Summary of SI2371 Special Relativity

Yashar Honarmandi
yasharh@kth.se

November 9, 2020

Abstract

Contents

1 Basic Concepts

What is Special Relativity? Special relativity is not just a way to do mechanics. It is a way to do theoretical physics involving a new way of thinking about space and time.

Fundamental Postulates of Special Relativity Einstein built the theory of special relativity based on the following postulates:

- The laws of physics are the same inertial frames.
- The speed of light is the same in all inertial frames.

Inertial Frames An inertial frame of reference is a kind of frame of reference, meaning that it is a certain way to consider time and space. In an inertial frame free particles move in straight lines with constant velocity. We will only be talking about such frames.

It can be shown that all inertial frames move uniformly with constant velocity relative to each other. When considering two inertial frames, we will usually choose the coordinate systems in each frame such that the axes are parallel to each other and one axis is parallel to the relative velocity.

The Galilei Group The Galilei group is the group of transformations of a physical system that do not change the fundamental physics of the system. It is composed of

- Rotations.
- Translations.
- Galilei boosts, to be described.

Galilei Boosts Consider two frames of reference S and S' moving with relative velocity v in the x-direction. Galilei boosts are of the form $x' = x - vt$.

Using such transformations, velocities transform by simply adding or subtracting $v\mathbf{e}_x$, meaning that these transforms do not leave the speed of light invariant. When constructing a group of transformations that leave physics invariant under the laws of special relativity, we cannot construct it using Galilei boosts.

Simultaneity To observe a consequence of special relativity, consider a light source at the origin in its rest frame S sending a pulse of light towards two detectors in $x = \pm x_0$. In S the light reaches each detector at the same time.

Consider now a frame S' moving in with velocity $v\mathbf{e}_x$. According to an observer in this frame, one detector will approach the light source and the other recede from it, making the light reach one detector before the other. This makes it obvious that the classical concept of absolute time cannot persist in special relativity.

Extending Inertial Frames The fact that simultaneity does not exist in special relativity forces us to assign a measure of time to every point in space in a particular frame. This time is synchronized with respect to an observer at rest in the frame.

One way to do this is to emit a pulse of light from the origin, setting the time at any point to $t_0 + \frac{r}{c}$, where r is the distance of the point from the origin.

Light Clocks A light clock is a device that can be used to measure time. It consists of a light source and a mirror separated by a distance r. The time taken from a light pulse being emitted to it returning to the source is

$$\Delta t = \frac{2L}{c}.$$

This can be used to ”standardize” a measurement of time.

Time Dilation From Light Clocks Consider two frames of reference S and S' moving with relative velocity v in the x-direction, and suppose that there exists a light clock at rest in S oriented orthogonally to the x-direction. The time taken in S' for a pulse of light to hit the mirror and return is as before. In S, the constancy of light requires

$$\sqrt{(2L)^2 + (v\Delta t)^2} = c\Delta t,$$

with solution

$$\Delta t = \gamma \frac{2L}{c},$$

where we introduce the Lorentz factor

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

Length Contraction From Light Clocks Consider the same setup as that above, but suppose instead that the light clock is oriented along the x-direction. In S' no difference is observed. In S the following sequence takes place:

1. At time 0 the pulse is emitted.
2. At time Δt_1 the pulse hits the mirror.
3. At time $\Delta t_1 + \Delta t_2$ the pulse returns.

