Ph 1b Recitation Notes
Section 7

Jaeden Bardati
jbardati@caltech.edu

January 18, 2024

1 The Language of Special Relativity

1.1 Lorentz Transformations

1.1.1 System of Equations Representation

We have seen so far that our intuition drastically fails us often when solving problems in special relativity. This arises from the fact that there may be multiple different compounding relativistic effects (i.e. time dilation, length contraction, light travel time, etc.) that occur due to a change in inertial frame, making dealing with complex systems difficult. It would be convenient to have a simple transformation to apply when changing frames, packaging all of the effects of SR into one structure and abstracting away its postulates so we do not have to rely on our faulty intuition. These are Lorentz transformations.

In classical mechanics, we implicitly use the Galilean transformation without even knowing it. Namely, for a transformation from a frame O to another frame O', moving at a speed v relative to O in the x-direction, we intuitively think of

$$\begin{cases} t' = t \\ x' = x - vt \\ y' = y \\ z' = z \end{cases} \tag{1}$$

In other words, we simply add the distance between the two observers at any given time (at constant velocity this is vt, of course) to get the transformed frame. From what we have seen already, this cannot be true (at high speeds at least) since nothing can move faster than the speed of light, where as there is nothing in the equations stopping us from letting $v > c$. We also know that there should be a difference in the magnitude of time and space from time dilation and length contraction.

The correct way to transform between frames is with the Lorentz transformation:

$$\begin{cases} t' = \gamma(t - \frac{vx}{c^2}) \\ x' = \gamma(x - vt) \\ y' = y \\ z' = z \end{cases} \tag{2}$$

where $\gamma \equiv \frac{1}{\sqrt{1-(v/c)^2}}$ is the Lorentz factor[1]. This transformation is a direct consequence of the postulates of special relativity. For a derivation, see §8.1 of Helliwell (2010), §3 (pp. 76-80) of French (1968), §2.3 of Faraoni (2013), §12.1.3 of Griffiths (2017), §1.9 of Schutz (2023), §11.3A of Jackson (1998), or many others.

Something important to note is that the inverse of the Lorentz transformation is simply the Lorentz transformation with the substitution $v \to -v$. Namely,

$$\begin{cases} t = \gamma(t' + \frac{vx'}{c^2}) \\ x = \gamma(x' + vt') \\ y = y' \\ z = z' \end{cases} \tag{3}$$

This is, of course, required by the principle of reciprocity.

The Lorentz transformation equations (2 or 3) might seem strange at first, but it is really just a packaging of all the concepts we have seen so far. If you let $x = 0$, then the time transformation equation becomes the familiar time dilation equation with proper time t and the second is simply how the position of the frame O would move relative to the frame O' (namely, $x' = -vt'$). Similarly, if you let $t' = 0$, then the spatial transformation becomes length contraction with proper length x, and the time equation becomes the relativity of simultaneity equation (namely, $t = (v/c^2)x$). You can also show that, upon setting $t = 0$ and (separately) $x' = 0$, you can obtain the same phenomena with respect to a proper time and length in the primed frame.

The equations 2 and 3, as written, simply performs a "Lorentz boost" (i.e. transformation) to another frame that is oriented to have the same origin (in time and space) and spatial rotation. You would have to also make such transformations if it is relevant to the problem you are solving. Note also that the equations become much more symmetrical if you chose units such that $c = 1$,

[1]Note that this "triple equals" sign typically denotes definitions in physics.

or if you consider the equation written in terms of ct, ct' and $\beta \equiv \frac{v}{c}$.

1.1.2 Tensor Representation

We can package the Lorentz transformation into a 4 by 4 matrix transformation as follows.

$$\begin{bmatrix} ct' \\ x' \\ y' \\ z' \end{bmatrix} \equiv \begin{bmatrix} \gamma & -\frac{v}{c}\gamma & 0 & 0 \\ -\frac{v}{c}\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} ct \\ x \\ y \\ z \end{bmatrix} \tag{4}$$

Sometimes it is convenient to give this matrix a symbol. Typically, this is given with

$$\Lambda^{\mu'}{}_{\nu} = \begin{bmatrix} \gamma & -\frac{v}{c}\gamma & 0 & 0 \\ -\frac{v}{c}\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{5}$$

where μ and ν are simply indices that reference elements of the matrix (e.g. $\Lambda^{x'}{}_{t} \equiv \Lambda^{1}{}_{0} = -\frac{v}{c}\gamma$). Then, we could package the (ct', x', y', z') vector into the notation $x^{\mu'}$ and similarly for x^{ν}. This is a 4-vector (named so since it has four components, 3 space and 1 time). Then, we could compactly write

$$x^{\mu'} = \sum_{\nu=0}^{4} \Lambda^{\mu'}{}_{\nu} x^{\nu} \tag{6}$$

Four-vectors have unique properties that make them useful for special relativity (and even more so for general relativity).[2] We will explore these in the section 1.2.4.

