Lectures 5 & 6 – The Special Theory of Relativity

Physics for Pedestrians

30th July, 2019

I have Dr. Bikram Phookun (from St. Stephen's College, Delhi) to thank for a large part of these notes.

1 Properties of Space and Time

Imagine that you are in the metro, underground. The PA system announces that the doors will open to the right at your stop. You are the only passenger getting off. As you get ready to disembark, you find that you don't know which side is on the right! And the reason for your confusion, you realise, is that you don't know which way the train is moving.

You may also have observed that so long as the train moves steadily through the tunnel it is virtually impossible to determine how fast it is moving (unless the sound that it makes is proportional to its speed).

In other words, in the perceptions of our environment there is nothing that reveals the speed and direction – i.e. our velocity – at which we are moving, at which we are moving – unless there are external indicators. In other words, it is only with reference to external markers that we can determine our velocity (unlike acceleration, which you experience when the train starts or stops, or when the train turns at constant speed, in a very palpable fashion). This is a property of space and time that goes much deeper than our perceptions: there is in fact no experiment that can reveal the absolute velocity at which the equipment used to perform the experiment is moving. In some fundamental sense, therefore, absolute velocity has no meaning: velocity is always relative. This is called the *Principle of Relativity.*

To this property of space let us add three others that are more obvious: (i) all points in space are equivalent; (ii) all directions from a point are equivalent; and (iii) all instants are equivalent. The first implies that if an experiment is carried out at two different points in space, it produces the same result. The second implies that the results of an experiment do not depend on the orientation of the equipment. The third implies that they also do not depend on whether the experiment is done today or tomorrow.

2 The Laws of Physics and the Principle of Relativity

The laws of physics are theoretical constructs that encapsulate the results of experiments. They too must therefore show the same invariances that experiments do. Since a physical law is a relationship between quantities, what we mean by invariance is invariance of this relationship, i.e. of the form of the law. This is best understood through an example. Consider the application of Newton's second law of motion to a body of mass m_1 interacting gravitationally with another body of mass m_2 separated by it by the distance r_{12}:

$$m_1\mathbf{a_1} = \frac{G\,m_1\,m_2}{r_{12}^2}\hat{\mathbf{r}}_{\mathbf{12}} \tag{1}$$

Notice that, within the common-sense way of the world within which this relationship holds, the acceleration of an object is independent of the velocity of the observer, so long as the latter is constant. So are masses, and distance. So is the direction of force, given by $\hat{\mathbf{r}}_{\mathbf{12}}$. Thus, both sides of the equation are independent of velocity of the observer with respect to the system, so long as it is constant. The form of the equation is thus the same for all observers moving at constant velocities with respect to the system.

3 Inertial Frames of Reference

Notice that to verify that the form of a law of physics is the same for two observers, we will need to figure out things like the distance between objects and the acceleration of objects, with respect to these two observers. Such measurements will need measurements of position and time. It is useful to imagine a coordinate system equipped with devices – rulers and clocks – that allow measurements of position and time, and stuck to the observer and moving with him/her; this is the rest frame of the observer.

We have described space as being homogeneous, isotropic, and time-invariant. Notice that this is true only for an observer moving at a constant velocity. The moment an observer accelerates, one direction becomes special. (When a train accelerates or decelerates, you feel a force in one direction even though there is nothing physically attracting or repelling you in that direction – so that direction becomes special.) A frame of reference in which all positions, directions, or moments are equivalent is called an *Inertial Frame*. It can equally be thought of as a frame in which any force is the result of a physical interaction rather than just the result of acceleration of the observer. Because of the principle of relativity any other frame moving an constant velocity with respect to an inertial frame is also an inertial frame.

The problem of verifying whether the laws of physics have the same form to all inertial observers thus reduces to that of verifying that they have the same form in all inertial frames of reference. And, since comparison of the views of different observers requires position and time measurements made by them, we must begin with the comparison of position and time measurements in different inertial frames of reference.

