Special Relativity: Lecture 2

1 The Lorentz transformation

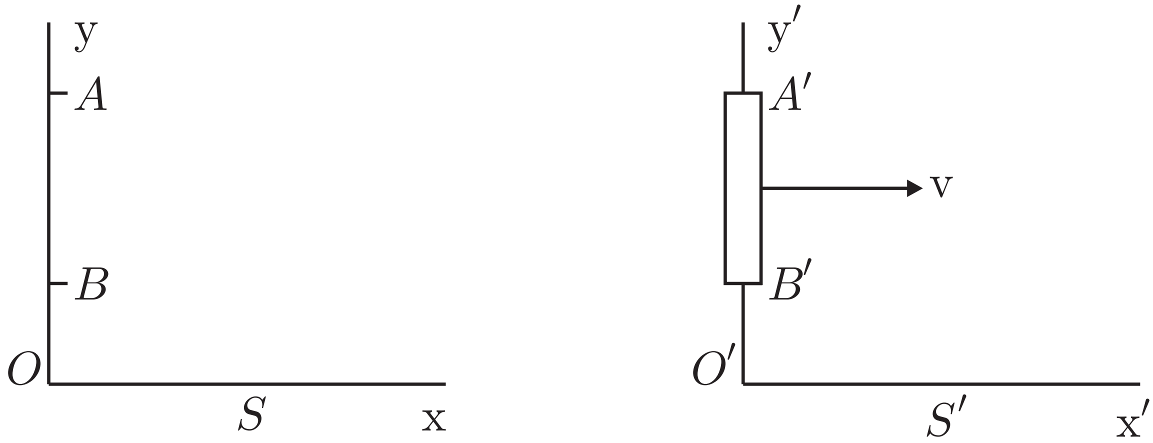
There are many ways to derive the form of the Lorentz transformation. We shall start from the second of Einstein's postulates, namely that the speed of a light signal, c, is the *same* in all IFs. Take the standard IFs S and S' and consider the profile of an outgoing spherical wave as seen in both frames. Remember that O and O' coincide at $t = t' = 0$, so in S and S',

$$x^2 + y^2 + z^2 - c^2t^2 \; = \; x'^2 + y'^2 + z'^2 - c^2t'^2 = 0. \qquad (II.1)$$

We shall assume that there is a linear mapping between the coordinates in S and those in S' so for example:

$$x' = a_{11}x + a_{12}y + a_{13}z + a_{14}ct.$$

This assumption is equivalent to assuming that space is homogeneous (`do you agree?`) and will also ensure that a constant velocity in S maps into a constant velocity in S'. S and S' are related by a relative velocity v directed along Ox, so coordinates y and z are transverse to the direction of motion and we consider these first.



Transverse Coordinates

Consider a rigid rod $A'B'$ fixed along $O'y'$ in S' (see figure above). As S' passes the stationary frame S, marks A and B are made along Oy. The length ℓ' of the rod as measured in S' can be compared to an identical rod of length ℓ as measured in S. Either $\ell = \ell'$ or $\ell < \ell'$ or $\ell > \ell'$. Suppose that $\ell < \ell'$ i.e. the stick appears longer in S'. Now consider the same problem as an observer at rest in S'. By the same method we find $\ell > \ell'$. But Einstein's first postulate was that physical results should be the same in all IFs! The only way to avoid a contradiction is thus to accept that $\ell = \ell'$. The same argument will apply to any direction perpendicular to the relative velocity. (Notice that it is crucial for this argument to be valid that the measurement of the length of $A'B'$ in S be made *at one instant of time*.)