1.2 Spacetime & Four-Vectors

1.2.1 Minkowski Spacetime

Classically, we treat time t independently from space $\vec{x}$. In special relativity, however, it is useful to instead treat space and time together in what is called spacetime. In classical mechanics, we use a 3-component vector $\vec{x}$ to represent space and a scalar t value to represent time. Since we mix space and time together in relativity (evidenced by the Lorentz transformation), spacetime is

[2] We could compactify it even more with Einstein summation convention. With it, we simply adopt the convention that repeated indices on one side of an equation are summed over. Thus, equation 6 becomes the quicker to write $x^{\mu'} = \Lambda^{\mu'}{}_{\nu} x^{\nu}$.

described with 4-vectors[3] such as $x^{\mu'}$.

You are no doubt familiar with the Pythagorean theorem, where $c^2 = a^2 + b^2$. This effectively calculates the distance between two points where a is the difference in position in some direction (call it the x direction), and b is the difference in the y direction. This is derived directly from Euclidean geometry and assumes the fifth postulate[4]. We can generalize this to 3 spatial dimensions giving

$$(\Delta s)^2 = (\Delta x)^2 + (\Delta y)^2 + (\Delta z)^2 \tag{7}$$

This is called the "Euclidean line element", "Euclidean metric" or "Euclidean invariant interval". It is called invariant because, classically, it does not change when translating or rotating space (about x, y or z). It is the *definition* of length in Euclidean space.

In SR, we know this Euclidean distance is no longer constant. It changes upon Lorentz transformations[5] We are therefore looking for something that is invariant under Lorentz transformations too. This will necessarily depend on time, since the Lorentz transformations do. It turns out with some reasoning [6] you can find that the quantity that stays invariant under spacetime translations (about t, x, y or z), spatial rotations (about x, y or z) or Lorentz transformations (about t, x, y or z)[7] is

$$(\Delta s)^2 = -(c\Delta t)^2 + (\Delta x)^2 + (\Delta y)^2 + (\Delta z)^2 \tag{8}$$

This is the Minkowski metric. We call $\Delta l \equiv \Delta s$ the *proper length* (since if $\Delta t = 0$, we regain the Euclidean metric, which is a distance). We call $\Delta \tau \equiv -\Delta s/c^2$ the *proper time*. You can show that these hold consistent with the definitions we had earlier.

While in the Euclidean metric, $\Delta s = 0 \implies \Delta x = \Delta y = \Delta z$ and $\Delta s \geq 0$, these are not true for the Minkowski metric. Now, Δs can be negative. This gives us an important delineation between so-called "space-like" intervals (whose $\Delta l > 0$ and $\Delta \tau < 0$), "time-like" intervals (whose $\Delta l < 0$ and $\Delta \tau > 0$) and "light-like" or "null" intervals (whose $\Delta l = \Delta \tau = 0$). Time-like intervals describe time differences, velocities, etc. and space-like intervals describe space differences, accelerations, etc. Null intervals describe the trajectory of anything that moves at the speed of light. Any time-like, space-like or null 4-vector will stay so under the transformations that leave Δs invariant.

---

[3] Or, more generally, tensors.

[4] Namely, the one stating that parallel lines never intersect or diverge away.

[5] It is a good exercise to check that the Euclidean line element is invariant under translations or rotations of space, but not Lorentz transformations.

[6] See e.g. §9.2 of Helliwell (2010) for details.

[7] It's also a good exercise to check this!

We can construct a spacetime diagram, where ct is on the y-axis and x is on the x-axis[8]. Then, you can create lines that propagate outwards at $45°$ angles from the ct-x axes (ie. $x = \pm ct$) representing the trajectories of light emitted in all directions. This is called the "light-cone", anything inside of which will be causally connected to the O observer (at the origin) and anything outside of which will be causally disconnected to the O observer. This is because if you cross the light-cone (at an angle larger than $45°$ from the ct-axis) it will have to travel at a speed greater than the speed of light for the observer O to reach the event in time. Time-like intervals originating at the origin will be within the light-cone, space-like intervals will lie outside the light-cone, and null intervals will lie on the light-cone.

1.2.2 Aside: Imaginary Time & Symmetry

Something interesting to note is that the Minkowski metric becomes a 4D Euclidean metric under the complex number transformation $t \to it \in \mathbb{C}$. What effect does this have? Well, it should change the form of the transformation that it is invariant under. The Euclidean metric is invariant under translations and rotations about all of its coordinates. The translations are, of course, unchanged and trivially satisfied for both the original Lorentz transformation and the $t \to it$ Lorentz transformation. The rotations, however, are more interesting. A 4D Euclidean metric should have $\binom{4}{2} = 6$ rotations since it needs to rotate two of its four dimensions, compared to the $\binom{3}{2} = 3$ rotations in the 3D Euclidean metric. Thus, we expect the analogous $t \to it$ Lorentz transformation to have 6 rotations also, in addition to the 4 translations (for a total of 10). The regular Minkowski metric is left invariant under 4 translations, 3 spatial rotations and 3 Lorentz boosts. Thus, the Lorentz boosts must act like rotations under $t \to it$.

To get a better view of this, we will define the rapidity $\phi \in \mathbb{R}$ such that $\tanh\phi \equiv \frac{v}{c}$, then you can manipulate equation 5 to get

$$\Lambda^{\mu'}_{\ \nu} = \begin{bmatrix} \cosh\phi & -\sinh\phi & 0 & 0 \\ -\sinh\phi & \cosh\phi & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{9}$$

This resembles closely to the equation for a rotation, but with hyperbolic sinusoidal functions instead of the regular ones (and some sign changes). This makes sense since the equation for Δs in Minkowski spacetime is hyperbolic, instead of the Euclidean space, which is ellipsoidal. It turns out that you can let $\theta = i\phi$ and do the $t \to it$ transformation, and making the Lorentz transformation exactly a rotation in the plane of imaginary time it and x.