Let us begin by comparing how a single point in space at a single moment in time – what in physics is called an *event* – looks in different frames of reference. An event in a frame is the triplet of its position coordinates and its time coordinate: (x, y, z, t). For simplicity we will deal with situations in which nothing changes with y or z, and thus each of our events will be indicated by (x, t).

4 Common-sense Transformations

Consider two inertial frames of reference S and S', moving with velocity v with respect to each other.

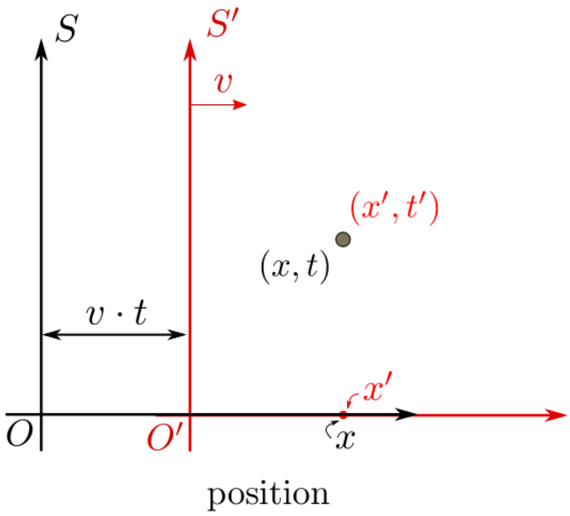


Figure 1: Two inertial frames S (black) and S' (red)

Common sense tells us that

$$\begin{aligned} x' &= x - vt \\ t' &= t \end{aligned} \tag{2}$$

and, in the other direction,

$$\begin{aligned} x &= x' + vt' \\ t &= t'. \end{aligned} \tag{3}$$

If a object is moving along the x direction with velocity u as observed in S and with velocity u' as observed in S', then

$$u' = u - v \tag{4}$$

Question: Use the common-sense transformations for coordinates and time to derive the transformation for speed given above. (Speed is the rate of change of position; so you will have to consider two successive positions and two successive moments.)

Question: Use the transformation for speed to show that the acceleration of an object is the same in both S and S'.

4.1 Invariance of Length

Our everyday experience leads us to believe that the length of an object is the same whether it is stationary or moving. We therefore expect this invariance of length to be consistent with common-sense relativity. The length of a rod laid along the x axis in the frame in which it is at rest is the difference between the coordinates of its ends. When the the rod is moving with with respect, however, its length is the difference between the coordinates of its ends only if these coordinates are measured simultaneously. (Imagine trying to measure the length of a moving train, and you'll immediately understand why this so.)

Suppose that a rod is at rest in S'. It stretches between x'_A and x'_B. Thus its length is

$$L' = x'_B - x'_A. \tag{5}$$

Now consider how this looks to an observer in S. As mentioned above we must now measure the coordinates of the two ends simultaneously, i.e. $t_A = t_B$.

$$L = (x_B - x_A)_{t_A = t_B} = (x'_B - x'_A) + v(t'_B - t'_A) \tag{6}$$

where t'_A and t'_B are the moments at which the coordinates of the ends of the rod are measured. You can see that in general, as explained above, $L = L' = x'_B - x'_A$ only if $t'_B = t'_A$, i.e. if the measurements of the coordinates are simultaneous in the rest frame of the rod as well as in the rest frame. Thus, for the rod to have the same length in all frames, simultaneity in one frame must imply simultaneity in other frames as well. And of course within common-sense relativity this is true.

