Summary of SI2371 Special Relativity

Yashar Honarmandi
yasharh@kth.se

January 11, 2021

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 (referred to as the standard configuration). 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. Also note that intervals transform according to the above regardless of configuration.

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

$$u_y' = \frac{\mathrm{d}y'}{\mathrm{d}t'} = \frac{\frac{\mathrm{d}y}{\mathrm{d}t}}{\frac{\mathrm{d}t'}{\mathrm{d}t}} = \frac{1}{\gamma\left(1 - \beta\frac{u}{c}\right)}u_y.$$

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 Spacetime Interval 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. This quantity, particularly when studied infinitesimally, is called the spacetime interval or invariant interval.

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 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 Spacetime Interval and Poincare Transformations The Lorentz transform transforms infinitesimal intervals in spacetime. We would now like to study 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 as well as translations, 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.

Tensors A tensor is a multilinear map between a set of vectors and real numbers. It 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 $\Lambda_\mu{}^\nu = \partial'_\mu x^\nu$. 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

$$\begin{aligned}\Lambda_{\mu'}{}^\mu \Lambda^{\mu'}{}_\nu &= \partial_{\mu'} x^\mu \partial_\nu x^{\mu'} = \partial_\nu x^\mu = \delta^\mu_\nu, \\ \Lambda^{\mu'}{}_\mu \Lambda_{\nu'}{}^\mu &= \partial_\mu x^{\mu'} \partial_{\nu'} x^\mu = \delta^{\mu'}_{\nu'}.\end{aligned}$$

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 $\partial_\mu A^{\alpha\dots}_{\dots}$. It transforms according to

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

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

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

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.

To prove this I propose the following: Let $A^\mu = 0$ in all frames. I may identify three different frames and write $A^{\mu'} = \Lambda^{\mu'}{}_\nu A^\nu = 0$. As the frames are different, these transformations must be linearly independent, and the only solution is thus that all components of A are zero.

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 speed 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 as the proper time is measured in the rest frame of the particle, 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 4-vector. This is true because it is the derivative of a 4-vector with respect to a scalar (the proper time, which is invariant under Lorentz transforms). This is termed the 4-velocity, and may be written such that it contains the 3-velocity $u^i = \frac{\mathrm{d}x^i}{\mathrm{d}t}$. Its norm is

$$U_\mu U^\mu = c^2\gamma_u^2(1-\beta^2) = c^2,$$

hence it is time-like. There is a simpler way to calculate this, namely in the rest frame of the particle, where $\mathbf{u} = \mathbf{0}$. This trick might be useful for other 4-vectors.

Transformation of 4-Velocities As an exercise and demonstration of the machinery we will now re-obtain the transformation rules for velocities using the machinery of 4-vectors. Consider a 4-velocity $u^\mu = \gamma_u(c, \mathbf{u})$ in some frame. In another frame moving with velocity $v\mathbf{e}_x$ relative to this frame we have

$$U^{0'} = \gamma_v(U^0 - \beta_v U^1) = \gamma_u\gamma_v(c - \beta_v u^1),\ U^{1'} = \gamma_v(U^1 - \beta_v U^0) = \gamma_u\gamma_v(u^1 - \beta_v c),$$

with the two other components unaffected. As the 4-velocity transforms like a tensor, this must be equal to $\gamma_u(c, \mathbf{u})$ in all frames, allowing us to identify

$$\gamma_{u'} = \gamma_u\gamma_v\left(1 - \beta_v\frac{u^1}{c}\right),\ u^{1'} = \frac{1}{\gamma_{u'}}U^{1'} = \frac{u^1 - \beta_v c}{1 - \beta_v\frac{u^1}{c}},\ u^{2'} = \frac{1}{\gamma_{u'}}U^{2'} = \frac{u^2}{\gamma_v\left(1 - \beta_v\frac{u^1}{c}\right)},$$

which are the familiar transformation rules.

4-Acceleration We define the 4-acceleration as $A^\mu = \frac{\mathrm{d}U^\mu}{\mathrm{d}\tau}$. Using the above we may write this as

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

where we have introduced the 3-acceleration $\mathbf{a} = \frac{\mathrm{d}\mathbf{u}}{\mathrm{d}t}$.

In particular, in the instantaneous rest frame of the particle we have $A^\mu = (0, \mathbf{a})$, and we define the proper acceleration α in this frame as $\alpha^2 = |\mathbf{a}|^2$. Thus we have

$$A_\mu A^\mu = -\alpha^2.$$

As inner products are Lorentz invariant, we conclude that the above holds in all frames and the 4-acceleration is space-like.