The constancy of the speed of light implies

$$c\Delta t_1 = L + v\Delta t_1,\ -c\Delta t_2 = -L + v\Delta t_2,$$

allowing us to solve for them. The total time elapsed is

$$\Delta t = L\left(\frac{1}{c+v} + \frac{1}{c-v}\right) = \frac{2cL}{c^2 - v^2}.$$

Note that the distance L is measured in S, but is not necessarily equal to the length in S' - in fact, as time dilates we must have

$$\frac{2L_0}{c} = \gamma \frac{2cL}{c^2 - v^2},$$

and thus

$$L = L_0 \frac{c^2 - v^2}{\gamma c^2} = \frac{L_0}{\gamma}.$$

Do the lengths in other directions contract as well? The answer is no. To understand this, consider the following thought experiment: Suppose you throw a ball through a slit of length L_0 (measured in its rest frame). In the rest frame of the ball, this length may contract, expand or be unaltered. If the lengths were to contract or expand, observers in the rest frames of the slit and the ball would disagree on whether the ball passed through it or not. This would seem to violate some law of physics, which cannot be allowed. Hence perpendicular lengths are not transformed by Lorentz boosts.

A Brief Note on Experiments I here briefly mention a few experiments which match with the predictions of special relativity that we have seen.

- Muons are particles with extremely short half-lives that travel at high speeds. Atmospheric muons should not have a sufficiently long half-life to be found on the surface of the Earth according to classical mechanics. In special relativity, we may explain this in two ways: According to an observer on Earth, the muons' clocks tick slower, meaning that they have the time to reach Earth. According to the muons, the distance to the Earth is contracted, meaning that they travel fast enough to make it to Earth before decaying.
- Very fast airplanes have been sent in different directions, each with their own atomic clock. After travelling a distance comparable to the circumference of the Earth, they meet and the times of the clocks are read off. The difference between the clocks was on the order of magnitude of nanoseconds.

Deriving the Lorentz Transformation The Lorentz transformation is the transformation that takes us from one inertial frame to another which is boosted (has a velocity) relative to the first. We will now derive it using the following assumptions:

- The transformation is linear, so as to not cause non-accelerating motion in one frame to be accelerating in another.
- Perpendicular lengths do not enter into the transformation and are themselves left unaltered.
- Inverting the transformation corresponds to swapping coordinates in different frames and changing the sign of the relative velocity.
- The speed of light must be preserved by the transformation.

To derive it, consider two frames of reference which coincide at $t = 0$ and where the primed frame moves with a speed v in the x-direction relative to the other (this will be standard notation from now on). The transformation is now of the form

$$x' = Ax + Bt,\ t' = Cx + Dt.$$

The point $x' = 0$ is obviously described by $x = vt$, which implies $\frac{B}{A} = -v$ and

$$x' = A(x - vt).$$

To impose the requirement that the speed of light be constant, we consider a light pulse emitted at $t = 0$ from the origin. Both the Lorentz transform and its inverse should transform between ct and ct', yielding

$$ct' = A(c - v)t,\ ct = A(c + v)t'.$$

Thus we obtain

$$c^2tt' = A^2(c^2 - v^2)tt',$$

implying

$$A^2 = \gamma^2.$$

This constant must needs be positive, hence we conclude $A = \gamma$.

To obtain the remaining coefficient, we compose the Lorentz transform with its inverse. The inverse spatial transform is

$$x = \gamma(x' + vt'),$$

which expressed in terms of the time transformation is

$$x = \gamma(x' + vCx + vDt).$$

Solving it yields

$$x' = \left(\frac{1}{\gamma} - vC\right)x - vDt.$$

This must simply be the original transform, hence

$$\frac{1}{\gamma} - vC = \gamma,\ -vD = -v\gamma.$$

The solutions to this are

$$D = \gamma,\ C = \frac{1}{v}\frac{1 - \gamma^2}{\gamma} = \frac{1}{v}\frac{1 - \frac{1}{1-\beta^2}}{\gamma} = -\frac{1}{v}\frac{\frac{\beta^2}{1-\beta^2}}{\gamma} = -\frac{\beta\gamma}{c}.$$

In conclusion, the Lorentz transform is

$$x' = \gamma(x - vt),\ t' = \gamma\left(t - \frac{\beta}{c}x\right).$$

Note that it approaches Galilei transforms in the limit of $\beta \to 0$, for which $\gamma \to 1$. This is why physicists managed to develop classical mechanics without realizing the existence of relativistic phenomena.