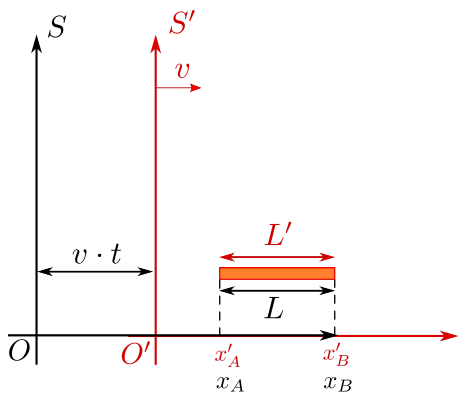


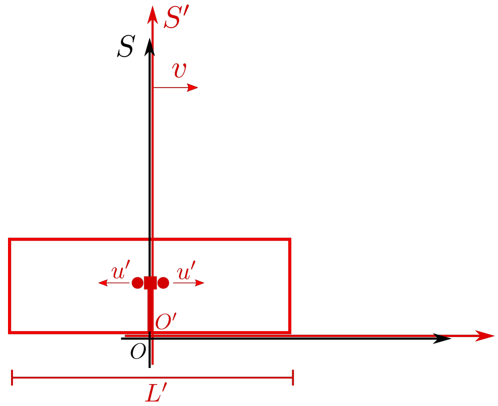
Figure 2: Length measurements in S (black) and S' (red): L' is the length measured by an observer in S' and L is the length measured by an observer in S.

Question: Consider three different inertial frames and check that length of a rod at rest in one of them has the same length in all three frames.

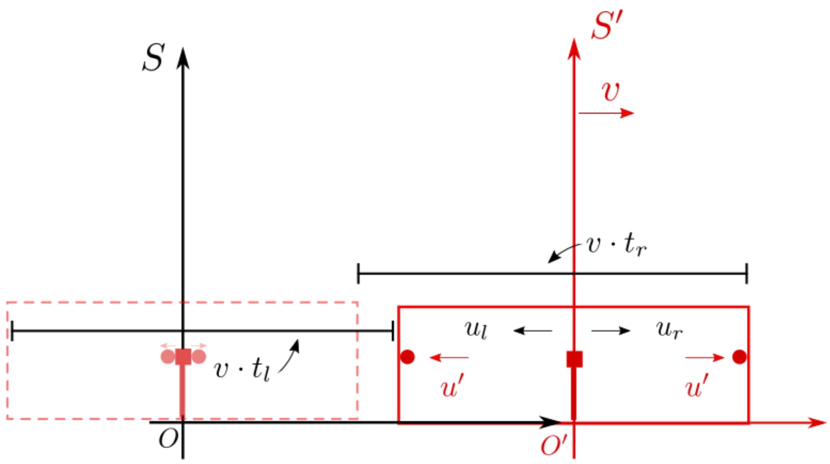
4.2 Invariance of Simultaneity

We have seen that the frame-independence of length requires the frame-independence of simultaneity. We must therefore ask what the frame-independence of simultaneity depends on. To investigate this, let us consider the following thought experiment (a favourite theoretical device of Einstein's). A box of length L is at rest in frame S', which is moving with respect to S with velocity v. The centre of the box coincides with the centre of the common $x - x'$ axis, as shown in Figure (3a) below. At $t = t' = 0$, two objects/signals are emitted with equal speeds u' (as measured in S') in opposite directions from the centre.

In S' the distance travelled in each direction to the wall is $L'/2$, and therefore the time taken is $L'/2u'$. The strikes on the walls are obviously simultaneous in frame S'. This implies that they must be simultaneous in S as well.



(a) Emission ($t = t' = 0$)



(b) Impact

Figure 3: Simultaneity in S (black) and S' (red)

In S the object/signal travelling to the right must travel a distance $L/2 + vt_r$, where t_r is the time it takes to travel from the point of emission to the right wall. The object/signal travelling to left travels $L/2 - vt_l$, where t_l is the time it takes to travel from the point of emission to the left wall. The travel times are given by

$$t_r = \frac{L/2 + vt_r}{u_r} \tag{7}$$

and

$$t_l = \frac{L/2 - vt_l}{u_l}. \tag{8}$$

Notice that we must have $t_r = t_l$, since the events were simultaneous in the rest frame of the box, and since simultaneity, in common-sense relativity, is frame-independent. But you can see that that depends on how velocities add. Unless

$$t_r\,(u_r - v) = t_l\,(u_l + v), \tag{9}$$

we will not be able to satisfy this condition. Fortunately, within this framework, $u_r = u' + v$ and $u_l = u' - v$, and so everything works out.