Thus, $y' = y, \quad z' = z$ and eq.(II.1) reduces to:[1]

$$x^2 - c^2t^2 = x'^2 - c^2t'^2 = 0. \tag{II.2}$$

Therefore the transformation we seek is:

$$\begin{aligned} x' &= a_{11}x + a_{14}ct, \\ ct' &= a_{41}x + a_{44}ct. \end{aligned} \tag{II.3}$$

The coefficients a_{ij} are independent of x, t, x', t' but may depend on v. The origin O' $(x' = 0)$ moves with velocity v in O, thus $a_{11}x + a_{14}ct = 0$ and,

$$\frac{\mathrm{d}x}{\mathrm{d}t}|_{x'=0} = v = -\frac{a_{14}c}{a_{11}}. \tag{II.4}$$

So $x' = a_{11}(x - \beta ct)$ where $\beta = v/c$. Similarly consider the velocity $-v$ of O $(x = 0)$ wrt S', $x' = a_{14}ct$ and $ct' = a_{44}ct$ so $x' = a_{14}ct'/a_{44}$, thus,

$$\frac{\mathrm{d}x'}{\mathrm{d}t'}|_{x=0} = -v = \frac{a_{14}c}{a_{44}}.$$

Using eq.(II.4) we thus have that $a_{11} = a_{44}$. Using these results in eq.(II.2) gives,

$$(a_{11}(x - \beta ct))^2 - (a_{41}x + a_{11}ct)^2 = x^2 - c^2t^2,$$

so equating coefficients ,

$$\begin{aligned} a_{11}^2 - a_{41}^2 &= 1, \\ a_{11}^2\beta^2 - a_{11}^2 &= -1, \\ -2\beta a_{11}^2 - 2a_{41}a_{11} &= 0. \end{aligned} \tag{II.5}$$

These three eqs have the consistent solution: $a_{11} = (1 - v^2/c^2)^{-1/2}$ and $a_{41} = -\beta a_{11}$. It is standard to write γ for a_{11}; doing so, the full Lorentz transformation is:

$$\begin{aligned} x' &= \gamma(x - \beta ct), \\ ct' &= \gamma(ct - \beta x), \\ y' &= y, \\ z' &= z, \end{aligned} \tag{II.6}$$

where $\beta = v/c$ and $\gamma = (1 - \beta^2)^{-1/2}$. The inverse transformation is,

$$\begin{aligned} x &= \gamma(x' + \beta ct'), \\ ct &= \gamma(ct' + \beta x'), \\ y &= y', \\ z &= z', \end{aligned} \tag{II.6'}$$

which can be derived by direct inversion of eq.(II.6) or by the replacement $v \to -v$.

Note that as $v/c \to 0$, $\gamma \to 1$, $x' \to (x - vt)$ and $t' \to (t - vx/c^2)$ so the Galilean transformation is recovered in the limit of velocities small compared to that of light.

[1] Check out the more sophisticated argument based on rotational invariance given by Bowler.

Another way of stating the content of the Lorentz transformation is that it is the coordinate transformation that leaves the *interval* S_{12}^2 between two events invariant, where,

$$S_{12}^2 \;=\; (x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2 - c^2(t_1 - t_2)^2. \qquad (II.7)$$

Lorentz transformations can be shown to satisfy the axioms of group theory namely, the existence of an identity, inverse and the combination of two LTs to give a transformation that is also an LT. `(Those who are mathematically inclined may wish to pursue this further.)`

The form of the LT given in eq.(II.6) can be written in a form that is reminiscent of a coordinate rotation, namely:

$$x' = x\,\cosh\phi \;-\; ct\,\sinh\phi,$$
$$ct' = -x\,\sinh\phi \;+\; ct\,\cosh\phi, \qquad (II.8)$$

where ϕ is the *boost parameter* and is given by $\tanh\phi{=}\beta$.