Length Contraction and Time Dilation From the Lorentz Transform The phenomena of length contraction and time direction previously emerged as a consequence of us considering some particular geometry, but we may show that they emerge more naturally from the Lorentz transform itself.

Consider two events happening at $x' = \alpha$ within the time interval $\Delta t'$ in S'. We then have

$$\alpha = \gamma(x_1 - vt_1),\ t_1' = \gamma\left(t_1 - \frac{\beta}{c}x_1\right),\ \alpha = \gamma(x_2 - vt_2),\ t_2' = \gamma\left(t_2 - \frac{\beta}{c}x_2\right),$$

hence

$$\Delta x = v\Delta t,\ \Delta t' = \gamma\left(\Delta t - \frac{\beta}{c}\Delta x\right) = \gamma\left(\Delta t - \beta^2\Delta t\right) = \frac{1}{\gamma}\Delta t,$$

and time dilates in the non-rest frame of the event.

Next, consider two events which happen at time τ at a distance $\Delta x'$ in S'. We have

$$x_1' = \gamma(x_1 - vt_1),\ \tau = \gamma\left(t_1 - \frac{\beta}{c}x_1\right),\ x_2' = \gamma(x_2 - vt_2),\ \tau = \gamma\left(t_2 - \frac{\beta}{c}x_2\right),$$

hence

$$\Delta t = \frac{\beta}{c}\Delta x,\ \Delta x' = \gamma(\Delta x - v\Delta t) = \gamma\Delta x(1 - \beta^2) = \frac{\Delta x}{\gamma},$$

and lengths contract in the non-rest frame of the event.

New Variables Let us introduce $x^0 = ct$. The Lorentz transform is thus

$$x' = \gamma(x - \beta x^0),\ (x^0)' = \gamma\left(x^0 - \beta x\right).$$

This has a more obvious symmetry with respect to the involved variables, and this notation is also more enticing for physics in general.

Lorentz Transforms and Hyperbolic Functions As $\gamma^2(1-\beta^2) = 1$, we may define $\gamma = \cosh(\theta),\ \gamma\beta = \sinh(\theta)$, where θ is the so-called rapidity, to write the Lorentz transform as

$$x' = \cosh(\theta)x - \sinh(\theta)x^0,\ (x^0)' = \cosh(\theta)x^0 - \sinh(\theta)x.$$

We may also identify the rapidity according to $\beta = \tanh(\theta)$.

Transforming Velocities Using the chain rule, velocities along the x-axis are transformed according to

$$u' = \frac{\mathrm{d}x'}{\mathrm{d}t'} = \frac{\frac{\mathrm{d}x'}{\mathrm{d}t}}{\frac{\mathrm{d}t'}{\mathrm{d}t}} = \frac{\gamma(u - v)}{\gamma(1 - \frac{v}{c^2}u)} = \frac{u - v}{1 - \beta\frac{u}{c}},$$

where we have used the Lorentz transform as well as the fact that $u = \frac{\mathrm{d}x}{\mathrm{d}t}$ to obtain this expression.

The velocities in other directions, for instance the y-direction, transform according to

$$u_y' = \frac{\mathrm{d}y'}{\mathrm{d}t'} = \frac{\frac{\mathrm{d}y}{\mathrm{d}t}}{\frac{\mathrm{d}t'}{\mathrm{d}t}} = \frac{1}{1 - \beta\frac{u}{c}}u_y.$$

