Postulates of the special theory of relativity

At the turn of the 20th century it became clear that the classical mechanics of Newton had some issues. It could not be recoinciled with the electromagnetic theory of Maxwell. Then a break down of the Newtonian mechanics became even more apparent when Michelson and Morley performed their experiment that showed that the speed of light did not depend on any relative uniform motion of the source and observer. The resolution of the problem lied in complete reevaluation of the nature of time. Newton's concept of absolute time had to be replaced by a concept of time that can be different for different observers. Two events occuring at different places, which may appear to be simultaneous to one observer are not necessarily simultaneous to a second observer that moves with respect to the first observer. Einstein formulated two basic postulates based on which the theory known as the Special Theory of Relativity (or just Special Relativity) could be constructed. These are:

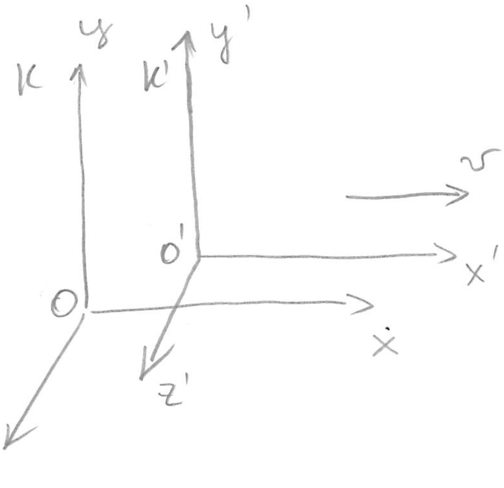
1) No observer can detect his absolute velocity through space. In other words, the laws of physical phenomena are the same in all inertial reference frames. That means only the relative motion of inertial frames can be measured and the concept of motion with respect to the "absolute origin" is meaningless.

2) The speed of light is the same for observers in different inertial frames. That is it is a universal constant independent on any relative motion of the source and observer.

The definition of the inertial reference frame is such that it is a frame that moves uniformly (without acceleration) with respect to any other inertial reference frame.

Galilean transformation

Let us consider two inertial reference frames, K and K', which move with a uniform relative velocity v along their x-axes (see the sketch below).
The transformation of coordinates from one frame to another is given by



$$\begin{cases} x' = x - vt \\ y' = y \\ z' = z \end{cases}$$

Also

$$t' = t$$

These equations define a Galilean transformation.
The element of length in the two systems is the same:

$$ds^2 = dx^2 + dy^2 + dz^2 = \sum_i dx_i^2$$
$$= dx'^2 + dy'^2 + dz'^2 = \sum_i dx_i'^2 = ds'^2$$

(we will use interchangeably the notations x, y, z and x_1, x_2, x_3 for spatial coordinates)

Newton's equations of motion are invariant with respect to Galilean transformations, i.e. they preserve their form:

$$F_i = m\ddot{x}_i \Longrightarrow F_i' = m\ddot{x}_i'$$

The Galilean transformations, however, are inconsistent with the second Einstein's postulate that states that the speed of light is constant in any reference frame. Indeed, according to the Galilean transformation

$$\dot{x}' = \dot{x} - v \qquad \text{or} \qquad \dot{x} = \dot{x}' + v$$

If in system k' the speed is $\dot{x}' = c$ (speed of light) then in system K it would be $\dot{x} = c + v > c$

Lorentz transformation

The transformation that maintains relativistically covariant (i.e. transforming according to the same scheme) form of physical laws is the Lorentz transformation.

Suppose a light pulse is emitted from the common origin of k and k' when they are coincident. Then according to the second postulate the wavefronts observed in the two systems must be described by

$$\begin{aligned} K:&\quad x^2 + y^2 + z^2 - c^2t^2 = 0 \\ k':&\quad x'^2 + y'^2 + z'^2 - c^2t'^2 = 0 \end{aligned} \qquad (*)$$

We can see that these equations are inconsistent with with Galilean transformation, which allows a spherical wavefront in one system, but requires the center of the spherical wave front in the second system to move at vecocity v with respect to the first system. In equations $(*)$ each observer sees that his spherical wavefront has it center fixed at his own coordinate origin as the wavefront expands.