[8] Or any other spatial direction you like. Optionally, you could have two spatial directions for a 3D plot.

The 4 translations, 3 spatial rotations and 3 Lorentz boosts form a symmetry group called the Poincaré group for the invariant quantity Δs. The Lorentz boost symmetries form a subgroup of the Poincaré group called the Lorentz group, which is related to SO(3), the symmetry group of rotations.

1.2.3 Aside: General Relativity

It may also be interesting to note that the metric represents the shape of the spacetime. Specifically, the Minkowski metric is only valid in SR (or locally in GR). It is not valid, for example, around a black hole and instead there is a different metric. The metric for a spherically symmetric static black hole (the so-called "Schwarzschild" black hole) is actually fairly simple:

$$\mathrm{d}s^2 = -\left(1 - \frac{2M}{r}\right) c^2 \mathrm{d}t^2 + \left(1 - \frac{2M}{r}\right)^{-1} \mathrm{d}r^2 + r^2 \mathrm{d}\theta^2 + r^2 \sin^2\theta \mathrm{d}\phi^2 \tag{10}$$

where the discrete Δs is replaced with the more generalized differential $\mathrm{d}s$ and the equation is represented in spherical coordinates[9]. There are other types of metrics that describe other kinds of black holes such as the Kerr black hole which rotates, or the Reissner–Nordström black hole which is electromagnetically charged. If you are more interested in this you may want to look into an introductory general relativity book such as Schutz (2023).

1.2.4 Aside: Four-Vectors & Tensors

A 4-vector is not just a list of four numbers as you might expect. It is a bit more special. It has the particular property of transforming in a particular way under a change of coordinates. Namely, a 4-vector A^ν under the coordinate transformation $x^\nu \to x^{\mu'}$ transforms like

$$A^{\mu'} = \sum_{\nu=0}^{4} \frac{\partial x^{\mu'}}{\partial x^\nu} A^\nu \tag{11}$$

This is how a vector is defined. Any four-vector by this definition will necessarily always transform in this way. This includes under Lorentz transformations [10]. The key piece of information that is gained by knowing that a list of four numbers transforms like a 4-vector (and thus is a four-vector by definition) is that it . For example, since 4-velocity and 4-acceleration (constructions on top of the usual 3-velocity and 3-acceleration) are four-vectors, when you boost to another frame, you know that they should transform in a predictable way: As in equation 11 with transformation matrix $\frac{\partial x^{\mu'}}{\partial x^\nu} = \Lambda^{\mu'}{}_{\nu}$. If you instead transform from Cartesian coordinates to polar coordinates, you could also transform any

[9] You might recognize some of the terms already!

[10] You can check this manually if you want.

4-vector with the same transformation using the respective coordinate transformation matrix defined by $\frac{\partial x^{\mu'}}{\partial x^{\nu}}$ where μ and ν are indices of the matrix (note that, under that view, equation 11 is just a matrix multiplication). There are examples of objects that have four elements, but are not 4-vectors because they do not have this sort of "coupling" with the position transformation. For example, four constants (i.e. 0, 1, π, etc.) do not transform at all and so are not 4-vectors.

The mathematical generalization of 4-vectors are tensors. In particular, 4-vectors are a form of "rank 1" tensors. Rank 0 tensors are simply scalars. Scalars are defined to be invariant of . For example π is a scalar and the trace of a rank 2 tensor is a scalar. Rank 2 tensors are effectively matrices, as you know them, but with another transformation restriction. A rank 2 (fully contravariant) tensor $T^{\mu\nu}$ transforms as

$$T^{\mu'\nu'} = \sum_{\mu,\nu} \frac{\partial x^{\mu'}}{\partial x^{\mu}} \frac{\partial x^{\nu'}}{\partial x^{\nu}} T^{\mu\nu} \tag{12}$$

You can see how these tensors can be generalized to an arbitrary amount of dimensions. There are also covariant (or dual) vectors and rank 2 tensors, denoted with lower indices. A covariant vector A_ν is defined as transforming like

$$A_{\mu'} = \sum_{\nu} \frac{\partial x^{\nu}}{\partial x^{\mu'}} A_{\nu} \tag{13}$$

And similarly for a general rank N (fully covariant) tensor. Note that this definition of dual 4-vectors, makes that the summed multiplication of a vector with its corresponding dual vector is a scalar:

$$\sum_{\mu'} A^{\mu'} A_{\mu'} = \sum_{\mu'} \left(\sum_{\nu} \frac{\partial x^{\nu}}{\partial x^{\mu'}} A_{\nu} \right) \left(\sum_{\nu} \frac{\partial x^{\mu'}}{\partial x^{\nu}} A^{\nu} \right) = \sum_{\nu} A^{\nu} A_{\nu}$$