Derivations Related to Acceleration We have

$$\frac{\mathrm{d}\gamma_u}{\mathrm{d}t} = \frac{\partial\gamma_u}{\partial u^i}\frac{\mathrm{d}u^i}{\mathrm{d}t} = -\frac{1}{2\left(1 - \frac{u^2}{c^2}\right)^{\frac{3}{2}}} \cdot -2\frac{u^i}{c^2}\frac{\mathrm{d}u^i}{\mathrm{d}t} = \gamma_u^3\frac{\mathbf{u}\cdot\mathbf{a}}{c^2}.$$

Another useful relation may be found by considering a particle in linear motion with proper acceleration α. In a frame with respect to which the particle moves with instantaneous velocity u we use the transformation rule for velocities to find

$$\frac{\mathrm{d}u'}{\mathrm{d}t'} = \frac{\mathrm{d}u'}{\mathrm{d}u}\frac{\mathrm{d}u}{\mathrm{d}t'} = \frac{\mathrm{d}u}{\mathrm{d}t'}\frac{1 - \beta_v\frac{u}{c} + (u - \beta_v c)\frac{\beta_v}{c}}{\left(1 - \beta_v\frac{u}{c}\right)^2} = \frac{\mathrm{d}u}{\mathrm{d}t'}\frac{1 - \beta_v^2}{\left(1 - \beta_v\frac{u}{c}\right)^2}$$

in some arbitrary frame moving colinearly with the particle at velocity v relative to the original frame. In particular, if this colinearly moving frame is the instantaneous rest frame of the particle, we find

$$\frac{\mathrm{d}u'}{\mathrm{d}t'} = \frac{\mathrm{d}u}{\mathrm{d}t'}\gamma_u^2 = \gamma_u^3\frac{\mathrm{d}u}{\mathrm{d}t}.$$

Compare this with the quantity

$$\begin{aligned}
\mathrm{d}(\gamma_u u) &= \left(-u\frac{-2\frac{u}{c^2}}{2(1 - \beta_u^2)^{\frac{3}{2}}} + \frac{1}{\sqrt{1 - \beta_u^2}}\right)\mathrm{d}u \\
&= \left(\frac{\beta_u^2}{(1 - \beta_u^2)^{\frac{3}{2}}} + \frac{1}{\sqrt{1 - \beta_u^2}}\right)\mathrm{d}u \\
&= \frac{1}{(1 - \beta_u^2)^{\frac{3}{2}}}\left(\beta_u^2 + 1 - \beta_u^2\right)\mathrm{d}u \\
&= \gamma_u^3\,\mathrm{d}u\,,
\end{aligned}$$

implying

$$\alpha = \frac{\mathrm{d}}{\mathrm{d}t}(\gamma_u u).$$

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 take m_0 to be 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 in the absence of external forces, a generalization of Newton's first law. 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 conservation of $m_0 \gamma_u c$, 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 $P_\mu P^\mu = (m_0 c)^2$, and thus the mass invariant

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

The 4-momentum is thus time-like.

Potential Energy The classical total energy of a free particle comes in the low-velocity limit of $P_\mu P^\mu = (m_0 c)^2$. Attempting the replacement

$$(m_0 c^2)^2 = (E - U)^2 - (c\mathbf{p})^2$$

when adding a potential does not work, however, as it is does not represent a manipulation of 4-vectors. Instead we introduce the 4-potential

$$Q^\mu = \left(\frac{U}{c}, \mathbf{q}\right),$$

which contains the potential in its time component and the so-called potential momentum ind its space components. In the presence of a potential the proper equation is

$$(m_0 c^2)^2 = (E - U)^2 - (c\mathbf{p} - c\mathbf{q})^2.$$

It carries strong similarities to expressions taken from particles moving in electromagnetic field, an early warning sign.

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 $E = |\mathbf{p}|c$, and we thus obtain $P^\mu = \frac{E}{c}(1, \mathbf{e})$, where $\mathbf{e}$ is a unit vector. Hence the 4-momentum is light-like for massless particles.

de Broglie de Broglie combined experiments showing $E = \hbar\omega$ for photons, which had been discovered to be particles, with the work done above to show that photons could be attributed a momentum $|p| = \hbar|\mathbf{k}|$. His hypothesis was that this extends to all particles, revealing the wave-particle duality of matter to the world.

Center-of-Momentum Frames If the 4-momentum is time-like, there exists a frame in which it has no spatial components. This frame is termed the center-of-momentum frame. In many scenarios it is useful to consider because it has a high degree of symmetry.

4-Force and 3-Force In general we have

$$F^\mu = \frac{\mathrm{d}}{\mathrm{d}\tau}P^\mu = \frac{\mathrm{d}t}{\mathrm{d}\tau}\frac{\mathrm{d}}{\mathrm{d}t}P^\mu = \gamma_u\left(\frac{1}{c}\frac{\mathrm{d}E}{\mathrm{d}t}, \mathbf{f}\right),$$

a generalization of Newton's second law, where $\mathbf{f}$ is the 3-force, defined as $\mathbf{f} = \frac{\mathrm{d}}{\mathrm{d}t}\mathbf{p}$.

An Invariant From 4-Force We use Lorentz invariance to construct the invariant

$$F^\mu U_\mu = \gamma_u^2\left(\frac{\mathrm{d}E}{\mathrm{d}t} - \mathbf{f}\cdot\mathbf{u}\right).$$

In particular we have in the rest frame, where $\gamma_u = 1$, that

$$F^\mu U_\mu = \frac{\mathrm{d}E}{\mathrm{d}\tau} = \frac{\mathrm{d}}{\mathrm{d}\tau}m_0c^2.$$

Using the chain rule we then have in an arbitrary frame

$$F^\mu U_\mu = \gamma_u\frac{\mathrm{d}}{\mathrm{d}t}m_0c^2.$$

Pure and Heat-Like Forces Consider a force such that $F^\mu U_\mu = 0$, i.e. a force that preserves rest mass. For such forces we have for a general frame that

$$\frac{\mathrm{d}E}{\mathrm{d}t} = \mathbf{f}\cdot\mathbf{u},$$

and thus

$$\mathrm{d}E = \mathbf{f}\cdot\mathrm{d}\mathbf{r} = \mathrm{d}W,$$

where W is the work as defined in classical mechanics. Such forces are termed pure.

