Special Relativity, Causality and Quantum Mechanics – 1

Guruprasad Kar, Samir Kunkri and Sujit K Choudhary

We discuss the significance of Einstein's second postulate of the special theory of relativity (STR) stipulating the constancy of the speed of light in vacuum. The causality that follows from the STR may be a more general principle to organize our knowledge of all phenomena. In particular, quantum dynamics can be derived using the causality principle. Causality does not prohibit quantum mechanics to be non-local, and this has implications in the real world.

All are from Physics and Applied Mathematics Unit, ISI, Kolkata
Left. G Kar – current research interest is the foundation of quantum mechanics and quantum information theory in general and quantum non-locality and entanglement in particular.
Right. S Kunkri – current research interest is the role of entanglement in quantum information processing and the connection between quantum operations and causality.
Centre. S K Choudhary – current research interest is the study of various versions of Bell's theorem on nonlocality of quantum mechanics for various quantum systems.

Introduction

In 1905, Einstein changed the face of the world of physics by making two postulates about nature, which gave birth to the special theory of relativity (STR). After various electromagnetic phenomena got a concrete and unified shape in the hands of Maxwell, physics entered a period of crisis which remained for decades. The existence of electromagnetic waves and a fixed speed for it followed from Maxwell's equations. Interestingly, the various results for the speed of light from experiments coincided with that value, and it was also shown that light is an electromagnetic wave.

The problem arose when it was shown that Maxwell's equations change their form when written in other inertial frames following the Galilean transformations which keep the laws of mechanics invariant in all inertial frames. So the principle of relativity seemed to be at stake with the entry of the electromagnetic theory in physics.

In the face of the crisis when physicists could not find a consistent solution, Einstein put forward the following

Keywords
Second postulate of relativity, causality, no-cloning theorem, entangled state, quantum non-locality.

two postulates:

1. The laws of physics take the same form in all inertial frames.

2. The speed of light in vacuum is the same in all inertial frames.

These two postulates changed the whole conceptual structure of time and space introduced by Newton. The new concept of space and time, shaped by these postulates, remained equally consistent with all observed phenomena. Not only that, in the process of modifying mechanics to make it compatible with a new space-time transformation (Lorentz transformation), it predicted new results which were eventually verified.

Importance of the Second Postulate

We start with the comment made in [1] and elsewhere, which says "Actually, the second postulate follows from the first since the laws of physics include Maxwell's equations of electromagnetism which contain the speed of light." We contend with this comment and argue for the inevitability of the second postulate once the first postulate has been proposed. For that purpose we first present the concrete situation that existed before 1905.

1) No observable phenomena, either in mechanics or in electromagnetism, can distinguish one inertial frame from another.

2) Maxwell's theory predicts a constant value c for the speed of propagation of electromagnetic waves in free space. This is not compatible with the Galilean transformations which were known to apply between inertial frames. Thus it seems that if Maxwell's theory holds in one inertial frame, it

No observable phenomena, either in mechanics or in electromagnetism, can distinguish one inertial frame from another.

will not hold in another inertial frame which has a relative motion with respect to the first.

A theory called the ether theory attempted to resolve the dilemma by postulating that Maxwell's theory is correct only in a frame which is fixed in a medium called the ether.

3) Experimental results contradict the existence of any medium like the ether. This, along with Maxwell's equations, established that electromagnetic waves move in the absence of any medium.

4) Maxwell's theory predicts a constant speed of light in vacuum and, if the theory is correct and relativity holds, then the democracy which holds amongst all inertial frames according to the Galilean transformations cannot be maintained. Galilean relativity, which kept the laws of mechanics invariant, was at stake.