While 4-vectors can be thought of as arrows pointing in spacetime in a given direction, dual 4-vectors (also called one-forms) can be thought of as contours in a given direction, denoting the definition of distance. This should ring some bells in your head. The notion of distance is found from the line element equation (in SR, the Minkowski spacetime). So there must be some object that carries the notion of distance from our chosen spacetime metric. This is the metric tensor $g_{\mu\nu}$. The metric tensor is defined from the line element as

$$ds^2 = \sum_{\mu\nu} g_{\mu\nu} x^{\mu} x^{\nu} \tag{14}$$

In Minkowski spacetime metric is typically denoted with $\eta_{\mu\nu}$ and is found from equations 14 and 8 to be $g_{\mu\nu} = \eta^{\mu\nu} = \eta_{\mu\nu} \equiv \mathrm{diag}(-1, 1, 1, 1)$ (i.e. the matrix formed where the diagonal elements are -1, 1, 1, 1, and the non-diagonal elements are 0)[11]. Then, we can raise and lower the indices (convert between the two vector types) with

$$A_\mu = \sum_\nu g_{\mu\nu} A^\nu$$
$$A^\nu = \sum_\mu g^{\mu\nu} A_\mu$$

and vice versa. Note that the metric tensor is always symmetric (i.e. $g_{\mu\nu} = g_{\nu\mu}$) and $\sum g_{\mu\nu} g^{\mu\nu} = 4$. We can define a dot-product between vectors u^μ and w^ν, that converges to the regular dot product when using the Euclidean metric, as

$$\sum_{\mu,\nu} u^\mu w_\nu = \sum_{\mu,\nu} g_{\mu\nu} u^\mu w^\nu \tag{15}$$

It turns out that defining the dot product in this way makes it a scalar. You can also have mixed tensors T^μ_ν whose individual components transform like vectors or one-forms accordingly.

$$T^{\mu'}_{\nu'} = \sum_{\mu,\nu} \frac{\partial x^{\mu'}}{\partial x^\mu} \frac{\partial x^\nu}{\partial x^{\nu'}} T^\mu_\nu \tag{16}$$

These typically represent transformations matrices. You might have noticed that this is similar notation to when I defined the Lorentz transformation $\Lambda^{\mu'}_{\nu}$ in equation 5. This is because the Lorentz transformation is a mixed rank 2 tensor[12]. Note that typically we do not include the sums in these tensor equations and implicitly as

If you want to understand 4-vectors and tensors further (and especially if you intend to study general relativity), I would recommend reading Neuenschwander (2015) for a concise text on tensor calculus for physics, or Faraoni (2013) for a more SR text that dives deep into tensors and how they are used in relativity. Please do note, however, that for Physics 1b you will only need to know 4-vectors for practical uses (in relativistic mechanics) and you will not need to know tensors at all.

[11] Here, I adopt what is called the -+++ signature. Typically, relativists use this signature. There is also the +--- signature, often used by particle physicists, where the definition of the metric is exactly negative what is shown in eq. 14 if you use the proper length as ds, or by just simply using the proper time as ds instead.

[12] Note that it is common to write that it is a $\binom{1}{1}$ tensor, where the upper number indicates the number of upper indices (or contravariant indices), and vice versa for the covariant indices.

1.3 Example: Velocity Transformation

Now, equipped with a new system to study special relativity, we can somewhat easily investigate how a 3-velocity transforms under a Lorentz boost.

1.3.1 Three-Vector Method

The standard way to do this is by using the Lorentz transformation as defined by the system of equations (2). Consider a 3-velocity $\vec{v} \equiv (v_x, v_y, v_z) \equiv (\frac{\mathrm{d}x}{\mathrm{d}t}, \frac{\mathrm{d}y}{\mathrm{d}t}, \frac{\mathrm{d}z}{\mathrm{d}t})$ in frame O, which corresponds to the 3-velocity $\vec{v'} \equiv (v'_x, v'_y, v'_z) \equiv (\frac{\mathrm{d}x'}{\mathrm{d}t'}, \frac{\mathrm{d}y'}{\mathrm{d}t'}, \frac{\mathrm{d}z'}{\mathrm{d}t'})$ in frame O', which is moving with speed V in the x direction away from O. From the Lorentz transformation, we know how to relate t', x', etc. to t, so we can use the chain rule to get that

$$\frac{\mathrm{d}x'}{\mathrm{d}t'} = \frac{\mathrm{d}x'}{\mathrm{d}t}\frac{\mathrm{d}t}{\mathrm{d}t'}$$

And similarly for y' and z'. By equations (2), we know that $\frac{\mathrm{d}x'}{\mathrm{d}t} = \gamma(\frac{\mathrm{d}x}{\mathrm{d}t} - V)$, $\frac{\mathrm{d}t}{\mathrm{d}t'} = \gamma(1+(V/c^2)\frac{\mathrm{d}x'}{\mathrm{d}t'}) \implies \frac{\mathrm{d}t'}{\mathrm{d}t} = \gamma(1-(V/c^2)\frac{\mathrm{d}x}{\mathrm{d}t})$ by the principle of reciprocity. Moreover, $\frac{\mathrm{d}y'}{\mathrm{d}t} = \frac{\mathrm{d}y}{\mathrm{d}t}$, and $\frac{\mathrm{d}z'}{\mathrm{d}t} = \frac{\mathrm{d}z}{\mathrm{d}t}$. Thus, we obtain the equations of 3-velocity transformation

$$\begin{cases} v'_x = \frac{v_x - V}{1 - V v_x/c^2} \\ v'_y = \frac{v_y\sqrt{1-(V/c)^2}}{1 - V v_x/c^2} \\ v'_z = \frac{v_z\sqrt{1-(V/c)^2}}{1 - V v_x/c^2} \end{cases} \tag{17}$$