If a force is not pure, we will instead obtain an extra term in the final expression above, and we ascribe that contribution to heat. Such forces are termed heat-like.

In the general case we have

$$F^\mu = \frac{\mathrm{d}}{\mathrm{d}\tau}P^\mu = m_0A^\mu + \frac{\mathrm{d}m_0}{\mathrm{d}\tau}U^\mu.$$

As heat-like forces must be orthogonal to the 4-velocity, the first term may be interpreted to arise from pure forces and the other from heat-like forces.

Newton's Second Law With 3-Vectors Using the previously constructed invariant we have

$$\frac{\mathrm{d}E}{\mathrm{d}t} = \mathbf{f}\cdot\mathbf{u} + \frac{c^2}{\gamma_u}\frac{\mathrm{d}}{\mathrm{d}t}m_0.$$

By this we obtain

$$\mathbf{f} = \frac{\mathrm{d}}{\mathrm{d}t}(m_0\gamma_u\mathbf{u}) = \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{E}{c^2}\right)\mathbf{u} + m_0\gamma_u\frac{\mathrm{d}\mathbf{u}}{\mathrm{d}t} = \left(\frac{\mathbf{f}\cdot\mathbf{u}}{c^2} + \frac{1}{\gamma_u}\frac{\mathrm{d}}{\mathrm{d}t}m_0\right)\mathbf{u} + m_0\gamma_u\mathbf{a},$$

which expresses Newton's second law with 3-vectors.

4 Electrodynamics

Lorentz Force We would like to extend the notion of conservative forces to special relativity. The corresponding notion would be a potential ϕ such that its derivative created a pure force. However, we have

$$F_\mu U^\mu = -\partial_\mu\phi\frac{\mathrm{d}x^\mu}{\mathrm{d}\tau} = -\frac{\mathrm{d}\phi}{\mathrm{d}\tau} \neq 0,$$

hence this does not work in general. The next attempt would be of the form $F^\mu = \phi U^\mu$, but that doesn't work either. The next attempt is of the form $F_\mu = CF_{\mu\nu}U^\nu$, where the tensor components $F_{\mu\nu}$ of the so-called field strength, or Faraday tensor, are functions on spacetime.

We then have $F^\mu U_\nu = CF_{\mu\nu}U^\mu U^\nu$, which is zero if $F_{\mu\nu}$ is antisymmetric. To obtain this antisymmetry we define $F_{0i} = e_i,\ F_{ij} = -\varepsilon_{ijk}cb_k$. We then obtain

$$F_i = -\gamma_u f_i = CF_{i\mu}U^\mu = -CF_{\mu i}U^\mu = -\gamma_u C(ce_i - u_j\varepsilon_{jik}cb_k) = -\gamma_u cC(e_i + \varepsilon_{ijk}u_jb_k),$$

and thus

$$f_i = cC(e_i + \varepsilon_{ijk}u_jb_k)$$

which is on the same form as the Lorentz force. As we only know of one force of this kind in nature, we identify the simplest possible pure force as the Lorentz force and choose $C = \frac{q}{c}$. With this choice, e_i and b_i are indeed the components of the electric and magnetic field.

A More Proper Discussion The last few steps above were somewhat dubious, so let us repeat them in a way that relies less on the use of Cartesian coordinates. More specifically, treat the time coordinate as before and use an arbitrary coordinate system for the spatial coordinates. Define

$$F^{0i} = -e^i,\ F^{ij} = -\varepsilon^{ijk}cb_k,\ F_{0i} = -e_i,\ F_{ij} = -\varepsilon_{ijk}cb^k.$$

Here the "contravariant´´ components of the fields are the ones to recognize as the proper Cartesian coordinates, although the transformation properties of the above are really not those of 4-vectors. Note also that as $\varepsilon_{123} = 1$, this implies $\varepsilon^{123} = -1$. We then have

$$F^i = CF^{i\mu}U_\mu = -CF^{\mu i}U_\mu = C\gamma_u(ce^i + \varepsilon^{jik}cb_ku_j) = cC\gamma_u(e^i - \varepsilon^{ijk}u_jb_k).$$

$$F_i = CF_{i\mu}U^\mu = -CF_{\mu i}U^\mu = -C\gamma_u(ce_i - \varepsilon_{jik}cb^ku^j) = -C\gamma_u(ce_i + \varepsilon_{ijk}cb^ku^j) = -cC\gamma_u(e_i + \varepsilon_{ijk}u^jb^k).$$

This is the form of the Lorentz force in a general coordinate system, I think.

