Special Relativity

Haydn Cheng

November 23, 2024

Contents

1 Kinematics

1.1 Lorentz Transformation

An event is something that an observer can specify with spatial and time coordiantes. For example somebody clap their hands, two spaceships crashes onto each other, or more abstractly a specific point at a given time. Something that is not an event is when the observer cannot specify the point in space where or the time when it occurs, such as a person staring at a pebble for one minute (because the time coordinate is not specified), or a train moving by an observer (because the train has spatial extent).

Lorentz transformation attempts to relate the coordinates of two different events[1] observed in two different inertial frames (S and S') which have a relative speed v. That is, we want to find the constants A, B, C, D in the relations[2][3]

$$\begin{aligned} \Delta x &= A\Delta x' + B\Delta t' \\ \Delta t &= C\Delta t' + D\Delta x'. \end{aligned} \tag{1.1}$$

The four independent equations depending on four independent facts, which are

1. S' moves with velocity v with respect to S.
2. S moves with velocity $-v$ with respect to S'.
3. Time dilation (or length contraction) looks the same from either frame.
4. A light pulse with speed c in S' also has speed c in S.

[1] Since there is no preferred origin, an event can be described by arbitrary x and t values simply by shifiting the origin. Therefore there is no physical significance to compare the coordinates between two frames describing one event (except if the origin is specified but this already constitutes as an event).

[2] We have assumed A, B, C and D are constants (does not depend on x, t, x' or t'), since all points in space are indistinguishable, the presence of dependency on absolute coordinates would mean that the absolute location in space (and not just the relative position) is important. In short, the absolute coordinate is arbitrarily defined based on the choice of the origin; only the change in coordinates is what we concern.

[3] We also assumed that Δx and Δt are linear functions of $\Delta x'$ and $\Delta t'$. This can be justified by noting that any interval can be built up from a series of many infinitesimal intervals. But for infinitesimal intervals $\Delta x'$ and $\Delta t'$, any nonlinear terms such as $(\Delta t')^2$ are negligible. So if we add up all the infinitesimal intervals to obtain the given interval, we will be left with only the linear terms.

The first two facts translate into (dropping the Δs)

1. $\Delta x' = 0$ and $v = \frac{\Delta x}{\Delta t} \implies \frac{B}{C} = v$.
2. $\Delta x = 0$ and $-v = \frac{\Delta x'}{\Delta t'} \implies A = C$.

The third fact can be used by asking how fast does a person in S see a clock in S' ticks and the analogous question how fast does a person in S' see a clock in S ticks?

The answer to the first question can be solved by letting $\Delta x' = 0$, which gives $\Delta t = \Delta t'$. The second question can be solved by letting $\Delta x = 0$, which gives $\Delta t' = \frac{\Delta t}{A - Dv}$. Comparing the two equations, we have $D = \frac{1}{v}(A - \frac{1}{A})$.

The fourth fact can be used to say that if $\Delta x' = c\Delta t'$, then $\Delta x = c\Delta t$. Solving for A gives $A = \gamma$.

Gathering all the constants found, we have the Lorentz transformation

$$\begin{aligned} \Delta x &= \gamma(\Delta x' + v\Delta t') \\ \Delta t &= \gamma(\Delta t' + v\Delta x'). \end{aligned} \tag{1.2}$$

Solving for $\Delta x'$ and $\Delta t'$, we have the same form as above except v is replaced by $-v$ (as expected, as there is no superior frame).

If two events occur simultaneously in frame S, then $\Delta t = 0$, so $(\Delta x', \Delta t') = (\gamma\Delta x, -\gamma v\Delta x)$. Thus the two events does not occur simultaneously in frame S' and length is contracted.

If two events occur at the same place in S', then $\Delta x' = 0$, so $\Delta x = \gamma v\Delta t', \Delta t = \gamma\Delta t'$. Thus time is dilated.

Note that proper length (the length of an object in its rest frame) is always longest and the proper time (the time interval in which the events occurs at the same place) is always to shortest. This distinguish whether we should multiply or divide by γ.

Both longitudinal and transverse velocity addition can be easily done with Lorentz transformation, as

$$\begin{aligned} v_x &\equiv \frac{\Delta x}{\Delta t} = \frac{\Delta x' + v\Delta t'}{\Delta t' + v\Delta x'} = \frac{v_x' + v}{1 + vv_x'} \\ v_y &\equiv \frac{\Delta y}{\Delta t} = \frac{\Delta y'}{\gamma(\Delta t' + v\Delta x')} = \frac{v_y'}{\gamma(1 + vv_x')}, \end{aligned} \tag{1.3}$$

where v is the relative speed of the two frames and $\mathbf{v}'$ is the velocity of the object observed in the primed frame.