For transformations between IFs with parallel coordinate axes but with a relative velocity $\mathbf{v}$ not necessarily parallel to an axis, the LT takes the form:

$$\mathbf{x}' = \mathbf{x} + \frac{(\gamma-1)}{v^2}\mathbf{v}(\mathbf{v}.\mathbf{x}) - \gamma\mathbf{v}ct,$$
$$ct' = \gamma(ct - \mathbf{v}.\mathbf{x}/c). \qquad (II.9)$$

`(Check the result by identifying the components parallel and perpendicular to v.)` The most general transformation between IFs with arbitrary orientations is constructed from the parallel LT of eq.(II.6) (known as a boost) and 3-rotations of the spatial coordinates.

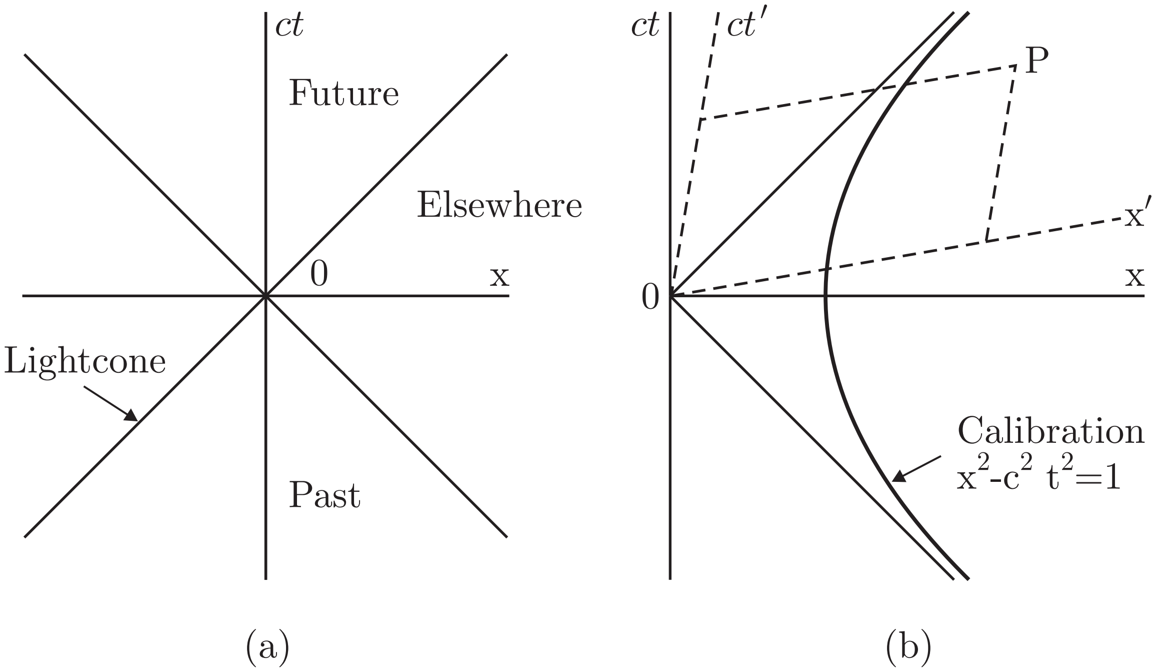
The major *conceptual* difference between Einstein's relativity and the Galilean relativity of Newtonian mechanics is that time is no longer absolute. Consider two events (x_1', t_1') and (x_2', t_2') in S'. What is the time ordering of the corresponding events in S? Applying the inverse transformation (II.6'), we get

$$c(t_2 - t_1) = \gamma[c(t_2' - t_1') + \beta(x_2' - x_1')].$$

Now even if $t_2' > t_1'$ in S', $(t_2 - t_1)$ in S can take any sign or even be zero depending on the sign and magnitude of $(x_2' - x_1')$. However if the event (x_2', t_2') is *caused* by the event (x_1', t_1') then $|x_2' - x_1'|/(t_2' - t_1') \;<\; c$ and $t_2 > t_1$ *always*, i.e. causality is preserved as it must. This leads us to the next topic.

2 Minkowski Diagrams and The Light Cone

We have already seen that the interval S_{12} (II.7) between two events is an invariant. The sign of the square of this quantity provides a useful way of classifying events in the *space-time diagram* shown below (for clarity we have only shown one space variable).