Transforming Accelerations We can study other quantities in a similar fashion. For instance, an acceleration parallel to the relative velocity transforms according to

$$\begin{aligned}
a' &= \frac{\mathrm{d}^2x'}{\mathrm{d}(t')^2} \\
&= \frac{\frac{\mathrm{d}}{\mathrm{d}t}\frac{\mathrm{d}x'}{\mathrm{d}t'}}{\frac{\mathrm{d}t'}{\mathrm{d}t}} = \frac{1}{\gamma\left(1 - \beta\frac{u}{c}\right)}\frac{\mathrm{d}}{\mathrm{d}t}\frac{u - v}{1 - \beta\frac{u}{c}} \\
&= \frac{1}{\gamma\left(1 - \beta\frac{u}{c}\right)}\frac{a\left(1 - \beta\frac{u}{c}\right) + (u - v)\beta\frac{1}{c}a}{\left(1 - \beta\frac{u}{c}\right)^2} \\
&= \frac{1}{\gamma\left(1 - \beta\frac{u}{c}\right)^3}(1 - \beta^2)a \\
&= \frac{1}{\gamma^3\left(1 - \beta\frac{u}{c}\right)^3}a.
\end{aligned}$$

The Garage Paradox Consider a garage of length L_0 and a car of length $l_0 > L_0$. If the car drives into the garage at a high speed, the garage will contract in the car's frame, making sure that it does not fit, while the car will contract in the garage's frame, making it (possibly) fit. How is this possible?

The resolution comes in the form of simultaneity. We consider two events: the front hitting the wall and the rear entering the garage. Assume that these are simultaneous in the rest frame S of the garage. Applying the Lorentz transformation formula yields

$$\Delta x' = \gamma \Delta x,$$

corresponding to the length of the car being contracted in S. Hence it fits in S. Furthermore, the time between the two events in S' is

$$\Delta t' = -\frac{\beta}{\gamma c}\Delta x,$$

meaning that the rear of the car enters the garage before the front hits the wall and matching the prediction in S.

The Twin Paradox Consider a pair of twins, where one twin remains on Earth and the other travels into space at a high velocity, eventually returning to Earth. According to the twin on Earth, time will pass more quickly for him, while according to the twin in space, time will pass more quickly for him. Which twin is older when they meet again?

The answer is that the travelling twin must change inertial system, removing the assumed symmetry of the scenario.

Minkowski Diagrams and Minkowski Space A Minkowski diagram is a diagram of a scenario with ct on the vertical axis and space, usually represented by x, on the other. Trajectories of particles, called world lines, must have a slope greater than 1 in this representation.

In this space we define certain distinct regions. First there is the light cone, which consists of all points such that $(ct)^2 - \sum\limits_i (x^i)^2 = 0$. Next there is the future and past, which consists of all points between the light cone and the ct-axis with positive and negative time components respectively. Finally there is the elsewhere, which is the rest.

Lorentz Transformations in Minkowski Diagrams A Lorentz transformation in a Minkowski diagram corresponds to making the (transformed) space and time axes move closer together, approaching the identity line at equal rates.

2 4-Vector Formalism

4-vector formalism makes more explicit use of invariance relations that have previously been identified, most importantly the invariance of c, or so-called Lorentz invariance.

The Lorentz Group Consider a light pulse sent out from the origin at $t = 0$. The wavefront in the rest frame of the emitter satisfies $r^2 = (ct)^2$. Next, for any frame in the standard configuration we must also have $(r')^2 = (ct')^2$, implying the invariance of the quantity

$$(ct)^2 - r^2$$

under any transformation that preserves the laws of physics. The Lorentz group is defined as the group of linear transforms that preserves the above quantity (linearity is to preserve important physical properties such as isotropy of space).

The Lorentz Boost in Matrix Form The Lorentz boost may now be written as

$$\Lambda = \begin{bmatrix} A & 0 \\ 0 & 1 \end{bmatrix},\ A = \begin{bmatrix} \gamma & -\beta\gamma \\ -\beta\gamma & \gamma \end{bmatrix}$$

such that $(x')^\mu = \Lambda^\mu{}_\nu x^\nu$, with Einstein summation from 0 to 3. This is also the general notation for Lorentz transformations, where Λ may now be taken to be any element in the Lorentz group.