For the enthusiastic reader

Question: In the thought experiment let the emission take place from a point one-third of the way between the walls. Then the time interval between the strikes on the right and left walls is non-zero. Check that it is still the same in both frames of reference.

4.3 The Velocity Addition Law

The thought experiment above demonstrates how important the velocity-addition law is. It shows that every object or signal must obey this law – else we would violate the requirement of common-sense relativity that events simultaneous in one frame are simultaneous in all frames. Notice that if $u' = u - v$ for everything that moves, then anything that moves with a finite velocity *must* have different velocities in two inertial frames moving with respect to each other in the same direction as the object/signal. The only velocity that can be the same in all frames is ∞.

4.4 The Problem with the Common-sense Transformations

What is the difficulty with this? The difficulty lies in the fact that that there does exist a frame-independent velocity. Experiments show that light is observed to travel at the same speed, irrespective of the speed of the emitter or of the observer.

The only conclusion that can be drawn from this is that coordinate transformations arrived at using common sense are incorrect. Yet, they are so obvious that if we try to modify the sense (called common) that was used to arrive at them, we may go further astray. One way out may be to approach the problem more formally – to ask ourselves what characteristics are desirable in a set of transformations relating inertial frames, and see if we can thus arrive at the correct transformations.

5 The Desirable Properties of Coordinate Transformations

Since common sense has lead us astray, we will instead try to identify what is properties the transformation must have to make physical sense. What we are looking for is a set of transformations between the coordinates in S and S' that has the following properties: (i) one event is mapped onto one event, since one event *is* one event, no matter which frame we view it from; (ii) it is invertible, and the inverse transformation corresponds to changing v to $-v$, since if S' moves at v with respect to S, then S must move at $-v$ with respect to S'; and (iii) the parameters of the transformation obey the basic properties of space, i.e. they are independent of position, direction, and instant in time, implying that these parameters can only depend on the relative velocity. The simplest transformation that do this are linear transformations:

$$\begin{aligned} x' &= a_{11}\,x + a_{12}\,t \\ t' &= a_{21}\,x + a_{22}\,t \end{aligned} \tag{10}$$

where the a_{ij} depend only on the relative velocity v. Our job is to figure out the four a_{ij} using physical arguments. (Reminder: we are using a convention in which when $t = t' = 0$, $x = x' = 0$, i.e. the origins of the frames coincide at the initial instant in both frames.). The inverse transformations are

$$\begin{aligned} x &= \frac{a_{22}}{D}\,x' - \frac{a_{12}}{D}\,t' \\ t &= -\frac{a_{21}}{D}\,x' + \frac{a_{11}}{D}\,t' \end{aligned} \tag{11}$$

where $D = a_{11}a_{22} - a_{12}a_{21}$.

Question: Check that the inverse transformations are correct.

When we say that S' moves at a constant velocity v with respect to S along the x axis, what we mean is that the rate at which the origin of S' moves with respect to the origin of S is v; thus the x coordinate of the origin of S' is vt. The origin of S' means $x' = 0$. Thus,

$$\text{when } x' = 0, \text{ we have that } x = v\,t.$$

Putting this into the first of the equations above, we get $a_{12} = -v\,a_{11}$.

Now let us look at S from S'. Now, using the same argument as above, we see that the x' coordinate of the origin of S is $-v\,t'$. (Notice that it's now t', and t and t' may not any longer be the same.) Similar to before,

$$\text{when } x = 0, \text{ we have that } x' = -v\,t'.$$

Putting this into the second set of equations above, we get $a_{12} = -v\,a_{22}$. Combining this with the result in the last paragraph, we get $a_{11} = a_{22}$.

Question: Check that $a_{11} = a_{22}$.