Transformation of Fields A Lorentz boost of the Faraday tensor takes the form

$$F^{\mu'\nu'} = \Lambda^{\mu'}{}_\mu\Lambda^{\nu'}{}_\nu F^{\mu\nu}.$$

This can be simplified to a matrix multiplication operation for any particular choice of coordinates - namely, the transformed components of F are given by $F' = \Lambda F \Lambda^T$. For a boost in the x-direction we find

$$\begin{bmatrix} 0 & -e^{1'} & -e^{2'} & -e^{3'} \\ e^{1'} & 0 & -cb^{3'} & cb^{2'} \\ e^{2'} & cb^{3'} & 0 & -cb^{1'} \\ e^{3'} & -cb^{2'} & cb^{1'} & 0 \end{bmatrix} = \begin{bmatrix} \gamma & -\beta\gamma & 0 & 0 \\ -\beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & -e^1 & -e^2 & -e^3 \\ e^1 & 0 & -cb^3 & cb^2 \\ e^2 & cb^3 & 0 & -cb^1 \\ e^3 & -cb^2 & cb^1 & 0 \end{bmatrix} \begin{bmatrix} \gamma & -\beta\gamma & 0 & 0 \\ -\beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} -\beta\gamma e^1 & -\gamma e^1 & -\gamma(e^2 - \beta cb^3) & -\gamma(e^3 + \beta cb^2) \\ \gamma e^1 & \beta\gamma e^1 & -\gamma(cb^3 - \beta e^2) & \gamma(cb^2 + \beta e^3) \\ e^2 & cb^3 & 0 & -cb^1 \\ e^3 & -cb^2 & cb^1 & 0 \end{bmatrix} \begin{bmatrix} \gamma & -\beta\gamma & 0 & 0 \\ -\beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 0 & -e^1 & -\gamma(e^2 - \beta cb^3) & -\gamma(e^3 + \beta cb^2) \\ e^1 & 0 & -\gamma(cb^3 - \beta e^2) & \gamma(cb^2 + \beta e^3) \\ \gamma(e^2 - \beta cb^3) & \gamma(cb^3 - \beta e^2) & 0 & -cb^1 \\ \gamma(e^3 + \beta cb^2) & -\gamma(cb^2 + \beta e^3) & cb^1 & 0 \end{bmatrix}.$$

One may now read off the transformed field components and compose the above boost with a spatial rotation to infer

$$\mathbf{e}'_{\parallel} = \mathbf{e}_{\parallel},\ \mathbf{e}'_{\perp} = \gamma(\mathbf{e}_{\perp} + c\boldsymbol{\beta} \times \mathbf{b}),\ \mathbf{b}'_{\parallel} = \mathbf{b}_{\parallel},\ \mathbf{b}'_{\perp} = \gamma\left(\mathbf{b}_{\perp} - \frac{1}{c}\boldsymbol{\beta} \times \mathbf{e}\right),$$

where the parallel and perpendicular subscripts refer to orientations with respect to the direction of the Lorentz boost.

Alternatively, expressed in terms of tensor components and transformation coefficients, we have the non-zero transformation coefficients

$$\Lambda^{0'}{}_0 = \Lambda^{1'}{}_1 = \gamma,\ \Lambda^{1'}{}_0 = \Lambda^{0'}{}_1 = -\beta\gamma,\ \Lambda^{2'}{}_2 = \Lambda^{3'}{}_3 = 1,$$
$$\Lambda_{0'}{}^0 = \Lambda_{1'}{}^1 = \gamma,\ \Lambda_{1'}{}^0 = \Lambda_{0'}{}^1 = \beta\gamma,\ \Lambda_{2'}{}^2 = \Lambda_{3'}{}^3 = 1$$

Thus,

$$e^{i'} = F^{i'0'} = \Lambda^{i'}{}_\mu \Lambda^{0'}{}_\nu F^{\mu\nu},\ \varepsilon_{i'j'k'} b^{k'} = -\frac{1}{c} F_{i'j'} = \frac{1}{c} \Lambda_{j'}{}^\mu \Lambda_{i'}{}^\nu F_{\mu\nu}.$$

First:

$$e^{1'} = \Lambda^{1'}{}_\mu \Lambda^{0'}{}_\nu F^{\mu\nu} = \Lambda^{0'}{}_0 \Lambda^{1'}{}_1 F^{10} + \Lambda^{0'}{}_1 \Lambda^{1'}{}_0 F^{01} = e^1(\gamma^2 - \beta^2\gamma^2) = e^1,$$
$$e^{2'} = \Lambda^{2'}{}_\mu \Lambda^{0'}{}_\nu F^{\mu\nu} = \Lambda^{0'}{}_\mu F^{2\mu} = \gamma e^2 - \beta\gamma F^{21} = \gamma(e^2 + \beta c\varepsilon^{213} b_3) = \gamma(e^2 - \beta cb^3),$$
$$e^{3'} = \Lambda^{3'}{}_\mu \Lambda^{0'}{}_\nu F^{\mu\nu} = \Lambda^{0'}{}_\mu F^{3\mu} = \gamma e^3 - \beta\gamma \cdot F^{31} = \gamma(e^3 + \beta c\varepsilon^{312} b_2) = \gamma(e^3 + \beta cb^2).$$