We can notice that the 3-velocity changes even perpendicular to the velocity of the frame (i.e. v_y and v_z). This change in magnitude comes from both time dilation and the relativity of simultaneity (see Prof. Patterson's note on velocity transforms for details). The velocity in the parallel direction to the frame motion (v_x) is further changed from the perpendicular directions due to length contraction and the regular intuitive understanding of velocity addition.

We can also notice that, in the Newtonian limit, where $v \ll c$, we obtain the intuitive $v'_x = v_x - V$ with v_y and v_z left unchanged.

1.3.2 Four-Vector Method

To show 4-vectors in use, we will derive this again with them. We will conveniently package our velocities $\vec{v}$ and $\vec{v'}$ into special 4-vectors called "4-velocities". In a rest frame, an object moves at zero 3-velocity, but not zero 4-velocity. This is because it also moves in time. Therefore, an object observes its own 4-velocity as $u^\nu_{\text{rest}} = (c, 0, 0, 0)$.[13] This is a 4-vector, so we can boost it to the O frame by

[13] Or, the shorthand $u^\nu_{\text{rest}} = (c, \vec{0})$.

multiplying it with the Lorentz transformation matrix. For now, we will only consider a velocity in the x direction so that the matrix in equation 5 will hold, and then we can rotate the frame to match a general velocity observed in O afterwards. So, the transformed 4-velocity becomes

$$u^\mu = \Lambda^\mu{}_\nu u^\nu_{\text{rest}}$$

where there is an implicit summation over the value ν on the right-hand-side from the Einstein summation convention. Actually doing this multiplication with the matrix from equation 5 gives that $u^\nu = (\gamma c, -\gamma v, 0, 0)$. We can then rotate the velocity spatially so that it matches with the observed $\vec{v}$ in frame O, getting

$$u^\nu = \gamma(c, \vec{v}) \tag{18}$$

This is the general 4-velocity formula for an object moving at velocity $\vec{v}$ relative to O. From symmetry, we expect the v_z velocity addition to be equivalent to the v_y velocity addition. We will therefore only consider $u^\nu = \gamma_{\vec{v}}(c, v_x, v_y, 0)$, where I distinguish $\gamma_{\vec{v}} = \left(1 - (|\vec{v}|/c)^2\right)^{-1/2}$ from $\gamma_V = \left(1 - (V/c)^2\right)^{-1/2}$. Then, we simply transform the 4-velocity to become

$$u^{\mu'} = \sum_\nu \Lambda^{\mu'}{}_\nu u^\nu = \gamma_{\vec{v}} \left(\gamma_V c - \gamma_V (V v_x/c), \gamma_V v_x - \gamma_V V, v_y, 0\right)$$

We already know from equation 20 that this should also be equal to $u^{\mu'} = \gamma_{\vec{v'}}(1, \vec{v'})$. So, we can solve for $\gamma_{\vec{v'}}$ by the equating the time component of the 4-vector. Namely,

$$\gamma_{\vec{v'}} = \gamma_{\vec{v}}\gamma_V (1 - V v_x/c^2)$$

Comparing now the x and y components gives

$$v'_y = \frac{\gamma_{\vec{v}}\gamma_V (v_x - V)}{\gamma_{\vec{v}}\gamma_V (1 - V v_x/c^2)}$$
$$v'_y = \frac{\gamma_{\vec{v}} v_y}{\gamma_{\vec{v}}\gamma_V (1 - V v_x/c^2)}$$

Remembering that $y \to z$ will give the equation for v'_z, we can plug in the γs and obtain equation 17, as expected.

2 Relativistic Mechanics

2.1 Conservation of Momentum and Energy

Conservation of momentum and conservation of energy are two important concepts in classical physics. In order to use them in SR, however, we require them to follow the two postulates. This means that if momentum is conserved in one frame, it must be conserved in all frames (and likewise for energy). Let's check if the regular momentum and energy conservation equations are invariant of the frame you are in.

Classically, 3-momentum $\vec{p} \equiv m\vec{v}$ between two particles of mass m_A and m_B is

$$m_A \vec{v}_{Ai} + m_B \vec{v}_{Bi} = m_A \vec{v}_{Af} + m_B \vec{v}_{Bf} \tag{19}$$

where i denotes the initial momentum and f denotes the final momentum. This is invariant to the Galilean transformation (and hence is a valid conservation equation classically). We know how velocities transform (equation 17), so we can simply plug in the transformation into 19 and we find that both sides are not equal anymore. Thus, conservation of 3-momentum is violated in SR. We can notice that for a photon $p = E/c$, so energy conservation is also violated.

We must therefore look for make the conservation laws invariant of Lorentz transformations. The way to do this is with our newfound tool of 4-vectors.