Now let us write rewrite the transformations:

$$
\begin{aligned}
x' &= a_{11}\left(x - v\,t\right)\\
t' &= a_{11}\left(\frac{a_{21}}{a_{11}}\,x + t\right) = a_{11}\left(\frac{a_{21}}{a_{12}}\,\frac{a_{12}}{a_{11}}\,x + t\right)\\
&= a_{11}\left(-\,\frac{a_{21}}{a_{12}}\,v\,x + t\right)
\end{aligned}
\tag{12}
$$

Now let us get the velocity transformation law corresponding to this set of transformations. To get velocity, we need two events – two positions x_1 and x_2 and the two corresponding times t_1 and t_2. We will write $\Delta x = x_2 - x_1$ etc. Thus

$$
\frac{\Delta x'}{\Delta t'} = \frac{\frac{\Delta x}{\Delta t} - v}{1 - v\,\frac{a_{21}}{a_{12}}\,\frac{\Delta x}{\Delta t}}
\tag{13}
$$

or

$$
u' = \frac{u - v}{1 - \frac{u\,v}{k^2}},
\tag{14}
$$

where $k^2 = a_{12}/a_{21}$. What we have above is the general velocity addition law consistent with the Principle of Relativity and the properties of space and time

Question: Check (i) that in this velocity-addition law the dimension of k is velocity; and (ii) that if $u' = k$ so is u.

The fact that if $u' = k$, then $u = k$, implies that in this more general set of transformations of coordinates between inertial frames, there exists a *finite* frame-independent velocity. You may remember that in the common-sense transformations, where the velocity transformation law was $u' = u - v$, the only frame-independent velocity was ∞.

This general velocity transformation law allows both possibilities: (i) if $k = \infty$ then we are just back to the common-sense of Galilean transformations; and (ii) if k is finite, then we have another set of transformations called the *Lorentz Transformations*. It is up to nature to decide which of these two sets of transformations is correct. Experiments show that there exists a finite frame-independent velocity – the velocity of light c. (So from now on we will write c in place of k.) Thus the correct coordinate transformations between inertial frames are the Lorentz transformations.

5.1 Why do the Galilean Transformations seem Obvious?

This questions remains. The answer is that the frame-independent velocity that we see in nature, the velocity of light, is very, very large: $c = 300,000$ km/s! When the velocities of objects and observers are small, i.e. if u and v are small, compared to c, then we effectively have the Galilean transformations. Since our common sense has developed within this domain, it is not surprising that it arrives naturally at the Galiliean transformations rather than the Lorentz transformations.

5.2 The Last Constant

We still need to figure out a_{11}. Inserting what we have just worked out, we have

$$\begin{aligned} x' &= a_{11}\left(x - v\,t\right) \\ t' &= a_{11}\left(t - \frac{v}{c^2}\,x\right) \end{aligned} \tag{15}$$

and

$$\begin{aligned} x &= \frac{a_{11}}{D}\left(x' + v\,t'\right) \\ t &= \frac{a_{11}}{D}\left(t' + \frac{v}{c^2}\,x'\right). \end{aligned} \tag{16}$$

Since the only difference between the forward and inverse transformations is that v goes to $-v$, we have $a_{11} = a_{11}/D$. Thus $D = 1$. Also

$$D = a_{11}\,a_{22} - a_{12}\,a_{21} = a_{11}^2\left(1 - \frac{a_{12}\,a_{21}}{a_{11}^2}\right) = a_{11}^2\left(1 - \frac{v^2}{c^2}\right). \tag{17}$$

It follows that $a_{11} = 1/\sqrt{1 - v^2/c^2}$. This is universally denoted by the symbol γ. Thus the Lorentz transformations are

$$\begin{aligned} x' &= \gamma\left(x - v\,t\right) \\ t' &= \gamma\left(t - \frac{v}{c^2}\,x\right) \end{aligned} \tag{18}$$

Question: (The second part of this exercise is for those who know matrices.) Show that you can also write this set of transformations as

$$\begin{aligned} x' &= \gamma\left(x - \frac{v}{c}\,c\,t\right) \\ c\,t' &= \gamma\left(c\,t - \frac{v}{c}\,x\right) \end{aligned} \tag{19}$$

What is the advantage of writing the transformations like this?