Next:

$$b^{1'} = \frac{1}{c} \Lambda_{3'}{}^\mu \Lambda_{2'}{}^\nu F_{\mu\nu} = \frac{1}{c} F_{32} = b^1,$$
$$b^{2'} = \frac{1}{c} \Lambda_{1'}{}^\mu \Lambda_{3'}{}^\nu F_{\mu\nu} = \frac{1}{c} \gamma(F_{13} + \beta F_{03}) = \frac{\gamma}{c}(cb^2 - \beta e_3) = \frac{\gamma}{c}(cb^2 + \beta e^3)$$
$$b^{3'} = \frac{1}{c} \Lambda_{2'}{}^\mu \Lambda_{1'}{}^\nu F_{\mu\nu} = \frac{1}{c} \gamma(F_{21} + \beta F_{20}) = \frac{1}{c} \gamma(cb^3 + \beta e_2) = \frac{1}{c} \gamma(cb^3 - \beta e^2).$$

The Dual Field Strength We may construct a dual field strength $\tilde{F}^{\mu\nu} = \frac{1}{2} \varepsilon^{\mu\nu\rho\sigma} F_{\rho\sigma}$, where ε is the completely antisymmetric tensor with the convention $\varepsilon_{0123} = 1$. We find:

$$\tilde{F}^{0i} = \frac{1}{2} \varepsilon^{0i\rho\sigma} F_{\rho\sigma} = \frac{1}{2} \varepsilon^{0ijk} F_{jk} = -\frac{1}{2} c\varepsilon^{0ijk} \varepsilon_{jkm} b^m.$$

Explicitly:

$$\tilde{F}^{01} = -\frac{1}{2} c(\varepsilon^{0123} \varepsilon_{231} + \varepsilon^{0132} \varepsilon_{321}) b^1 = cb^1,$$

and cyclic permutation of the spatial indexes yields $\tilde{F}^{0i} = cb^i$. Next:

$$\tilde{F}^{ij} = \frac{1}{2}\varepsilon^{ij\rho\sigma}F_{\rho\sigma}.$$

One particular choice is

$$\tilde{F}^{23} = \frac{1}{2}\varepsilon^{23\rho\sigma}F_{\rho\sigma} = \frac{1}{2}(\varepsilon^{2301}F_{01} + \varepsilon^{2310}F_{10}) = e_1.$$

To obtain the other terms, take the first occurring sequences 23 and permute in a 1 from the left or the right. The former will add a minus sign while the latter will not, yielding

$$\tilde{F}^{ij} = -\varepsilon^{ijk}e_k.$$

In other words, we obtain the dual by switching $\mathbf{e} \to -c\mathbf{b}$ and $c\mathbf{b} \to \mathbf{e}$ in the Faraday tensor.

Invariants From the Field Strength We may now construct Lorentz scalars by contracting indices of the field strength and its dual.

Consider first the quantity $F^{\mu\nu}F_{\mu\nu} = F^{0\nu}F_{0\nu} + F^{i\nu}F_{i\nu} = F^{0i}F_{0i} + F^{i0}F_{i0} + F^{ij}F_{ij} = 2F^{0i}F_{0i} + F^{ij}F_{ij}$, where we have used the fact that the Faraday tensor is antisymmetric. The first term is merely $2e^ie_i$. The second is:

$$F^{ij}F_{ij} = \varepsilon_{ijk}\varepsilon^{ijm}c^2b^kb_m.$$

Evidently there is only a contribution when $k = m$, implying that every combination of indices such that a non-zero term is produced adds the square of some component. Furthermore, all components appear equally often, so we may fix both k and m to some value and sum over i and j to find the final value. By doing that, we find that each component appears exactly twice, finally yielding

$$\frac{1}{2}F^{\mu\nu}F_{\mu\nu} = e^ie_i - c^2b^ib_i = c^2|\mathbf{b}|^2 - |\mathbf{e}|^2.$$

By performing the above described replacement, we find $F^{\mu\nu}F_{\mu\nu} = -\tilde{F}^{\mu\nu}\tilde{F}_{\mu\nu}$.

Next we form the quantity $F^{\mu\nu}\tilde{F}_{\mu\nu} = F^{0\nu}\tilde{F}_{0\nu} + F^{i\nu}\tilde{F}_{i\nu} = F^{0i}\tilde{F}_{0i} + F^{i0}\tilde{F}_{i0} + F^{ij}\tilde{F}_{ij} = 2F^{0i}\tilde{F}_{0i} + F^{ij}\tilde{F}_{ij}$. The first term is

$$F^{0i}\tilde{F}_{0i} = -ce^ib_i.$$

The second term is

$$F^{ij}\tilde{F}_{ij} = \varepsilon^{ijk}\varepsilon_{ijm}cb_ke^m = -2ce^ib_i,$$

yielding

$$\frac{1}{4}F^{\mu\nu}\tilde{F}_{\mu\nu} = -ce^ib_i = -c\mathbf{e}\cdot\mathbf{b}.$$

Maxwell's Equations We will now try to formulate Maxwell's equations using the introduced formalism. We start by introducing the 4-current $J^\mu = (c\rho, \mathbf{j})$, which we take to be a 4-vector. A general 4-current may be written as

$$J^\mu = \sum_i (c\rho_i, \rho_i\mathbf{u}_i),$$

where the charge densities ρ_i are defined in frames such that the corresponding current densities are zero.