In this situation, the appropriate question is whether the principle of relativity is true in the absence of any counter example. Any fundamental theory will have to answer this question. But how? Einstein did it by asserting the truth of the relativity principle as a postulate which draws its strength from the absence of any counter examples in realistic experiments. Now, what is the content of the first postulate? It says that 'laws' that do not have the same form for all inertial observers cannot be correct. But the first postulate does not single out any existing law to be correct. The first postulate is only one amongst the various tests to judge the correctness of a physical theory. If a theory is incompatible with the first postulate, one can at first sight rule it out.

Newton's laws were widely accepted as the laws of mechanics. They are also compatible with the first postulate provided one believes in the correctness of the Galilean transformation equations for space and time.

Einstein proposed his second postulate since the constancy of the speed of light follows from a well established theory, namely, Max- well's theory of electromagnetism, and experiments do not contradict it.

Similarly, Maxwell's equations were very well established as equations for electromagnetic phenomena. These are also not in clash with the first postulate, not according to the Galilean transformation equations, but rather according to a different set of transformation equations called the Lorentz transformation equations.

But how can there be different sets of transformation equations of space and time for a set of two observers (one moving with respect to other with a constant velocity)? The transformation equations between two inertial observers should be unique and should be independent of any particular physical theory. But how to find them? Nature gives no further clue apart from some physical quantities and the laws involving these quantities.

So once the first postulate has been made, the option that remains open is to find the correct transformation, either by (a) declaring some existing laws to be correct, or (b) proposing the invariance of a quantity involving time and space. Einstein opted for the second by declaring the constancy of the speed of light which of course follows if Maxwell's equations are taken to be true.

So the steps taken by Einstein are the following:

- **Einstein proposed his second postulate since the constancy of the speed of light follows from a well established theory, namely, Maxwell's theory of electromagnetism, and experiments do not contradict it.**

  There is some confusion regarding the results of laboratory experiments performed by Fizeau, Foucault and Michelson. Some people think that these experiments establish a constant speed for the propagation of light. Then one may ask: what was the need for postulating the constancy of the speed of light? But it should be noted that all these laboratory experiments study the two-way speed of light. Since such measurements are obviously possible with just one clock, the synchronization problem does not arise. Constancy of the speed of light in any direction is obviously a possible explanation of experimental results, but not the only possible one.

- Once the second postulate has been made, the constancy of the speed of light can be used to define synchronization of clocks in a single inertial frame, and it will easily follow that simultaneity is a frame dependent concept.

Einstein's first postulate says that the necessary condition for a physical theory to be correct is that it has to satisfy the principle of relativity.

How is the synchronization procedure adopted? Consider two points where clocks are located; at one point there is a light source, and at another point there is a mirror to reflect light back to its source. Now, if light starts from the source at time t_1 and returns there at time t_2 after reflection, the moment registered by the clock at the mirror is to be set equal to $t = t_1 + (t_2 - t_1)/2$.

One has to note that this synchronization procedure presumes that the speed of light is constant in any direction. However this constancy of the speed of light in any direction cannot be established by experiments, since our clocks at different points in space are synchronized asssuming a fixed value c for the speed of light.

We summarize the argument in the following way. Einstein's first postulate only says that the necessary condition for a physical theory to be correct is that it has to satisfy the principle of relativity. But the postulate itself is powerless to distill the correct physical theory from the existing theories since the relativity principle (*i.e.*, the first postulate) does not suggest the appropriate space-time transformation laws. Here arises the necessity of the second postulate which, along with the first, can determine the transformation equations. **Hence the second postulate has a status which**

is independent of the first, *i.e.*, this postulate is not to be proven, like the axioms in geometry, but can be exploited as we wish for any purpose.

Special Relativity and Causality

As we have seen, Einstein, instead of choosing Maxwell's laws to be correct, took one consequence of them, *i.e.*, the speed of light is constant. Why? It is very difficult to state the reason definitely. But he realized that a new feature of light, (*i.e.*, the quantum aspect of light where he himself played an important role) and the dynamics of atomic particles could not be explained in terms of mechanics and electrodynamics. New theories were expected to emerge. So he wanted to distill from electromagnetic phenomena the key feature whose significance transcends Maxwell's equations and which might be taken as a viable principle for organizing our knowledge of all phenomena, including mechanics, optics and new theories which were yet to appear.