In equations $(*)$ we did not assume $t = t'$. Each observer has his/her own clock. A clock can be placed at any point in space. All clocks can be synchronized. Because the motion takes place along the x-axes it is obvious that

$$y' = y$$
$$z' = z$$

Now at time $t = t' = 0$, when the flashbulb goes off, in system K the motion of the origin O' of K' is described by

$$x - vt = 0$$

while in system K' the motion of O' is described by

$$x' = 0$$

At time $t = t' = 0$ we have $x' = x - vt$. We know, however that the Galilean transformation is not valid in the relativistic case. The simplest linear transformation we can come up with next is of the form

$$x' = \gamma(x - vt)$$

where γ is some coefficient that may depend on v and c, but not on x, t. The linearity is essential because it ensures that each event in K corresponds to one and only one event in K'. Moreover, when v is small compared to the speed of light we should be able to reproduce the Galilean transformation, i.e.

$$\gamma \to 1 \quad \text{when} \quad v \ll c$$

In a similar way we can write for the motion of origin O of K in system K':

$$x = \gamma'(x' + v t')$$

The first postulate requires that the laws of physics are the same in any inertial frame. Therefore, $\gamma = \gamma'$ (Alternatively we can say that $\gamma'(v) = \gamma(-v)$, but $\gamma \to 1$ when $\frac{v}{c} << 0$, so γ is an even function of v). Then we can solve the following system of equations

$$\begin{cases} x' = \gamma(x - vt) \\ x = \gamma(x' + vt') \end{cases} \qquad (**)$$

substituting x' into the last expression gives

$$x = \gamma\left(\gamma(x - vt) + vt'\right)$$

$$\frac{x}{v\gamma} = \frac{\gamma}{v}(x - vt) + t'$$

$$t' = \gamma t + \frac{x}{\gamma v}(1 - \gamma^2)$$

The second postulate requires that the speed of light is the same in both systems. Therefore in both systems

$$x = ct$$
$$x' = ct'$$

When we substitute these into $(**)$ we will get

$$\begin{cases} x' = \gamma\left(x - \frac{v}{c}x\right) \\ x = \gamma\left(x' + \frac{v}{c}x'\right) \end{cases}$$

or

$$\begin{cases} \frac{x'}{x} = \gamma\left(1 - \frac{v}{c}\right) \\ \frac{x'}{x} = \frac{1}{\gamma\left(1 + \frac{v}{c}\right)} \end{cases}$$

Equating the latter two gives

$$\gamma\left(1-\frac{v}{c}\right)=\frac{1}{\gamma\left(1+\frac{v}{c}\right)}$$

$$\gamma^2\left(1-\frac{v^2}{c^2}\right)=1$$

and, lastly

$$\gamma=\frac{1}{\sqrt{1-\frac{v^2}{c^2}}}$$

Then the complete Lorentz transformation can be written as

$$\begin{cases} x'=\dfrac{x-vt}{\sqrt{1-\frac{v^2}{c^2}}} \\ y'=y \\ z'=z \\ t'=\dfrac{t-\frac{vx}{c^2}}{\sqrt{1-\frac{v^2}{c^2}}} \end{cases} \qquad (***)$$

We can also write the inverse Lorentz transformation (most easily obtained by replacing v with $-v$)

$$\begin{cases} x=\dfrac{x'+vt'}{\sqrt{1-\frac{v^2}{c^2}}} \\ y=y' \\ z=z' \\ t=\dfrac{t'+\frac{vx'}{c^2}}{\sqrt{1-\frac{v^2}{c^2}}} \end{cases}$$

In the literature, ratio $\frac{v}{c}$ is often denoted β. It is easy to see that in the limit $\frac{v}{c}\to 0$ the Lorentz transformations reduce to the Galilean ones.

Now let us find how the velocities measured in each system transform. We will denote them as $\vec{u}$ and $\vec{u}'$:

$$u_i = \frac{dx_i}{dt} \qquad u_i' = \frac{dx_i'}{dt'}$$

From (***) we determine

$$u_x' = \frac{dx'}{dt'} = \frac{dx - v\,dt}{dt - \frac{v}{c^2}dx} = \frac{u_x - v}{1 - \frac{v}{c^2}u_x}$$

In a similar way we can also determine

$$u_y' = \frac{u_y}{\gamma\left(1 - \frac{v}{c^2}u_x\right)} \qquad u_z' = \frac{u_z}{\gamma\left(1 - \frac{v}{c^2}u_x\right)}$$

It is easy to verify that when an observer in K measures the speed of light of the light pulse in the x-direction to be $u_x = c$ then an observer in K' will measure

$$u_x' = \frac{c - v}{1 - \frac{v}{c^2}c} = c$$

which is consistent with the second postulate.