Charge conservation takes the form $\partial_\mu J^\mu = \partial_t\rho - \vec{\boldsymbol{\nabla}}\cdot\mathbf{j} = 0$ in this formalism, and we will have to formulate alternative Maxwell equations consistent with this. We try an anzats

$$\partial_\mu F^{\mu\nu} = kJ^\nu,\ \partial_\mu\tilde{F}^{\mu\nu} = kJ^\nu_{\mathrm{m}}.$$

J_{m} is a quantity that will correspond to magnetic charge and current. Note that this guarantees charge conservation due to the Faraday tensor being antisymmetric.

Let us first study

$$\partial_\mu F^{\mu 0} = \partial_i e^i = kc\rho.$$

This already has the correct form if $k = \frac{1}{c\varepsilon_0}$. Repeating the same on the magnetic expression implies $J^0_{\mathrm{m}} = 0$ in some frame, at least.

Next, consider

$$\partial_\mu F^{\mu i} = \partial_0 F^{0i} + \partial_j F^{ji} = -\frac{1}{c}\partial_t e^i - \varepsilon^{jik} c\partial_j b_k = \frac{1}{c\varepsilon_0} j^i,$$

or equivalently

$$\varepsilon^{ijk}\partial_j b_k = \frac{1}{c^2}\partial_t e^i + \frac{1}{c^2\varepsilon_0} j^i = \mu_0 (j^i + \varepsilon_0 \partial_t e^i),$$

reproducing another one of Maxwell's equations. The corresponding expression for the dual tensor creates the final equation if $J^i_{\mathrm{m}} = 0$, and having argued that the magnetic 4-current is zero in one frame, the same must hold in all frames.

The 4-Potential We define the 4-potential as $\Phi^\mu = (c\phi, \mathbf{a})$, where ϕ is the electric potential and $\mathbf{a}$ is the magnetic potential. Using this, the Maxwell equation $\partial_\mu \tilde{F}^{\mu\nu} = 0$ is automatically satisfied. With this particular choice of 4-potential, the other will as well. Namely, we infer from the above that

$$F_{\mu\nu} = \partial_\mu \Phi_\nu - \partial_\nu \Phi_\mu,$$

and the other Maxwell equation takes the form

$$\partial_\mu \partial^\mu \Phi^\nu - \partial_\mu \partial^\nu \Phi^\mu = \frac{1}{c\varepsilon_0} J^\nu.$$

A Jacobi Identity We have

$$\begin{aligned}
\partial_\mu F_{\nu\sigma} + \partial_\nu F_{\sigma\mu} &= \partial_\mu(\partial_\nu \Phi_\sigma - \partial_\sigma \Phi_\nu) + \partial_\nu(\partial_\sigma \Phi_\mu - \partial_\mu \Phi_\sigma) \\
&= \partial_\nu \partial_\sigma \Phi_\mu - \partial_\mu \partial_\sigma \Phi_\nu \\
&= \partial_\sigma(\partial_\nu \Phi_\mu - \partial_\mu \Phi_\nu) \\
&= -\partial_\sigma F_{\mu\nu},
\end{aligned}$$

implying the Jacobi identity

$$\partial_\mu F_{\nu\sigma} + \partial_\nu F_{\sigma\mu} + \partial_\sigma F_{\mu\nu} = 0.$$

Gauge The choice of 4-potential is not unique - namely, if two 4-potentials differ by a term $\partial_\mu \Psi$, then they produce the same field strength. This means that there is a so-called gauge degree of freedom in the choice of the 4-potential.

The Lorentz Gauge The Lorentz gauge is the gauge such that $\partial_\mu \Phi^\mu = 0$.

Solutions in Vacuum In vacuum and in the Lorentz gauge the second set of Maxwell equations is simply $\partial_\mu \partial^\mu \Phi^\nu = 0$, with the solution

$$\Phi^\mu = \varepsilon^\mu e^{-ik^\nu x_\nu},$$

where k^μ is necessarily a light-like wavevector. Reinserting this into the definition of the Lorentz gauge we obtain $k^\mu \varepsilon_\mu = 0$. The vector ε^μ is called the 4-polarization, and it may always be constructed in a given frame such that $\varepsilon^0 = 0$.

General Solutions In vases where there are sources present, the general solution is constructed using the Green's function

$$G_{x_0^\mu}(x^\mu) = \frac{1}{4\pi|\mathbf{r} - \mathbf{r}_0|}\delta\left(t - t_{0-}\frac{1}{c}|\mathbf{r} - \mathbf{r}_0|\right).$$

Fields from Uniformly Moving Charges From a uniformly moving charge moving at speed v in the x-direction one can find

$$\Phi^0 = \frac{\gamma_v q}{4\pi\varepsilon_0 r'},\ \Phi^1 = \frac{\beta_v\gamma_v q}{4\pi\varepsilon_0 r'},\ r' = \sqrt{\gamma_v^2(x-vt)^2+y^2+z^2}.$$

The corresponding electric and magnetic field is

$$\mathbf{e} = \frac{q}{4\pi\varepsilon_0 r'}(\mathbf{r}-\mathbf{v}t),\ \mathbf{b} = \frac{1}{c^2}\mathbf{v}\times\mathbf{e}.$$

Maxwell's Equations From a Variational Principle It can be shown that the Lagrangian density

$$\mathcal{L} = -\frac{1}{4}\varepsilon_0 F^{\mu\nu}F_{\mu\nu} - \frac{1}{c}\Phi_\mu J^\mu$$

reproduces Maxwell's equation.