- $S_{12}^2 = 0$ — the light cone:
  This describes the trajectory of light rays from the origin.

- $S_{12}^2 < 0$ — timelike separation:
  For such events there exists a LT with $v < c$ such that $\Delta x' = 0$ i.e. we can go to a frame in which the events occur at the same place but separated in time — such events can be causally connected and the time ordering is necessarily preserved in any frame.

- $S_{12}^2 > 0$ — spacelike separation:
  For such events there exists a LT with $v < c$ such that $\Delta t' = 0$ i.e. we can go to a frame in which the events occur at the same time but are separated in space. Such events cannot be causally connected and the time-ordering is not necessarily preserved in all frames.

Clearly the light cone separates the two types of events, the timelike events being inside the light cone. The light cone itself can be separated into a "future" and a "past" wrt the origin. Points spacelike separated from the origin are said to be "elsewhere" (or the "unknowable present").

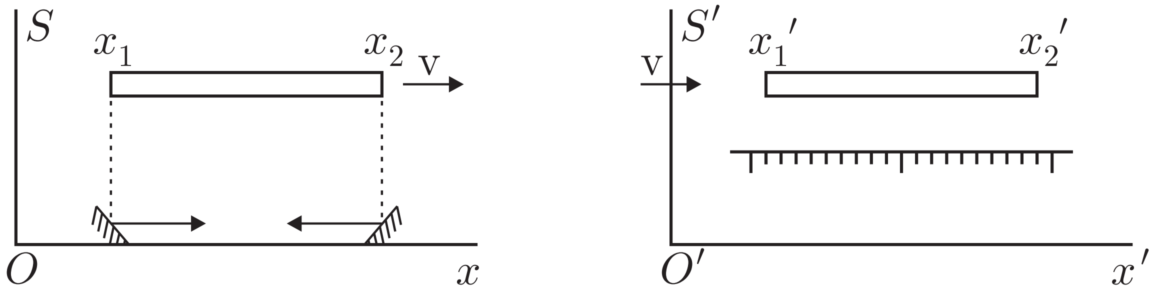
Minkowski diagrams can also be used to show the relationship between two frames conected by a simple LT. We have already mentioned that the LT can be written in the form of a pseudo-rotation. Consider panel (b) above. The horizontal axis represents the space axis (Ox in S) and the vertical axis the temporal (ct) axis in S. S' is moving with a velocity $\mathbf{v}$ wrt S along Ox thus the Ox' axis is represented by an a line (dashed) at an angle to the x-axis.The ct' axis is also at an angle to ct but notice that the x' and ct' axes are no longer at right angles. The diagram shows how the space and time coordinates of a point are to be read off in the frames S and S'. To make quantitative use of such diagrams it is necessary to calibrate them because unit length is not preserved in the mapping between S and S' in the diagram. Calibration is provided by hyperbolae representing the interval (essentially $\mathrm{S}_{12}^2 > 0$) e.g. $x^2 - c^2t^2 = 1$ as shown above. While not a substitute for calculation Minkowski diagrams can provide an intuitive insight into various problems.

3 Fitzgerald Contraction

The length of a rigid rod at rest can be measured by comparison with a metre rule. Consider such a rod at rest in S' with length $\Delta x' = x_2' - x_1'$. An observer in S will measure the length of the rod as the difference of the coordinates x_2 and x_1 measured *simultaneously* at time t in S. Thus,

$$\begin{aligned} x_1' &= \gamma(x_1 - vt), \\ x_2' &= \gamma(x_2 - vt), \\ \Delta x' &= \gamma \Delta x, \end{aligned}$$

or $\Delta x = \Delta x'\sqrt{1 - v^2/c^2}$, i.e. the length of the moving rod is measured to be less by the stationary observer in S than the rest length. This asymmetry is due to the fact that the two types of measurement are inherently different — see diagram below. The Fitzgerald contraction has so far not proved to be amenable to direct experimental verification.