As will be proved in later chapters, the quantity

$$(\Delta s)^2 = (\Delta t)^2 - (\Delta x)^2 - (\Delta y)^2 - (\Delta z)^2 \tag{1.4}$$

is an invariant (does not depend on which reference frame you are measuring it from).

2

4-Vectors and Dynamics

2.1 Definition and Examples of 4-Vectors

Definition 2.1.1. *A 4-tuple, $A = (A_0, A_1, A_2, A_3)$ is called a "4-vector" if the A_i transform between frames in the same way that $(\Delta t, \Delta x, \Delta y, \Delta z)$ do. In other words, A is a 4-vector if it transforms according to the Lorentz transformation*

$$\begin{pmatrix} A_0 \\ A_1 \\ A_2 \\ A_3 \end{pmatrix} = \gamma \begin{pmatrix} 1 & v & 0 & 0 \\ v & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} A'_1 \\ A'_2 \\ A'_3 \\ A'_4 \end{pmatrix}. \tag{2.1}$$

The displacement 4-vector between two events that are infinitesimally close together is

$$dS \equiv (dt, dx, dy, dz) \tag{2.2}$$

Dividing the displacement 4-vector with the proper time $d\tau = \frac{dt}{\gamma}$ (which is an invariant since it is independent of the fram in which it is measured), we obtain

$$V \equiv \frac{dS}{d\tau} = \gamma(1, \frac{dx}{dt}, \frac{dy}{dt}, \frac{dz}{dt}) = \gamma(1, \mathbf{v}). \tag{2.3}$$

We can then take the derivative of the velocity 4-vector with respect to τ. The result is a four vector because the numerator is just an infinitesimal change of the displacement 4-vector (which is also a 4-vector due to linearity) and the denominator is an invariant. So

$$A \equiv \frac{dV}{d\tau} = \gamma(\frac{d\gamma}{dt}, \frac{d(\gamma\mathbf{v})}{dt}) = (\gamma^4 v\dot{v}, \gamma^4 v\dot{v}\mathbf{v} + \gamma^2\mathbf{a}) \tag{2.4}$$

Multiply the velocity 4-vector by another invariant, mass m, we get the energy-momentum 4-vector

$$P \equiv mV = \gamma m(1, v) = (E, \mathbf{p}). \tag{2.5}$$

The force 4-vector is then the derivative of the momentum 4-vector with respect to τ. This gives

$$F \equiv \frac{dP}{d\tau} = \gamma(\frac{dE}{dt}, \mathbf{f}), \tag{2.6}$$

where $\mathbf{f} \equiv \frac{d(\gamma m \mathbf{v})}{dt}$ is the usual 3-force defined as the rate of change of momentum.

In terms of acceleration 4-vector, we can write

$$F = mA = m(\gamma^4 v\dot{v}, \gamma^4 v\dot{v}\mathbf{v} + \gamma^2 \mathbf{a}) = m(\gamma^4 v_x a_x, \gamma^4 a_x, \gamma^2 a_y, \gamma^2 a_z), \tag{2.7}$$

if $\mathbf{v} = v_x \hat{\mathbf{x}}$. Combining the two forms of F, we see that the 3-force is

$$\mathbf{f} = m(\gamma^3 a_x, \gamma a_y, \gamma a_z). \tag{2.8}$$

2.2 Properties of 4-Vectors

If A and B are 4-vectors, then $C \equiv aA + bB =$ is also a 4-vector, this is due to the linearity of Lorentz transformation. For example, the first component of C is

$$C_0 \equiv aA_0 + bB_0 = a(A'_0 + vA'_1) + b(B'_0 + vB'_1) = (aA'_0 + bB'_0) + v(aA'_1 + bB'_1) = C'_0 + vC'_1 \tag{2.9}$$

which is the correct transformation for the first component of C under Lorentz transformation.

The inner product of two 4-vectors A and B are defined as

$$A \cdot B \equiv A_0 B_0 - A_1 B_1 - A_2 B_2 - A_3 B_3. \tag{2.10}$$

It can be shown (by direct substitution) that the inner product is an invariant. This invariance under Lorentz Transformation (by multiplying the Lorentz matrix) is analogous to the invariance of the scalar (dot) product $\mathbf{A} \cdot \mathbf{B} = |A||B| \cos\theta$ under rotations (by multiplying the rotational matrix) in 3-space (since neither $|\mathbf{A}|, |\mathbf{B}|$ nor θ is affected by rotation).