Question: Now show that they can be written in the following form

$$\begin{pmatrix} x' \\ c\,t' \end{pmatrix} = \gamma \begin{pmatrix} 1 & -v/c \\ -v/c & 1 \end{pmatrix} \begin{pmatrix} x \\ c\,t \end{pmatrix} \tag{20}$$

For the enthusiastic reader

What is the advantage of writing the transformations in this matrix form? Consider three frames of reference S_o, S_1, and S_2. S_1 moves at velocity v_{10} to the right with respect to S_o, S_2 moves at velocity v_{20} to the right with respect to S_o and velocity v_{21} with respect to S_1. First write coordinates in S_2 with respect to those in S_1, using the matrix form give above. Then, on the right hand side of the equation, substitute the transformations for the coordinates of S_1 with respect to those in S_o, again using the matrix form. You will now have a relation between the coordinates in S_2 and those in S_o that has the product of two matrices in between. But you can also write the same transformation directly between S_2 and S_o, using a single matrix. Do that, and compare the two equations. What do you see? Explore.

6 Some Consequences of Special Relativity

We can now go back to the two thought experiments we conducted at the start. As we argued before, the fact that an observer in S and one in S' agreed on simultaneous events was because of the "common-sense" velocity addition law that we now know to be false. And similarly, the fact that these observers also agreed upon the lengths of objects was because they agreed on simultaneous events. Since the velocity addition law is no longer the same, it makes sense to re-examine these two concepts: length and time-intervals.

Before we start, let us write down the Lorentz Transformations:

$$\begin{aligned}
&\text{(A)} \quad \Delta x' = \gamma\left(\Delta x - v\Delta t\right)\\
&\text{(B)} \quad \Delta t' = \gamma\left(\Delta t - \frac{v}{c^2}\Delta x\right)\\
\\
&\text{(C)} \quad \Delta x = \gamma\left(\Delta x' + v\Delta t'\right)\\
&\text{(D)} \quad \Delta t = \gamma\left(\Delta t' + \frac{v}{c^2}\Delta x'\right)
\end{aligned} \tag{21}$$

6.1 Lengths

Let us consider, as before, that the object we are measuring is at rest in the frame S', and its length is being measured both from S (in which it is moving to the right with a velocity v) and S' in which it is at rest.

The observer in S requires to measure the endpoints of the object **simultaneously** in her frame of reference, as otherwise the object would move between measurements. In other words, for $(x_B - x_A)$ to be the length, we require that $\Delta t = t_B - t_A = 0$.[1] Thus, we need to find a relation between Δx and $\Delta x'$, when $\Delta t = 0$. We refer to Equation (21), and see that **(A)** is the transformation we should use, as it relates these quantities.

[1] Note that we are not placing any condition on $\Delta t'$. It may not be (and isn't!) zero.

$$\Delta x' = \gamma\left(\Delta x - v\Delta t\right)$$

$$\Delta\ x'\Big|_{\Delta t=0} = \gamma\left(\Delta x\Big|_{\Delta t=0} - v\ \Delta t\Big|_{\Delta t=0}\right)$$

$$L' = \gamma L$$

Thus, the length that an observer measures when she is at rest with respect to the object (i.e. sitting in S') L' is always greater than L, since $\gamma > 1$.

Thus, an observer sitting in S, with respect to whom the object is moving at a constant velocity will measure a length L which is **shorter**: *lengths contract*!

Remember: to the person sitting in S', the table will not look contracted at all! It is only with respect to an observer in S that the object will appear contracted.

6.2 Time-intervals

In order to understand time-intervals, we perform another thought experiment. Consider a 'light' clock, which we make using a rod and an emitter and detector of light. A pulse of light is emitted at one end of the rod, reflected at the other end, and detected back where it was emitted. Let us place this clock in the frame S' where it is moving with respect to S with a velocity v.