The Lorentz Group The Lorentz group is the group of all linear transformations that preserve the laws of physics. It consists of rotations and Lorentz boosts.

The Poincare Group The Poincare group is the group of all transformations that preserve the laws of physics. It consists of the Lorentz group as well as translations.

The Spacetime Interval The Lorentz transform transforms infinitesimal intervals in spacetime. We would now like to define the spacetime interval

$$\mathrm{d}s^2 = (\mathrm{d}x^0)^2 - \sum_i (\mathrm{d}x^i)^2$$

under Poincare transformations.

Clearly the spacetime interval is preserved by space-preserving transformations that do not alter time, hence we only need to consider Lorentz boosts. We have

$$\begin{aligned}(\mathrm{d}(x')^0)^2 - \sum_i (\mathrm{d}(x')^i)^2 &= (\gamma\,\mathrm{d}x^0 - \beta\gamma\,\mathrm{d}x^1)^2 - (\gamma\,\mathrm{d}x^1 - \beta\gamma\,\mathrm{d}x^0)^2 - (\mathrm{d}x^2)^2 - (\mathrm{d}x^3)^2 \\ &= \gamma^2(1-\beta^2)(\mathrm{d}x^0)^2 - \gamma^2(1-\beta^2)(\mathrm{d}x^1)^2 - (\mathrm{d}x^2)^2 - (\mathrm{d}x^3)^2 \\ &= (\mathrm{d}x^0)^2 - (\mathrm{d}x^1)^2 - (\mathrm{d}x^2)^2 - (\mathrm{d}x^3)^2,\end{aligned}$$

hence the spacetime interval is preserved under the Poincare group. This is one of its defining traits.

Tensors A tensor is something that transforms according to the familiar transformation rules under Lorentz transformations. We recall that the transformation coefficients for contravariant indices are $\Lambda^\mu{}_\nu = \partial_\nu (x')^\mu$ and the coefficients for covariant indices are $\Lambda_\nu{}^\mu = \partial'_\nu x^\mu$. A notation that will be introduced is that transformed tensor components are denoted with primed indices, rather than the symbol of the tensor having the prime. Using this notation we have

$$\Lambda_{\mu'}{}^\mu \Lambda^{\mu'}{}_\nu = \partial_{\mu'} x^\mu \partial_\nu x^{\mu'} = \partial_\nu x^\mu = \delta^\mu_\nu.$$

Similarly we may obtain $\Lambda^{\mu'}{}_\mu \Lambda_{\nu'}{}^\mu = \delta^{\mu'}_{\nu'}$.

The transformation rules for vectors are

$$A^{\mu'} = \Lambda^{\mu'}{}_\mu A^\mu,\ A_{\mu'} = \Lambda_{\mu'}{}^\mu A_\mu,$$

identifying the general transformation coefficients for covariant and contravariant indices.

The Metric Tensor The metric tensor g is defined by

$$\mathrm{d}s^2 = g_{\mu\nu}\,\mathrm{d}x^\mu\,\mathrm{d}x^\nu\,.$$

By definition it is symmetric.

Clearly in special relativity with Cartesian coordinates we have $g_{00} = 1,\ g_{ii} = -1$ and all other components are zero. In Cartesian coordinates we also have $g_{\mu\nu} = g^{\mu\nu}$.

In special relativity we take the metric to define the inner product between vectors, which implies that it can be used to raise and lower indices.

Classification of 4-Vectors 4-vectors are time-like if $V^2 > 0$, space-like if $V^2 < 0$ and light-like if $V^2 = 0$. Furthermore, if $V^0 > 0$, V is future-directed, and if $V^0 < 0$, it is past-directed.

This has some useful consequences; First, if V is time-like then there exists a Lorentz transform that eliminates all components of V but V^0. Likewise, if V is space-like then there exists a Lorentz transform such that all components but V^1 is eliminated, and if V is light-like then there exists a Lorentz transform such that $V^2 = V^3 = 0,\ V^0 = V^1$.