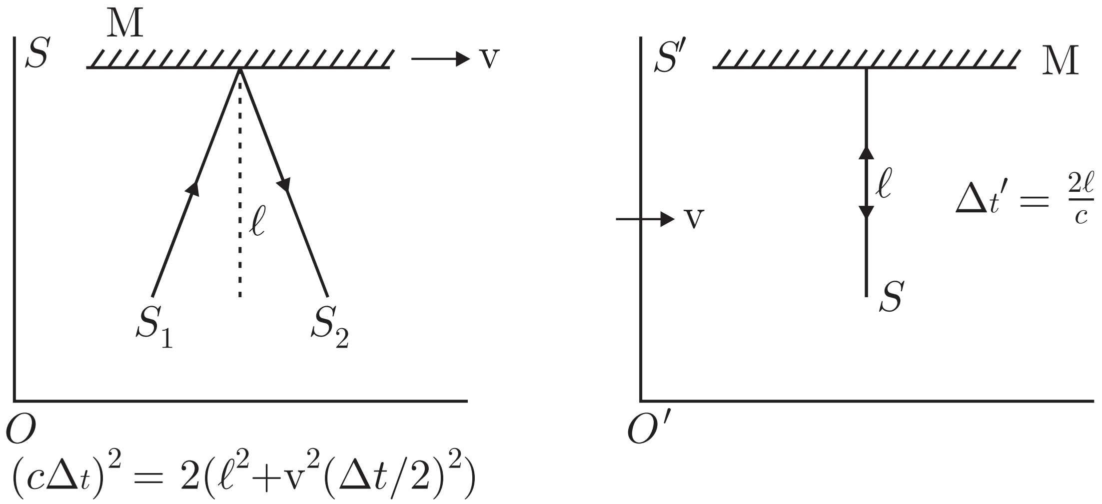
In S the length $x_2 - x_1$ is measured at the same instant, i.e. the coordinates of the end points in S are (x_1, t), (x_2, t). Use the direct LT since t is constant in S.

4 Time Dilation

One of the most remarkable predictions of special relativity is time dilation. Despite being verified daily in accelerator laboratories it has given rise to many so-called "paradoxes". Consider identical clocks in S and S', both at their respective coordinate origins. Since the LT is a linear transformation we can write down immediately how the time interval measured at a fixed point transforms:

$$\Delta t \;=\; \gamma \Delta t'.$$

So the clock in the moving frame S' appears to be running slow in S.



Again there is an asymmetry in the way the two measurements are made. In the moving IF S', a light signal is bounced off a mirror back to the same point to give a transit time $\Delta t'$. Seen from the stationary frame S the light signal will not be received at the point of emission and it is simple to show (see figure) that the measured time difference in S is $\gamma \Delta t'$.

For a clock at rest in S' two events will have coordinates (x', t'_1), (x', t'_2). Since x' is fixed in S', use the inverse LT to transform the time interval to S.

5 Proper Time

It would be convenient to have a measure of time that is *invariant* between IFs and this is provided by the *proper time*. Consider two events occuring at the origin in S' with a time interval $\Delta t'$ (e.g. clock ticks). Now transform the two events to S and use the invariance of S_{12}^2:

$$c^2 \Delta t'^2 \;=\; c^2 \Delta t^2 - \Delta x^2 - \Delta y^2 - \Delta z^2.$$

Divide by $c^2 \Delta t^2$ and take the square root to get[2],

$$\Delta \tau \;=\; \Delta t' \;=\; \Delta t \sqrt{1 - v^2/c^2}, \qquad (II.10)$$

and the result that $\Delta t'$ is an invariant. Proper time is usually designated by τ and is the time measured in the *rest frame* of a sequence of events.