The norm of a 4-vector is then defined as the square root of the inner product of A with itself:

$$|A|^2 \equiv A \cdot A \equiv A_0 A_0 - A_1 A_1 - A_2 A_2 - A_3 A_3 = A_0^2 - |\mathbf{A}|^2. \tag{2.11}$$

The invariance of the norm $\sqrt{A \cdot A}$ is analogous to the invariance of the norm $\sqrt{\mathbf{A} \cdot \mathbf{A}} \equiv |\mathbf{A}|$ for rotations in 3-space.

If a certain one of the components of a 4-vector is zero in every frame, then all four components are zero in every frame. This is analgous to how if someone comes along and says the she has a vector in 3-space that has no x component, no matter how you rotate the axes, then you would certainly say that the vector must be the zero vector. The situation in Lorentzian 4-sapce is the same, because all of the coordinates get intertwined with each other in the Lorentz (and rotation) transformations.

2.3 Force and Acceleration

In frame S' of a given particle travelling at speed v in the x direction relative to the lab frame S, we have the force 4-vector

$$F' = \gamma(\frac{dE'}{dt}, \mathbf{f}') = \gamma(\frac{dm}{dt}, \mathbf{f}') = (0, \mathbf{f}') \tag{2.12}$$

Transfomring F' into F, we get

$$\begin{aligned} F_0 &= \gamma \frac{dE}{dt} = \gamma(F_0' + vF_1') = \gamma v f_x', \\ F_1 &= \gamma f_x = \gamma(F_1' + vF_0') = \gamma f_x', \\ F_2 &= \gamma f_y = F_2' = f_y', \\ F_2 &= \gamma f_z = F_2' = f_z'. \end{aligned} \tag{2.13}$$

Thus we get

$$\frac{dE}{dt} = v f_x', \quad f_x = f_x', \quad f_y = \frac{f_y'}{\gamma}, \quad f_z = \frac{f_z'}{\gamma}. \tag{2.14}$$

The acceleration 4-vector, on the other hand, is

$$A' = (\gamma^4 v'\dot{v}', \gamma^4 v'\dot{v}'\mathbf{v}' + \gamma^2 \mathbf{a}') = (0, \mathbf{a}'), \tag{2.15}$$

Transforming A' into A, we get

$$\begin{aligned} A_0 &= \gamma^4 v a_x = \gamma(A_0' + vA_1') = \gamma v a_x', \\ A_1 &= \gamma^4 a_x = \gamma(A_0' + vA_0') = \gamma a_x', \\ A_2 &= \gamma^2 a_y = A_2' = a_y', \\ A_3 &= \gamma^2 a_z = A_3' = a_z'. \end{aligned} \tag{2.16}$$

Thus we get

$$a_x = \frac{a_x'}{\gamma^3}, \quad a_y = \frac{a_y'}{\gamma^2}, \quad a_z = \frac{a_z'}{\gamma^2}. \tag{2.17}$$

Example: Acceleration for Circular Motion

Question: A particle moves with constant speed v around the circle

$x^2 + y^2 = r^2$, $z = 0$, in the lab frame. At the instant the particle crosses the negative y axis, find the 3-acceleration and 4-acceleration in both the lab frame and the instantaneous inertial frame of the particle.

Solution: The 3-acceleration in S is simply

$$\mathbf{a} = (0, \frac{v^2}{r}, 0). \tag{2.18}$$

There is nothing fancy going on here; the standard nonrelativistic proof of the centripetal acceleration works just fine again in the relativistic case.
The 4-accleration in S is then

$$A = (0, 0, \gamma^2 \frac{v^2}{r}, 0). \tag{2.19}$$

Transforming to S', we have

$$A' = (0, 0, \gamma^2 \frac{v^2}{r}, 0). \tag{2.20}$$

So the 3-acceleration in S' is then the space part of A'

$$\mathbf{a}' = (0, \gamma^2 \frac{v^2}{r}, 0). \tag{2.21}$$

Of course, we can arrive at this result with a simple time-dilation argument, since

$$a'_y = \frac{d^2 y'}{dt'^2} = \frac{d^2 y}{(\frac{dt}{\gamma})^2} = \gamma^2 a_y = \gamma^2 \frac{v^2}{r}. \tag{2.22}$$

Note that in special relativity, if a statement has any chance of being true in all frames, it must involve only 4-vectors. Consider a 4-vector equation $A = B$ that is true in frame S. Aplying Lorentz transformation on both sides, we get $A' = B'$ so the law is therefore also true in frame S'. However, $\mathbf{f} = m\mathbf{a}$ cannot be a physical law, since the two sides of this equation transform differently when going from one frame to another, so the statement cannot be true in all frames.

