

$$L = -mc^2 \sqrt{1 - \dot{q}^2/c^2} - U(q) \ .$$

At low speeds, the kinetic term expands into the classical kinetic energy plus correction terms, with the rest energy subtracted away:

$$-mc^2 \sqrt{1 - \dot{q}^2/c^2} \approx -mc^2 + \frac{1}{2} m \dot{q}^2 + \cdots$$

The same result is attained by integrating the formula for the momentum, but slipping in the relativistic momentum instead of the classical form:

$$p = \gamma m \dot{q} = \frac{\partial L}{\partial \dot{q}} \qquad \rightarrow \qquad m \int \frac{\dot{q} \, d\dot{q}}{\sqrt{1 - \dot{q}^2/c^2}} = \int dL$$

The potential $U(q)$ arises as the integration constant.

8.4 Equations of Motion

In three dimensions, the Lagrangian of a particle having mass m and velocity vector $\vec{v} = v \, \hat{v}$ is

$$L = -mc^2 \sqrt{1 - v^2/c^2} - U(\vec{r}) \ ,$$

obeying the Euler-Lagrange equation

$$\frac{\partial L}{\partial \vec{r}} = \frac{d}{dt} \frac{\partial L}{\partial \vec{v}} \ ,$$

where $\vec{r}$ is the time-dependent position vector, and the link $\dot{\vec{r}} = \vec{v}$ has been assumed. Calculating this out, we discover

$$-\frac{\partial U}{\partial \vec{r}} = m\frac{d}{dt}\left(\frac{\vec{v}}{\sqrt{1 - v^2/c^2}}\right) = F\left(\vec{r}\right) \ ,$$

matching our previous form for the force vector. In terms of L, the momentum vector is

$$\vec{p} = \frac{\partial L}{\partial \vec{v}} = \frac{m\vec{v}}{\sqrt{1 - v^2/c^2}} \ ,$$

which gives a simultaneous expression for the velocity:

$$\vec{v} = \frac{\vec{p}/m}{\sqrt{1 - \left(p/mc\right)^2}}$$

To write the total energy E of the particle, we deploy Hamiltonian formalism by writing

$$E = \vec{v} \cdot \frac{\partial L}{\partial \vec{v}} - L \ ,$$

where simplifying, we land at a familiar expression for the energy, namely

$$E = \frac{mc^2}{\sqrt{1 - v^2/c^2}} + U\left(\vec{r}\right) \ .$$

To write the Hamiltonian H of the particle, we need to express the energy purely in terms of $\vec{r}$ and $\vec{p}$. Doing so, we have

$$H = \frac{mc^2}{\sqrt{1 - \left(p/mc\right)^2}} + U\left(\vec{r}\right) \ .$$

Borrowing again from Hamiltonian formalism, it's straightforward to show that the equations of motion of the particle are given by

$$\dot{\vec{r}} = \frac{\partial H}{\partial \vec{p}} = \frac{\vec{p}/m}{\sqrt{1 - \left(p/mc\right)^2}}$$

$$\dot{\vec{p}} = -\frac{\partial H}{\partial \vec{r}} = -\frac{\partial U}{\partial \vec{r}} = \vec{F}\left(\vec{r}\right)$$