2.2 Four-Momentum

To formalize relativistic mechanics, let's first start with the most basic element 4-position defined as $x^\mu = (ct, x, y, z) = (ct, \vec{x})$ for some point particle with mass m. We can then define its 4-velocity as

$$u^\mu \equiv \frac{\mathrm{d}x_\mu}{\mathrm{d}\tau} \tag{20}$$

where τ is the proper time of u_μ in a given frame. This effectively means that u_μ is always tangent to the so-called "worldline" x_μ, with a special parameter τ. The worldline is the (sometimes not straight) line in some observer's frame that views a particle with trajectory x^μ in spacetime. Thus, u_μ acts as we might expect a velocity to.

In the rest frame of the particle, the proper time is the observer time ($\tau = t$), so we get $= u_\mu \doteq (c, \vec{0})$.[14] This is exactly what we defined in section 1.3.2. Thus, in some general frame O, a 3-velocity $\vec{v}$ will have an equivalent 4-velocity

$$u^\mu = \gamma(c, \vec{v}) \tag{21}$$

where $\gamma \equiv \frac{1}{\sqrt{1-(|\vec{v}|/c)^2}}$. Now we have a way to transform velocities with an arbitrary amount of Lorentz transformations. Namely, we convert it from a

[14] Note that here I use the special notation $\doteq$ to denote "equal in some frame", distinct from $=$ which denotes "equal in all frames".

3-velocity to at 4-velocity, Lorentz transform and then convert back to a 3-velocity. Note that at any point, we can just take the spacial components of a 4-vector by using Latin indices instead of Greek ones. For example, with the 4-velocity

$$u^i = \gamma v^i \tag{22}$$

It is sometimes useful to know that the dot product $u^\nu u_\nu \equiv -(u^t)^2 + (u^x)^2 + (u^y)^2 + (u^z)^2 = -c^2$ and thus is a scalar[15]. We could also define the 4-acceleration in a similar way, such that

$$a^\mu \equiv \frac{\mathrm{d}u^\mu}{\mathrm{d}\tau} = \frac{\mathrm{d}^2 x^\mu}{\mathrm{d}\tau^2} \tag{23}$$

A useful property to know is that the dot product between the 4-velocity and 4-acceleration can be shown to be $u^\nu a_\nu = 0$.[16]

Now, we define the 4-momentum as you might expect

$$p^\mu \equiv m u^\mu \tag{24}$$

Since it is a 4-vector, it transforms like one, and so any linear combination of p^μs will also follow the same transformation. Thus, **four momentum is conserved**.

This is all rather abstract though, so let's turn the momentum equation into its 3-vector counterpart. Namely, take the spatial components of the 4-momentum

$$\begin{aligned} p^i &= m u^i = \gamma m v^i \\ \implies \vec{p} &= \gamma m \vec{v} \end{aligned} \tag{25}$$

Thus, the relativistic 3-momentum is a Lorentz factor away from the classical 3-momentum. Notice that in the Newtonian limit as $\gamma \to 1$, we recover the classical 3-momentum $m\vec{v}$. Thus, the momentum that is conserved is really $\gamma m \vec{v}$ not $m\vec{v}$.

You might have noticed that we have not looked at the time component of the 4-momentum yet. It should be conserved too. So, what conservation law does it hold information about? Let's find out. The time component is

[15] Note this dot product definition is only valid in special relativity (i.e. Minkowski space-time). See section 1.2.4 for the general definition.

[16] We call them "orthogonal vectors". This effectively means that if u^ν is space-like (it is for massive particles from equation 21), then a_ν is necessarily space-like and vice versa. Similarly, u^ν is null if and only if a_ν is null.

$$
\begin{aligned}
p^t &= mu^t = \gamma mc \\
\implies p^t &\equiv \frac{E}{c}
\end{aligned} \tag{26}
$$

In fact, this is energy E conservation. So, we can conclude that

$$E = \gamma mc^2 \tag{27}$$

is the relativistic energy. When rewritten in terms of the magnitude of the relativistic 3-momentum and the mass

$$E = \sqrt{p^2c^2 + m^2c^4} \tag{28}$$

This is a scalar equation and thus is true in any frame. Or, in the rest frame (the so-called rest energy), we obtain the famous

$$E \doteq mc^2 \tag{29}$$

This is why we call m the rest mass of the particle. This means that there is an equivalence between mass and energy. With this equation, you can predict the amount of energy released in nuclear reactions (which transfer mass into energy or vice versa). Note that it is sometimes convenient to package energy and momentum together as

$$p^\mu = \gamma m \left(c, \vec{v}\right) \tag{30}$$

which we can use to transform easily between frames.