These two postulates force us to have a fresh look at space and time. Newtonian physics was based on the notion both of absolute three-dimensional space and of time as a separate entity which was again absolute. But an entirely different picture of space and time emerges from the two relativistic postulates. Time is no longer absolute and the same is true for space too. Moreover, these two (*i.e.*, space and time) are not two separate entities; rather they are interwoven and form a four-dimensional space-time continuum.

Although this space is different from our earlier Newtonian space, it has a few features which are like the earlier one, *i.e.*, the Newtonian one. These features are:

Causality means that cause precedes effect. This implies an ordering in time.

(i) The interval (see *Box* 1) between two events remains invariant in this space under a transformation from one inertial frame to another.

Box 1.

An event (for example, the explosion of a bomb or the decay of a photon or the collision of two balls) can be specified completely by four numbers: the three spatial coordinates and the time of its occurrence.

Let (x_1, y_1, z_1, t_1) and (x_2, y_2, z_2, t_2) be the space-time coordinates assigned to two events by an observer in an inertial frame S. Then the quantity I, where

$$I = c^2(t_2 - t_1)^2 - (x_2 - x_1)^2 - (y_2 - y_1)^2 - (z_2 - z_1)^2, \qquad \text{(a)}$$

is called the interval between these two events (also called the Minkowskian squared distance) in frame S.

Let us introduce $w = ict$. Then equation (a) reads as

$$-I = (w_2 - w_1)^2 + (x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2.$$

At this point, we may interpret x, y, z and w as cartesian coordinates in a 4-dimensional space in which $\sqrt{-I}$ is like the distance between the points (x_1, y_1, z_1, t_1) and (x_2, y_2, z_2, t_2) A directed segment joining these points is like a vector having $(w_2 - w_1), (x_2 - x_1), (y_2 - y_1), (z_2 - z_1)$ as the components in a 4-dimensional space. This space is also called Minkowski space in honour of the Russian-German mathematician Hermann Minkowski who first realized it.

Notice that I can be positive, negative or zero.

If $I > 0$, then we call the interval time-like, and the two events are said to be time-like separated events. The corresponding displacement is called a time-like vector in this space.

If $I = 0$ then we call the interval light-like, and the two events are said to be light-like separated events. The corresponding displacement is called a light-like vector in this space.

(ii) For time-like and light-like separated events (see *Box* 1) in this space, causality is preserved for two inertial observers.

Causality means that cause precedes effect. This implies an ordering in time. If two events are time-like or light-like separated and are causally connected in one frame, then they will remain so connected in every frame. We name this the principle of causality. Mathematically speaking, if the space-time coordinates of two events when measured in frame S satisfy the inequality

$$c^2(t_2-t_1)^2-(x_2-x_1)^2-(y_2-y_1)^2-(z_2-z_1)^2\geq 0, \quad (1)$$

and if $(t_2 - t_1) > 0$, then $(t'_2 - t'_1) > 0$.

Here t'_2 and t'_1 are the moments of occurrence of these events as measured by another inertial observer S' who is in uniform relative motion with respect to observer S.

Lorentz Transformation Equations Respect Causality

Let (x, y, z, t) be the space-time coordinates assigned to an event by an observer in an inertial frame S. S' is another inertial frame moving with respect to S with a velocity v along the common $X - X'$ axis. An observer in S' assigns $(x', y'\ z', t')$ as the space-time coordinates to the same event. Then the transformation

$$x' = \gamma(x - vt), \quad y' = y, \quad z' = z, \quad t' = \gamma(t - \frac{v}{c^2}x), \quad (2)$$

where $\gamma = 1/(\sqrt{1-\frac{v^2}{c^2}})$, is known as the special Lorentz transformation. These equations contain all the kinematical information in the special theory of relativity. For example, the preservation of causality and the invariance of the interval are consequences of these equations.