Covariant and Contravariant Derivatives Consider the derivative $\partial_\mu A^{\alpha\cdots}_{\cdots}$. It transforms according to

$$\partial_{\mu'} A^{\alpha'\cdots}_{\cdots} = \partial_{\mu'} x^\mu \partial_\mu (\Lambda^{\alpha'}{}_{\alpha} \dots A^{\alpha\dots}_{\cdots}),$$

where we have denoted an extra set of transformation coefficients with dots. As all of these are space-independent, we have

$$\partial_{\mu'} A^{\alpha'\cdots}_{\cdots} = \Lambda^{\alpha'}{}_{\alpha} \dots \Lambda_{\mu'}{}^{\mu} \partial_\mu A^{\alpha\cdots}_{\cdots},$$

which transforms as a tensor with an extra covariant index provided by the derivative. Hence the partial derivative transforms covariantly. The space-independence of the metric may be used to derive the machinery in special relativity, but this is not the case in general relativity, and there this derivative does not transform covariantly.

There is also a contravariant derivative defined by $\partial^\mu = g^{\mu\nu}\partial_\nu$, which indeed transforms contravariantly.

I also briefly mention the operator $\partial^2 = \partial_\mu \partial^\mu = g_{\mu\nu}\partial_\mu\partial_\nu = \partial_t^2 - \nabla^2$.

The Quotient Rule The quotient rule states that, given a relation of the form

$$A^{\alpha\beta} = G_{\alpha\beta}{}^{\delta} B_\delta$$

for two tensors A, B in some frame, G must also be a tensor.

The Zero Component Lemma Suppose that some particular vector component is zero in all frames. Then the vector is itself the zero vector.

Proper Time The proper time of two events is the time between them measured in their rest frame.

4-Velocity While the event vector x^μ transforms as a tensor, the time derivative $\partial_t x^\mu = c\partial_0 x^\mu$ does not. Explicitly we have

$$\partial_{0'} x^{\mu'} = \Lambda_{0'}{}^{\nu} \partial_\nu (\Lambda^{\mu'}{}_{\mu} x^\mu) = \Lambda_{0'}{}^{\nu} \Lambda^{\mu'}{}_{\mu} \partial_\nu x^\mu,$$

which is not the transformation rule for a rank-1 tensor. The implication is that transforming $\partial_0 x^\mu$ to a new frame does not allow us to extract velocities from the transformed coefficients. However, as we would still like to be able to find velocities in transformed frames, we would still like to define it.

To do this, consider a particle in some motion. In the lab frame the spacetime interval is given by

$$\mathrm{d}s^2 = g_{\mu\nu}\partial_0 x^\mu \partial_0 x^\nu \,\mathrm{d}(x^0)^2 = (1 - \beta_u^2)\,\mathrm{d}(x^0)^2,$$

where $\beta_u = \frac{u}{c}$ and u is the instantaneous velocity of the particle in the lab frame. For particles, which move along space-like paths, we have

$$\Delta s = \int \mathrm{d}x^0 \,\frac{1}{\gamma_u},$$

where γ_u is the instantaneous Lorentz factor calculated using u.

As the spacetime interval is invariant, we may calculate it in the rest frame of the particle. Supposing that the particle measures that it takes a (proper) time τ to traverse the path, we must have $\Delta s = c\tau$. Note that by time dilation, we have $\mathrm{d}t = \gamma_u \,\mathrm{d}\tau$, implying $\tau < \Delta t$. Now, we may reparametrize the path in the lab frame in terms of the proper time τ, such that $x^\mu = x^\mu(t(\tau))$. Defining

$$u^\mu = \frac{\mathrm{d}x^\mu}{\mathrm{d}\tau} = c\gamma_u \frac{\mathrm{d}x^\mu}{\mathrm{d}x^0} = \gamma_u(c, \mathbf{u}),$$

we have a vector. This is a vector because it is the derivative of a vector with respect to a scalar (the proper time, which is invariant under Lorentz transforms). This is termed the 4-velocity. Its norm is

$$u_\mu u^\mu = c^2\gamma_u^2(1 - \beta^2) = c^2,$$

hence it is time-like.