2.3 Aside: Four-Force

While it is not possible to use gravity in special relativity (we need general relativity for that), we can cheat our way around to include another force, such as electrodynamics, into the mix. To do this, we define a relativistic 4-force f^μ as

$$f^\mu \equiv \frac{\mathrm{d}p^\mu}{\mathrm{d}\tau} \tag{31}$$

For constant mass m, this is just $f^\mu = ma^\mu$. We could find that its dot product with velocity is just $f^\mu u_\mu = -\dot{m}c^2$ where $\dot{m} \equiv \frac{\mathrm{d}m}{\mathrm{d}\tau} = 0$ if the mass of the particle is constant. The relativistic 3-force is defined from the relativistic 3-momentum $\vec{p}$ as $\vec{F} \equiv \frac{\mathrm{d}\vec{p}}{\mathrm{d}t}$.[17] So, to some observer,

[17]Note that this implicitly includes a Lorentz factor compared to the classical definition (i.e. $\vec{p} = \gamma m\vec{v}$). Also note that this is defined with respect to an observer time t not the proper time τ.

$$f^\mu \equiv \left(f^t, \vec{f}\right) = \left(\frac{\gamma}{c}\frac{\mathrm{d}E}{\mathrm{d}t}, \gamma\vec{F}\right) \tag{32}$$

You can therefore find that the energy change due to the work by $\vec{F}$ against the particle is $\frac{\mathrm{d}E}{\mathrm{d}t} = \vec{F}.\vec{v}$. And that the spatial component of the relativistic 4-force (distinct from the relativistic 3-force due to the t vs. τ difference) is $\vec{f} = \gamma\vec{F} = \gamma\frac{\mathrm{d}}{\mathrm{d}t}\left(\gamma m\vec{v}\right)$. That is how you construct a 4-force relative to a particular frame in special relativity. You can then just transform the force to any frame (or coordinate system) you like with the Lorentz transformation (or other transformation) and decompose to get the regular non-relativistic 3-vector quantities you are used to if you like.

From the 4-force, we can define relativistic kinetic energy by first defining work in the usual way

$$W = \int |\vec{F}|\mathrm{d}s \tag{33}$$

Using the work-energy theorem, we can eventually conclude that the relativistic kinetic energy of a particle with mass m moving with speed v is $T = (\gamma - 1)mc^2$. If we wanted, we could also define the relativistic angular momentum of a particle as a pseudo-tensor $l^{\mu\nu} = x^\mu p^\nu - x^\nu p^\mu \equiv 2x^{[\mu]}p^{\nu]}$. We could also define particle systems as having total mass, momentum and angular momentum as the sums of the four-vector/tensor analogues. We have then derived all of (massive particle) mechanics in special relativity.

2.4 Example: Particle Decays

Suppose we have a single particle of mass M that decays into two particles of mass m_1 and m_2. We want to find what the energy of the resulting particles are.

In the rest frame of the single particle, by conservation of energy

$$E = E_1 + E_2 = Mc^2$$

since $\gamma = 1$. And, by conservation of momentum,

$$\begin{aligned} \vec{0} = \vec{p} &= \vec{p_1} + \vec{p_2} \\ \implies p_1 &= p_2 \end{aligned}$$

Using equation 28 for particle 1 gives

$$E_2^2 = p_2^2c^2 + m_2^2c^4$$

$$
\begin{aligned}
&\implies (Mc^2 - E_1)^2 = p_1^2c^2 + m_2^2c^4 \\
&\implies M^2c^4 - 2Mc^2E_1 + E_1^2 = (p_1)^2c^2 + m_2^2c^4 \\
&\implies M^2c^2 - 2ME_1 + m_1^2c^2 = m_2^2c^2 \\
&\implies E_1 = \left(\frac{M^2 + m_1^2 - m_2^2}{2M}\right)c^2
\end{aligned}
$$

Thus, the energy of the resulting particle 1 will be E_1 as above. By symmetry, there will be a similar equation for E_2 with $(m_1, m_2) \to (m_2, m_1)$.

2.5 Relativistic Optics

So far we have only considered massive particles. However, photons (and many other particles) are massless. We cannot even define a proper time since, along a null trajectory, $\Delta s = 0 \implies \Delta\tau = 0$. Since you cannot Lorentz boost to a photon trajectory (vectors that are null in one frame must be null in all other frames), we can only define a photon trajectory from the perspective of some other frame. We therefore define it according to some parameterization λ. So, for a null trajectory

$$
u^\mu \equiv \frac{\mathrm{d}x_\mu}{\mathrm{d}\lambda} \tag{34}
$$

Necessarily, its magnitude is $u^\mu u_\mu = 0$. In Minkowski spacetime, light rays cannot accelerate and follow straight lines, so we can always choose a parameter λ such that $\frac{\mathrm{d}u^\mu}{\mathrm{d}\lambda} = 0$.[18] This means that it can be written in terms of a 3-velocity $\vec{u}$ as

$$
u^\mu = (|\vec{u}|, \vec{u}) \tag{35}
$$

We cannot now follow the usual 4-momentum equation (24). Instead we define It is sometimes useful to write a 4-vector that contains information about the frequency and wave number of the light. We define the wave 4-vector as

$$
k^\mu \equiv \frac{\omega}{c}u^\mu \tag{36}
$$

where ω is the angular frequency of the wave. For null trajectories, specifically, we have

$$
k^\mu = \left(\frac{\omega}{c}, \vec{k}\right) \tag{37}
$$

[18]Such a paramter is called “affine”.

where $\vec{k}$ is the usual wave number. If we match with the usual definition of photon 3-momentum $\vec{p} = \hbar\vec{k}$, then we can define the photon 4-momentum as

$$p^\mu \equiv \hbar k^\mu = \left(\frac{\hbar\omega}{c}, \hbar\vec{k}\right) = \left(\frac{E}{c}, \vec{p}\right) \tag{38}$$

We can then recover the energy equation for photons using the normalization $p_\nu p^\nu = 0$ to get, effectively, equation 28 with $m = 0$). Namely,

$$E = pc \tag{39}$$

There are lots of strange effects in relativistic optics such as a relativistic Doppler effect (where light changes frequency based on the relative speed of the observer from the source), relativistic beaming (where light is much brighter and centrally concentrated due to its motion), and how extended objects will appear to morph in shape depending on certain conditions. If you are interested, see §13.2 of Helliwell (2010) or §7 of Faraoni (2013).