Now applying a Lorentz transformation to the two events (x_1, y_1, z_1, t_1) and (x_2, y_2, z_2, t_2) in the S frame with $t_2 > t_1$, we can write

$$c(t'_2 - t'_1) = \gamma c\ (t_2 - t_1) - \gamma\beta\ (x_2 - x_1), \quad (3)$$

where $\beta = v/c$.
Now it is obvious that

$$[-\gamma\beta\ (x_2-x_1)]^2 \leq (\gamma\beta)^2[(x_2-x_1)^2+(y_2-y_1)^2+(z_2-z_1)^2]. \quad (4)$$

Using equation (1) and the fact $\gamma^2 - \gamma^2\beta^2 = 1$, we get

$$[-\gamma\beta\ (x_2 - x_1)]^2 \leq (\gamma^2 - 1)c^2(t_2 - t_1)^2 < \gamma^2 c^2(t_2 - t_1)^2 \tag{5}$$

which implies

$$\gamma c\ (t_2 - t_1) - \gamma\beta\ (x_2 - x_1) > 0.$$

So from equation (3) we get

$$c(t'_2 - t'_1) > 0,$$

i.e.,

$$(t'_2 - t'_1) > 0.$$

This is what we mean by causality preservation.

General Lorentz Transformation

The special Lorentz transformation takes on a simpler and much used form (at least by mathematicians) when they are written in terms of the new quantities

$$x^0 = ct, \quad x^1 = x, \quad x^2 = y, \quad x^3 = z.$$

With this notation, the Lorentz transformation equations as given by equation (2) take the form

$$(x^0)' = \gamma(x^0 - \beta x^1), \ \ (x^1)' = \gamma(x^1 - \beta x^0),$$
$$(x^2)' = x^2, \ \ (x^3)' = x^3, \tag{6}$$

or in matrix form,

$$(x^\mu)' = \Sigma_{\nu=0}^{3}(L^\mu_\nu x^\nu). \tag{7}$$

Here L^μ_ν are the elements of the Lorentz transformation matrix

$$L = \begin{pmatrix} \gamma & -\gamma\beta & 0 & 0 \\ -\gamma\beta & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

The special Lorentz transformation (mentioned above) is only a special case of a more general Lorentz transformation. The general Lorentz transformation is derived by assuming invariance of the Minkowskian distance and linearity. Linearity ensures that if an object is not accelerated in frame S, it is unaccelerated in frame S' too. The new coordinates $(x^\mu)'$ of a vector are connected with the old coordinates x^ν via the linear relation

$$(x^\mu)' = \Sigma_{\nu=0}^{3}(L_\nu^\mu x^\nu), \tag{8}$$

where L_ν^μ are the elements of the matrix of the general Lorentz transformation

$$L = \begin{pmatrix} L_0^0 & L_1^0 & L_2^0 & L_3^0 \\ L_0^1 & L_1^1 & L_2^1 & L_3^1 \\ L_0^2 & L_1^2 & L_2^2 & L_3^2 \\ L_0^3 & L_1^3 & L_2^3 & L_3^3 \end{pmatrix}$$

Now the invariance of the interval I, *i.e.*, $(x_0)^2 - \sum_{i=1}^{3}(x_i)^2 = (x_0')^2 - \sum_{i=1}^{3}(x_i')^2$, gives some conditions on the L matrix, namely,

$$det\ L = +1 \quad or \quad det\ L = -1,$$

and

$$L_0^0 \geq +1 \quad or \quad L_0^0 \leq -1.$$

Thus there are four sets of Lorentz transformations L, depending on whether $det\ L = \pm 1$ and $L_0^0 \geq +1$ *or* $L_0^0 \leq -1$.