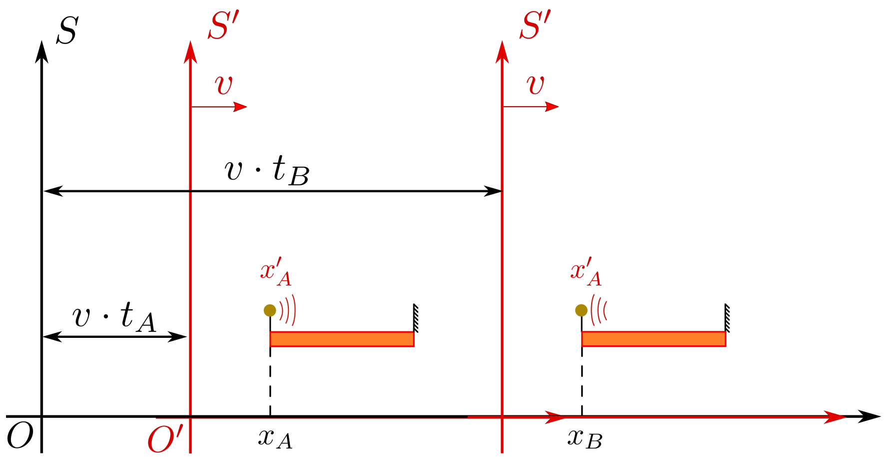


Figure 4: A light clock at rest in S', observed from S. The light pulse emitted at one end of the rod is reflected at the other end and detected back where it started from. The coordinates of emission and detection in S' are the same (x'_A) but different when viewed from S, as the clock is moving with respect to S.

In the frame S' an observer will see the light traverse the length of the rod twice and be detected after some time $\Delta t'$. What does an observer in S see? In order to relate these two observations, let us consider the two events:

Event	**In** S	**In** S'
Emission:	(x_A, t_A)	(x'_A, t'_A)
Detection:	(x_B, t_B)	(x'_A, t'_B)

Notice that since the rod is moving with respect to an observer in S, emission and detection occur in different points of space (x_A and x_B), but for an observer in S', both emission and detection occur at the *same* point (x'_A)! Thus, $\Delta x' = 0$, even though $\Delta x \neq 0$.

We would like to relate the time intervals that someone in S' and someone in S measure. i.e. we would like to relate Δt to $\Delta t'$, and we know that $\Delta x' = 0$ (in other words, the two events occur at the same point in space according to someone in S', the frame that it motionless with respect to the clock). We look at Equation (21) and see that the equation relating the quantities we are interested in is **(D)**.

$$\Delta t = \gamma \left(\Delta t' + \frac{v}{c^2} \Delta x' \right)$$

Since $\Delta x' = 0$,

$$\begin{aligned} \Delta t &= \gamma \left(\Delta t' + \frac{v}{c^2} \cancel{\Delta x'} \right) \\ \Delta t &= \gamma \Delta t' \end{aligned}$$

Thus, $\Delta t > \Delta t'$, in other words intervals of time observed in S would appear to take **longer** than the same intervals as measured in S': *time dilates!*

> **Remember:** as the person in S looks at her counterpart in S', she will see time slow down for the person in S', and his lengths contract. However, according to the person is S', he will look at the world as if it were normal, and would rather see time slow down for the observer in S (moving backwards) and *her* lengths contract!
>
> Their views of the world are **symmetric**, and there is no way for either of them to say which of them is "actually" moving. (You should now be convinced that this question makes no sense.)

Question: You may perform the same analysis above using only the fact that lengths contract, and that the speed of light is constant:

1. Start by observing the event in S': ask yourself how long the light takes to get reflected and come back as a function of L' and c; call this $\Delta t'$.
2. Now observe the same event in S, and calculate the time interval Δt as a function of L and c.
3. c being constant for both observers, relate $\Delta t'$ and Δt by relating L and L'.