2.6 Example: Compton Scattering

Suppose a photon at an initial wavelength λ collides with an electron of mass m_e. The photon will scatter at an angle θ relative to its original trajectory and the electron will scatter at some other angle -ϕ. We want to find out what the resulting photon wavelength λ' will be.

Using relativistic 4-momentum conservation along the t, x and y axis, we can find the energy and momentum of the electron. Doing this gives

$$\begin{cases} E + m_e c^2 = E' + E_e \\ p = p' \cos\theta + p_e \cos\phi \\ 0 = p' \sin\theta + p_e \sin\phi \end{cases}$$

We can solve the third equation for ϕ and substitute it back into the second equation to get

$$\begin{cases} E_e = E - E' + m_e c^2 \\ p_e^2 = p^2 + p'^2 - 2pp' \cos\theta \end{cases}$$

We can combine energy and momentum with $E_e^2 = p_e^2 c^2 + m_e^2 c^4$, giving

$$\left(E - E' + m_e c^2\right)^2 = \left(p^2 + p'^2 - 2pp' \cos\theta\right) c^2 + m_e^2 c^4$$

We can relate the momentum and energy of a photon to its wavelength with $E = h\nu = hc/\lambda$, $p = h\nu/c = h/\lambda$, $E' = h\nu' = hc/\lambda'$, and $p' = h\nu'/c = h/\lambda'$. Then, solving for λ' gives the Compton scattering equation:

$$\lambda' = \lambda + \frac{h}{m_e c}(1 - \cos\theta) \tag{40}$$

Thus, Compton scattering reddens the wavelength of light. Compton scattering is very important in astronomy. The related inverse electron scattering, where an electron scatters off of a photon, *increases* the frequency of light instead of decreasing it (hence the name). Around supermassive black holes (in the centers of galaxies), there is an accretion disk of material that emits a thermal spectrum (you can even now see a picture of such light around the black hole in the center of our galaxy, Sag A*). There is highly ionized matter that obscures this accretion disk, called the corona. The electrons from the corona scatter cause the thermal photons from the disk to extremely high energies that we can detect with X-ray telescopes.

2.7 Example: Energy-Momentum In Another Frame

Problem (Helliwell (2010) problem 13.1): We observe an electron of mass 0.511 MeV/c^2 moving in the +x direction at speed $v = (3/5)c$. Find its momentum and energy in our frame, and also in a frame moving in the +x direction at speed $v = (4/5)c$.

The power of 4-vectors is for exactly such a scenario. We have that the rest mass of a particle when moving at $v = (3/5)c$ to the right is $m = 0.511$ MeV/c^2. The velocity relative to our frame is $\vec{v} = ((3/5)c, 0, 0)$. We can therefore construct the 4-momentum from equation 30 as

$$\begin{aligned} p^\mu &= \gamma m(c, \vec{v}) \\ &= \frac{0.511\ \text{MeV}/c^2}{\sqrt{1-(3/5)^2}}(c, (3/5)c, 0, 0) \\ &= (0.6388\ \text{MeV}/c, 0.3833\ \text{MeV}/c) \end{aligned}$$

Thus, relative to us, the electron has an energy of $E = 0.639$ MeV and a momentum of $p = 0.383$ MeV/c. To find it in a frame moving at $v = (4/5)x$ in the x direction (call it the O' frame), we can Lorentz transform it with equation 5 to get

$$p^{\mu'} = \sum_{\nu=0}^{4} \Lambda^{\mu'}{}_{\nu} p'^{\mu}$$

$$= \frac{1}{\sqrt{1-(4/5)^2}}(0.6388 - (4/5) * 0.3833, 0.3833 - (4/5) * 0.6388)\ \text{MeV}/c$$

$$= (-0.214 \text{ MeV}/c, 0.554 \text{ MeV}/c)$$

Thus, relative to the O' frame, the electron has an energy of $E' = 0.554$ MeV and a momentum of $p' = -0.214$ MeV/c.

References

V. Faraoni. *Special Relativity*. Undergraduate Lecture Notes in Physics. Springer International Publishing, Cham, Switzerland, 2014 edition, Aug. 2013.

A. French. *Special Relativity*. Norton, 1968.

D. J. Griffiths. *Introduction to Electrodynamics*. Cambridge University Press, 2017.

T. Helliwell. *Special Relativity*. University Science Books, 2010.

J. D. Jackson. *Classical Electrodynamics*. John Wiley Sons, Inc., 1998.

D. E. Neuenschwander. *Tensor Calculus for Physics: A Concise Guide*. Johns Hopkins University Press, 2015.

B. F. Schutz. *A First Course in General Relativity*. Cambridge University Press, 2023.