4-Acceleration We define the 4-acceleration as $a^\mu = \frac{\mathrm{d}u^\mu}{\mathrm{d}\tau}$. Using the above we obtain

$$a^\mu = \gamma_u \left(\frac{\mathrm{d}\gamma_u}{\mathrm{d}t} c, \frac{\mathrm{d}\gamma_u}{\mathrm{d}t}\mathbf{u} + \gamma_u \mathbf{a} \right).$$

We have $\frac{\mathrm{d}\gamma_u}{\mathrm{d}t} = \gamma_u^3 \frac{\mathbf{u}\cdot\mathbf{a}}{c^2}$. In particular, in the instantaneous rest frame of the particle we have $a^\mu = (0, \mathbf{a})$, and we define the proper acceleration $\mathbf{a}$ according to this expression. Its norm is

$$a_\mu a^\mu = -a^2,$$

hence it is space-like.

The Relation Between Velocity and Acceleration The inner product $a_\mu u^\mu$ is a scalar, and we may compute it in the instantaneous rest frame. There we obtain $a_\mu u^\mu = 0$. This can also be obtained by computing $\frac{\mathrm{d}}{\mathrm{d}\tau} u_\mu u^\mu$.

3 Relativistic Mechanics

4-Momentum and 4-Force To formulate Newton's laws in a relativistic manner, we start with Newton's second law. We try extending it to relativistic mechanics by introducing the 4-momentum

$$p^\mu = m_0 u^\mu$$

for some scalar m_0, as well as the 4-force

$$F^\mu = \frac{\mathrm{d}}{\mathrm{d}\tau} p^\mu.$$

A First Postulate Any laws of classical mechanics must result from these definitions in the limit of γ_u approaching 1. In this limit we have $P^\mu = (m_0 c, m_0 \mathbf{u})$, imploring us to recognize m_0 as the mass of the particle measured at rest and thus the space components to be those of the classical momentum.

To generalize mechanics we thus postulate that p^μ is conserved. This will lead to the conservation of both the spatial components $m_0\gamma_u\mathbf{u}$, which is termed the relativistic 3-momentum $\mathbf{p}$, and the conversion of $m_0\gamma_u\mathbf{u}$, to be discussed.

Relativistic Energy In the classical limit we obtain

$$m_0 \gamma_u c^2 \approx m_0 c^2 + \frac{1}{2} m_0 u^2,$$

imploring us to define relativistic kinetic energy as

$$T = m_0 c^2 (\gamma_u - 1)$$

and relativistic total energy as

$$E = m_0 \gamma_u c^2.$$

Having done this, we may generally write $p^\mu = \left(\frac{E}{c}, \mathbf{p}\right)$. In particular, we obtain in the rest frame that $p_\mu p^\mu = m_0 c$, and thus

$$E^2 = (m_0 c^2)^2 + (c\mathbf{p})^2.$$

Potential Energy

A Note on Relativistic Mass I make a brief note of how the above is modified by defining the relativistic mass $m = m_0 \gamma_u$. In this case you would obtain $E = mc^2$ and $\mathbf{p} = m\mathbf{u}$. In this context m_0 is termed the rest mass. This way of doing this is generally not preferred in modern contexts.

Massless Particles The expression for the energy may be extended to massless particles. In these cases we have $E = |\mathbf{p}|c$.

de Broglie de Broglie noted that photons, which had been discovered to be particles, could be attributed a momentum $|p| = \hbar|\mathbf{k}|$ as a consequence of the above. He extended this to all particles, revealing the wave-particle duality of matter to the world.