But it can be shown that out of these four sets of Lorentz transformations, only those sets can preserve causality for which $L_0^0 \geq +1$. Such transformations are called 'orthochronous Lorentz transformations'. Orthochronous transformations with $det\ L = +1$ are called proper orthochronous transformations. The special Lorentz transformation belongs to this set. Orthochronous Lorentz transformations form a group called the orthochronous Lorentz group.

Here, in the midst of many definitions, we have seen that only the invariance of the square of the Minkowskian distance in 4-dimensional space together with the linearity of the transformation are not sufficient to lead to the orthochronous Lorentz transformations. To get orthochronous Lorentz transformations, one will also have to include causality as an additional condition.

One of Zeeman's research ambitions was to model the special theory of relativity, solely on the basis of the notion of causality.

Does Causality Imply Lorentz Transformation?

In this context, the question naturally arises: If one starts with causality as the physical principle of the world, can it lead to the orthochronous Lorentz transformations? Surprisingly, Chris Zeeman (very well known for his catastrophe theory) answered this question in the positive. In fact one of his research ambitions was to model the special theory of relativity, solely on the basis of the notion of causality.

In 1964, Zeeman showed that the transformations which preserve causality are combinations of the following three basic transformations:

- Translations: $(x^\mu)' = x^\mu + \lambda^\mu$, $\mu = 0, 1, 2, 3$, for some constants λ^μ.
- Positive scalar multiplication: $(x^\mu)' = kx^\mu$, $\mu = 0, 1, 2, 3$ for some positive constant k.
- Orthochronous Lorentz transformations.

More precisely speaking:

Let f be a transformation which maps a time-like (or light-like) vector oriented towards the future to another time-like (or light-like) vector also oriented towards the future, i.e., for time-like and light-like separated events,

$$t_1 > t_2 \Longleftrightarrow f(t_1) > f(t_2).$$

Then such transformations form a group called the causality group.

Let G be another group generated by

- The orthochronous Lorentz group.
- Translations in M (Minkowsi space).
- Dilations (*i.e.*, multiplication by a scalar).

Then the theorem says: *The causality group* $= G$.

We are not giving its proof here as it is beyond the scope of the present article. But the interesting part of Zeeman's proof is that it does not presume linearity of f. Linearity results as a consequence of the assumption of causality. Since two frames of reference related by dilation (multiplication by a factor) in equation (3) differ only by a trivial and unnecessary change of scale, they can be ignored. Moreover, since the constants λ^μ appearing in the translation can be regarded as the coordinates of the origin of the S frame with respect to the S' frame, λ^μ for $\mu = 0, 1, 2, 3$ can be made zero if observers in different frames select a common event to act as the origin. So finally what remains is the orthochronous Lorentz transformation as the admissible nontrivial transformation relating different inertial frames.

Suggested Reading

[1] V Natarajan and D Sen, The special theory of relativity, *Resonance*, Vol. 10, pp. 32-50, 2005.
[2] E C Zeeman, *Journal of Math. Phys.*, Vol. 5, pp. 490-493, 1963.
[3] C Simon, V Buzek and N Gisin, *Physical Review Letters*, Vol. 87, p.170405, 2001.
[4] G Brassard, *Foundations of Physics*, Vol. 33, pp.1593-1615, 2003.
[5] N D Mermin, *Reviews of Modern Phys.*, Vol. 65, p.803, 1993.
[6] D Rohrlich and S Popescu, quant-ph/9508009.
[7] S Chaturvedi, R Simon and N Mukunda, *Resonance*, Vol.11, No. 7, pp.14-29, 2006.

Address for Correspondence
Guruprasad Kar
Samir Kunkri and
Sujit K Choudhary
Physics and Applied Mathematics Unit
Indian Statistical Institute
203 B. T. Road
Kolkata 700108, India
Email
gkar@isical.ac.in
skunkri_r@isical.ac.in,
sujit_r@isical.ac.in