The Electromagnetic Energy Tensor In the case of continuous test charge distributions, the distribution experiences a force described by a force density

$$K_\mu = \frac{1}{c}F_{\mu\nu}U^\nu,$$

where U is the 4-velocity field. Furthermore, it can be shown that $K_\mu = -\partial_\nu {M_\mu}^\nu$ for

$${M_\mu}^\nu = -\varepsilon_0 F_{\mu\sigma}F^{\nu\sigma} + \frac{1}{4}\delta_\mu^\nu F^{\rho\sigma}F_{\rho\sigma}.$$

When elevating indices to the same height, we find that this tensor is symmetric.

Next, it can be shown that $K^0 = \frac{1}{c}\mathbf{e}\cdot\mathbf{j}$.

In the end, one obtains equations

$$\partial_t\sigma + \vec{\boldsymbol{\nabla}}\cdot c^2\mathbf{g} = -\mathbf{e}\cdot\mathbf{j},\ \partial_t g_i + \partial_j p_{ij} = -k_i.$$

5 Continuum Mechanics

The Energy-Momentum Tensor In general we define the energy-momentum tensor of a continuum such that

$$K^\mu = \partial_\mu T^{\mu\nu}.$$

It will look similar to the electromagnetic energy tensor, but notably its energy density is given by $T^{00} = c^2\rho$.

It turns out that the spatial integral of the quantity $T^{\mu 0}$ transforms like a 4-vector, and the quantity itself has the interpretation of momentum density.

6 Relativistic Optics

The 4-Wavevector The overall phase of a wave is given by $\phi_0 + \omega t - \mathbf{k}\cdot\mathbf{r}$. By defining the 4-wavevector $K^\mu = \left(\frac{\omega}{c}, \mathbf{k}\right)$, we may write the overall phase as $\phi_0 + K_\mu x^\mu$.

The 4-Frequency We may also define a 4-frequency $N^\mu = \frac{c}{2\pi}K^\mu$, where c is the wave speed.

Light and Matter For light we have $P^\mu = \hbar K^\mu = \frac{h}{c}N^\mu$. For matter the same holds, but we now have $\omega = w(\mathbf{k})k$, where w is termed the phase velocity.

Phase and Group Velocity The group velocity is defined as $u = \frac{\mathrm{d}\omega}{\mathrm{d}k}$. We have

$$u = \frac{\mathrm{d}\omega}{\mathrm{d}k} = \frac{\mathrm{d}E}{\mathrm{d}p} = \frac{c^2 p}{E} = \frac{p}{m_0\gamma_u},$$

reproducing the known expression $p = m_0\gamma_u u$. Furthermore we have

$$uw = \frac{c^2 p}{E}\nu\lambda = c^2,$$

and as u corresponds to the momentum and must therefore be smaller than c, the phase velocity is thus higher than c.

Doppler Effect for Light The Doppler effect is the observation of frequency change when moving relative to a source of a wave. Specifically for light, we may derive expressions for how much the frequency changes.

While it can be done by considering specific geometries, we instead consider the following scenario: In S there is a light source at rest. Light from this source is received by an observer from a direction $\mathbf{n}$ as seen in S. Suppose that the observer is moving with velocity $\mathbf{v}$, and let S' be its rest frame. We will find the Doppler shift by considering the invariance of the scalar $N^\mu V_\mu$. In S we have

$$N^\mu V_\mu = \nu\gamma_v(c - \mathbf{n}\cdot\mathbf{v}),$$

whereas in S' we have

$$N^{\mu'} V_{\mu'} = \nu' c,$$

yielding

$$\nu' = \nu\gamma_v\left(1 - \frac{\mathbf{n}\cdot\mathbf{v}}{c}\right).$$

Aberration of Light Consider a similar scenario to the above, and restrict $\mathbf{n}$ to the xy-plane. We may then write $\mathbf{n} = -(\cos(\theta)\mathbf{e}_x + \sin(\theta)\mathbf{e}_y)$. In S' we may write a similar expression. Applying the Lorentz transform to N^μ and equating it to $N^{\mu'}$ yields

$$\sin(\theta') = \frac{\sin(\theta)}{\gamma_v(1+\beta\cos(\theta))},\ \cos(\theta') = \frac{\cos(\theta)+\beta}{1+\beta\cos(\theta)}.$$

Furthermore we have

$$\tan\left(\frac{\theta'}{2}\right) = \frac{\sin(\theta')}{1+\cos(\theta')} = \frac{\sin(\theta)}{\gamma_v(1+\beta)(1+\cos(\theta))} = \sqrt{\frac{1-\beta}{1+\beta}}\tan\left(\frac{\theta}{2}\right).$$

7 Analytical Mechanics

On the Handling of Constraints Suppose that the system moves under some constraint $g(x^\mu) = 0$. There are two ways to handle this.