[2]Because $\mathrm{d}\mathbf{x}/\mathrm{d}t = \mathbf{v}$ the velocity of S' relative to S, $\Delta r^2/\Delta t^2 = v^2$.

6 Tests of Time Dilation

Although it is not easy to use ordinary clocks to check time dilation, Nature has provided us with unstable particles with characteristic half-lives that can be used instead. Time dilation implies that rapidly moving "clocks" run slow in the stationary laboratory, thus a rapidily moving unstable particle will travel a much greater distance in the lab before it decays.

According to the radioactive decay law, the number of particles remaining after time t' in frame S' in which the sample is at rest, is given by

$$N(t') = N(0)e^{-t'/\tau_m}, \qquad (II.11)$$

where τ_m is the mean lifetime in frame S', or, more correctly the proper lifetime in the rest-frame of the parent particle. Consider the decay from a frame S wrt which S' has a velocity $\mathbf{v}$. In S, $t = \gamma t'$, so the decay law appears to an observer in S to be,

$$N(t) = N(0)e^{-t/\gamma\tau_m}$$

which has the same form as eq.(II.11) but with an apparent mean lifetime of $\gamma\tau_m$. (For example a 50 GeV/c momentum K^+ meson with τ_m of 1.24×10^{-8} s will travel a mean distance of 376 m before decaying, compared to 3.7 m that would have been expected in the absence of time dilation. Here $\gamma = 101.2$.)

One of the first demonstrations of time dilation was made by Rossi and Hall (Phys. Rev. **59** (1941) 223) when they compared the cosmic ray flux at sea level with that measured on the top of Mt Washington. The sea level flux was much greater than could be accounted for by the known muon lifetime and the flux measured at high altitudes. However the most accurate checks come from high energy particle accelerators, in particular the CERN muon storage-ring experiment (Bailey *et al*, Nature **268** (1977) 301). (This was designed to check Quantum Electrodynamics but as by-product gives a very accurate check of time dilation.) Muons of momentum 3.094 GeV/c ($\gamma = 29.33, \beta = 0.9994$) were circulated in a storage ring of 14 m diameter. The muons were produced by the decay of charged pions ($\pi \to \mu + \nu$), which were produced by 10 GeV/c proton interactions. Muons β-decay to electrons by $\mu \to e + \nu_e + \nu_\mu$ and the decay electrons are registered by suitable counters. The measured lifetime in the lab ($\tau_{\rm L}$) was 64.368(29) μs. Using the known γ factor the muon proper lifetime can be calculated and compared to the measured rest liftime (τ_0) of 2.19711(8) μs giving, $(\tau_0 - \tau_L/\gamma)/\tau_0 = (2 \pm 9) \times 10^{-4}$. This result shows that time dilation is correct to 0.1%. Many other species of particles have been used to give checks to about 1%.

Does the storage ring experiment really check special relativity? Surely the circulating muon is in an accelerated frame `(try to estimate the centrifugal acceleration)`? The answer to both questions is yes and we shall return to the matter later — what is vital is that the functioning of the "clock" should not be affected by the acceleration. This is indeed the case with the "nuclear clocks" used here.

For the particular case of muons an experimental check can be made to show that their behaviour as clocks is unaffected by acceleration. When brought to rest in a low-Z material such as carbon, μ^- particles are captured into the lowest atomic orbit and experience very large accelerations. Nevertheless the μ^- proper lifetime agrees with that of the μ^+ which is not captured (see Bowler p.56 for more details.)

A very different sort of experiment was performed by Hafele and Keating (Science **177** (1972) 168) who compared the time intervals measured by a caesium atomic clock flown around the world (on commercial jets) with the intervals measured by an identical clock at rest (wrt Earth). There are two effects predicted — time dilation of special relativity, and a general relativistic effect of comparable magnitude which predicts that clocks in a higher gravitational potential will run slower. The experimental results were in good agreement with the overall theoretical predictionas we will discuss later.