All of this is exactly analogous to the situation in 3-dimensional space. In Newtonian mechanics, $\mathbf{f} = m\mathbf{a}$ is a possible law (and indeed a physcal law), because both sides are 3-vectors. But $\mathbf{f} = m(2a_x, a_y, a_z)$ has no chance of being a physical law, because the right hand side is not a 3-vector; it depends on which axis you label as the x axis.

Physcial laws may also take the form of scalar equations, such as $P \cdot P = m^2$ A scalar is by definition a quantity that is frame independent. So if a scalar statement is true in one inertial frame, then it is true in all inertial frames.

2.4 Collisions

From the energy-momentum 4-vector we see that the momentum and energy of a particle governed by special relativity are

$$\mathbf{p} = \gamma m\mathbf{v} \text{ and } E = \gamma m. \tag{2.23}$$

Taylor expanding γ , we get

$$\begin{aligned} \mathbf{p} &= (1 + \mathcal{O}(v^2))m\mathbf{v} = m\mathbf{v} + \mathcal{O}(v^3) \\ E &= (1 + \frac{1}{2}v^2 + \mathcal{O}(v^4))m = m + \frac{1}{2}mv^2 + \mathcal{O}(v^4), \end{aligned} \tag{2.24}$$

which, according to correspondence principle, must give the same answer in Newtonian physics.

Note that the energy of a particle even in Newtonian physics consists of the rest mass energy term m.

In an elastic collision, no kinetic energy is transformed into other forms of energy, so the mass is conserved (To not introduce another degree of freedom, we assume that individual masses are conserved, not just the sum of them, so particles cannot transfer their internal energy to one another), so we can neglect writing down the rest mass energies when applying the law of conservation of energy.

An inelastic collision, on the other hand, converts kinetic energy to internal energy (e.g., thermal energy) which increases the masses of the particles[1], so the mass is not conserved and thus the Newtonian kinetic energy $\frac{1}{2}mv^2$ is not conserved.

Similar to the relation $E = \frac{p^2}{2m}$ or $p = \sqrt{2mE}$ in Newtonian physics, we can relate the energy and the momentum of a particle by this "very-important-relationship"

$$E^2 = p^2 + m^2. \tag{2.25}$$

Whenver we know two of the three quantities E, p and m, this eqution gives you the third.

There is no need to use 4-vectors to solve collision problems, they can be always be solved solely by writing down the energy and momentum conservation laws, as well as the "very-important-relationship" but they often prove to be useful bookeeping tools, since the norm of the energy-momentum 4-vector of a particle (note that this does not work for a collection of particles) is simply the mass of the particle:

$$P \cdot P = E^2 - \mathbf{p}^2 = m^2. \tag{2.26}$$

The power of writing energy and momentum of a particle in a single energy-momentum 4-vector can be illustrated by the example below:

Example: Decay of a Particle

Question: A particle with mass M and energy E decays into two identi-

[1]All internal energy such as thermal, rotational and potential energy contributes to the rest mass of an object. In fact, the vast majority of the mass of an atom is due to the internal energy between quarks that make up the nucleus rather than the rest mass of the quarks themselves.

cal particles. In the lab frame, one of them is emitted at a 90° angle. What are the energy of the created particles?

Solution: Let P, P_1 and P_2 be the energy-momentum 4-vectors of the decayed particle, the particle emitted at 90° angle, and the remaining particle respectively.
By conservation of energy and momentum,

$$\begin{aligned} P &= P_1 + P_2 \\ P - P_1 &= P_2 \\ (P - P_1)(P - P_1) &= P_2 \cdot P_2 \\ P^2 - 2P \cdot P_1 + P_1^2 &= P_2^2 \\ M^2 - 2EE_1 + m^2 &= m^2 \\ E_1 &= \frac{M^2}{2E}. \end{aligned} \tag{2.27}$$

Again, the energy-momentum vector introduce no new physics, it simply write the conservation laws in a way that help simplify the algebra using the "norm squared equals mass" property.

The only non-trivial property of the energy-momentum vector is that since the sum of the energy-momentum 4-vectors of a collection of particles is also a 4-vector due to linearity,

$$\left(\sum P\right)^2 = \left(\sum E\right)^2 - \left(\sum \mathbf{p}\right)^2 \tag{2.28}$$

is also an invariant. Note that the sums are taken before squaring in the above equation. If we sqaure before the addition it would simply give the sum of the squares of the masses (which of course is also an invariant but is trivial why).