The first is to introduce a Lagrange multiplier according to

$$S = \int \mathrm{d}\theta\, \mathcal{L} + \lambda g$$

and extremize the action with the condition as an extra equation.

The other is to perform a coordinate transformation. Such transformations preserve the form of the equations of motion. The new coordinates are q^μ for all μ but the final one, and the final coordinate is g itself. You can somehow then obtain equations of motion which can be solved by explicitly constraining g.

Constructing an Action To do analytical mechanics in a way that respects relativity, we will need an action. This will reasonably have to be constructed from some non-trivial Lorentz scalars. We try

$$S = -\int \mathrm{d}\theta\, \sqrt{g_{\mu\nu}\dot{x}^\mu\dot{x}^\nu}m_0 c + \frac{q}{c}\Phi_\mu\dot{x}^\mu,$$

where we have parametrized the path in terms of some affine parameter. The involved quantities are the proper time and a line integral of the 4-potential. The dots now signify derivatives with respect to this parameter. The equations of motion are

$$\frac{\mathrm{d}}{\mathrm{d}\theta}\left(\frac{m_0 c}{\sqrt{g_{\mu\nu}\dot{x}^\mu\dot{x}^\nu}}g_{\mu\nu}\dot{x}^\nu + \frac{q}{c}\Phi_\mu\right) - \frac{q}{c}\dot{x}^\nu\partial_\mu\Phi_\nu = 0.$$

Reobtaining The Lorentz Force Law One way to proceed with the above is to choose $\theta = \tau$, for which we obtain

$$\frac{\mathrm{d}}{\mathrm{d}\tau}\left(m_0 g_{\mu\nu}\dot{x}^\nu + \frac{q}{c}\Phi_\mu\right) - \frac{q}{c}\dot{x}^\nu \partial_\mu \Phi_\nu = 0.$$

Expanding the derivative yields

$$m_0 \ddot{x}_\mu = \frac{q}{c}\dot{x}^\nu \left(\partial_\mu \Phi_\nu - \partial_\nu \Phi_\mu\right) = \frac{q}{c} F_{\mu\nu}\dot{x}^\nu,$$

which is the expected equation of motion.

The Action and 3-Vectors Another way to express the above is to choose t as the parameter, whence the action becomes

$$S = -\int \mathrm{d}t\, \sqrt{g_{\mu\nu}\dot{x}^\mu \dot{x}^\nu} m_0 c + \frac{q}{c}(c\phi - c\mathbf{u}\cdot\mathbf{a}) = -\int \mathrm{d}t\, \sqrt{1 - \frac{|\mathbf{u}|^2}{c^2}} m_0 c^2 + q(\phi - \mathbf{u}\cdot\mathbf{a}).$$

The corresponding generalized momentum is

$$\mathbf{p} = -\frac{1}{2\sqrt{1-\frac{|\mathbf{u}|^2}{c^2}}} m_0 c^2 \cdot -\frac{2}{c^2}\mathbf{u} + q\mathbf{a} = \frac{1}{\sqrt{1-\frac{|\mathbf{u}|^2}{c^2}}} m_0 \mathbf{u} + q\mathbf{a} = \gamma_u m_0 \mathbf{u} + q\mathbf{a}.$$

The equation of motion is

$$\frac{\mathrm{d}}{\mathrm{d}t}(\gamma_u m_0 \mathbf{u} + q\mathbf{a}) = -q\vec{\boldsymbol{\nabla}}\phi + u_j \vec{\boldsymbol{\nabla}} a_j.$$

It can be shown that this is equivalent to

$$\frac{\mathrm{d}}{\mathrm{d}t}(\gamma_u m_0 \mathbf{u}) = q\mathbf{e} + q\mathbf{u}\times\mathbf{b}.$$

A 3-Vector Hamiltonian It can be shown from that the above that a corresponding Hamiltonian is

$$\mathcal{H} = \sqrt{(m_0 c^2)^2 + (\mathbf{p} - q\mathbf{a})^2 c^2} + q\phi.$$

The Hamilton-Jacobi Equation I mention the Hamilton-Jacobi equation

$$\mathcal{H} + \partial_t S = 0.$$

This defines a partial differential equation for S, as the momenta are in this context to be taken to be $p_\mu = \partial_\mu S$. For details, see my summary of SI2360.

Nöether's Theorem The general formulation of Nöether's theorem is the following: Suppose that there exists a continuous transformation $\tau \to \tilde{\tau} = \tau + \delta\tau\,,\ x^\mu(\tau) \to \tilde{x}^\mu(\tilde{\tau}) = x^\mu(\tau) + \delta x^\mu(\tilde{\tau})$ that adds a term corresponding to a function F evaluated at the start and end to the action. Corresponding to such a transformation there exists a conserved quantity

$$J = F - \frac{\partial\mathcal{L}}{\partial x^\mu}\,\delta x^\mu + \left(\dot{x}^\mu \frac{\partial\mathcal{L}}{\partial \dot{x}^\mu} - \mathcal{L}\right)\delta\tau\,.$$

4-Angular Momentum It can be shown that corresponding to a Lorentz boosts there is a conserved quantity

$$L^{\mu\nu} = x^\mu p^\nu - p^\mu x^\nu.$$