2.5 Force

The Newton's law in special relativity is

$$F = \frac{dp}{dt} = \frac{d(\gamma m v)}{dt} = m(\dot{\gamma} v + \gamma \dot{v}) = m(\gamma^3 v a v + \gamma a) = ma\gamma(\gamma^2 v^2 + 1) = \gamma^3 ma, \tag{2.29}$$

for one-dimensional case while the work energy theorem is

$$\int_{x_1}^{x_2} F dx = \int_{x_1}^{x_2} (\gamma^3 ma) dx = m \int_{v_1}^{v_2} \gamma^3 v dv = m \int_{\gamma_1}^{\gamma_2} d\gamma = E_2 - E_1 = \Delta E. \tag{2.30}$$

which is the same as in nonrelativistic physics.

The Newtons's law can be easily generalized into two dimensions by

$$F = \frac{d\mathbf{p}}{dt} = \frac{d(\gamma m\mathbf{v})}{dt} = \frac{d}{dt}\left(\frac{m(v_x, v_y)}{\sqrt{1 - v_x^2 - v_y^2}}\right) = m(\gamma^3 a_x, \gamma a_y). \tag{2.31}$$

The asymmetric arises from the fact that $v_x = v$ and $v_y = 0$ at the start by default.

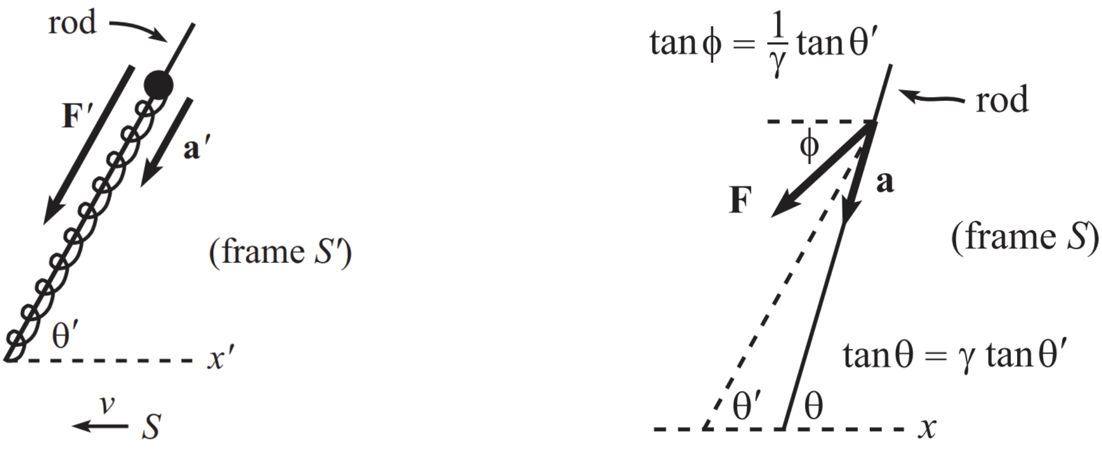
Example: Bead on a Rod

Question: A spring with a tension has one end attached to the end of a rod, and the other end attached to a bead that is constrained to move along the rod. The rod makes an angle θ' with respec to the x' axis and is fixed at rest in the s' frame (see fig. 2.1a). Right after the bead is release, find the directions of the rod, the acceleration of the bead and the force on the bead in frame S.

Solution: The horizontal span of the rod is decreased by a factor γ due to length contraction while the vertical span is unchanged. So we have $\tan\theta = \gamma \tan\theta'$.
The acceleration, on the other hand, must point along the rod, simply because the bead always lies on the rod (and the rod itself does not accelerate).
The force, however, does not point along the rod. This is because the y component of the force on the bead is decreased by a factor of γ while the x component is unchanged. So $\frac{F_y}{F_x}$ is smaller than $\frac{F'_y}{F'_x}$ by a factor γ. So we have $\tan\phi = \frac{1}{\gamma}\tan\theta'$.
To double check the direction of acceleration,

$$\frac{a_y}{a_x} = \gamma^2 \frac{F_y}{F_x} = \gamma^2 \frac{\frac{F'_y}{\gamma}}{F'_x} = \gamma \tan\theta' = \tan\theta. \tag{2.32}$$

Note that the rod does not exert a force of constraint. The bead does not need to touch the rod in S' so it does not need to touch it in S. **F** simply does not have to be collinear with **a** in relativistic physics.
The situation in frame S is shown in fig. 2.1b.



(a) (b)

Figure 2.1