From Galilean to Lorentz Transformations: A Pedagogical Viewpoint

Hiqmet Kamberaj[1,2*]

[1]*International Balkan University, Faculty of Engineering,*

Department of Computer Engineering,

Makedonsko-Kosovska Brigada BB, Skopje, North Macedonia and

[2]*National Institute of Physics, Fan Noli Square, Tirana, Albania*

(Dated: June 27, 2025)

* Corresponding author: H. Kamberaj; Email: h.kamberaj@gmail.com

Abstract

Galilean transformations and Lorentz transformations are part of the curricula in modern physics at the upper-secondary schools (such as Grade 11, 12, or Grade 13) and in physics undergraduate programs at any university. However, learning outcomes are different for these courses taught at the upper-secondary schools and physics undergraduate programs at the universities because the prerequisites are different. Furthermore, Einstein's special theory of relativity is related to Lorentz transformations and is part of the physics curricula taught at upper-secondary schools and undergraduate physics degrees at universities. Moreover, that theory is based on the framework of the four-dimensional space-time concept and its fibre-absolute time. These concepts are necessary for a high level of mathematics; therefore, the difficulty in acquiring the desired outcome knowledge can be different for the students of upper-secondary schools and physics undergraduate programs at the universities. The need to balance the mathematical skills necessary to understand materials has increased efforts in designing syllabuses with a gradual development and implementation of modern physics courses in curricula.

In this paper, I discuss the pedagogical frameworks that can be established in the syllabuses of physics courses for teaching the topics, such as Galilean and Lorentz transformations and special theory of relativity, at different levels of secondary school and undergraduate university curricula aiming for a gradual transition from the secondary school to the careers in science, technology, engineering, and mathematics.

Keywords: Galilean transformations, Lorentz transformations, Special theory of relativity, Space-time, Lorentz invariant, Metric tensor

I. INTRODUCTION

The development of curricula that gradually implement new skills in learning physics is one of the main focuses of teaching the physical sciences [1]. These curricula aim to develop students' scientific and critical thinking; additionally, they should strengthen students' interests and affinities for scientific reasoning and research [2–5].

Furthermore, the syllabus should include a variety of supportive learning techniques in the classroom (such as scientific reading materials, data structures, tables and graphs, videos, interactive discussion, artificial intelligence, virtual and augmented realities, and software) combined with problem- and project-based learning strategies [3–5].

More importantly, the students should be prepared to identify physical quantities, measure and identify trends in the user data, utilising data processing tools and graphics [6]. That helps develop students' problem-solving skills, collaborative and group learning skills, and critical thinking, enabling them to become independent. The introduction of modern physics courses in upper-secondary schools and undergraduate university programs has attracted many students to the physical sciences courses; however, the need to balance the mathematical skills necessary to gain an understanding of materials has increased efforts in designing syllabi with a gradual development and implementation of such courses in curricula.

In our previous work [7, 8], we aimed to develop reading materials on the fundamental concepts of physics, focusing on how different theories were developed from physical observations and phenomena, with a greater emphasis on examples and problem-solving techniques. The students gain firsthand experience of how physics theories are applied to engineering and scientific problems, primarily aimed at upper-secondary schools and undergraduate students in engineering and science.

This study aims to develop in-class materials to balance the mathematical skills necessary to understand the reading material. In particular, Galilean and Lorentz transformations, including the special theory of relativity, are an integral part of the science and engineering curriculum; therefore, the focus is on Galilean and Lorentz transformations and the special theory of relativity. The objective is to help secondary school students achieve good grades and prepare them for the challenges they will face during their university studies.

This paper is organised as follows. In Section 2, we discuss Galilean transformations for the coordinates and velocities; besides, we give the formulas of these transformations by introducing the concept of Cartesian coordinates, Euclidean geometry, and inertial reference frame, $S(t, x, y, z)$, with components, the absolute time t and relative Cartesian coordinates (x, y, z). The framework of Sections 2 and 3 can be established in the curricula of the upper-secondary schools for the physics courses. Moreover, the concept of the velocity vector is introduced. Here, we have also introduced the concept of a three-dimensional Euclidean geometry interval as an invariant measure under Galilean transformations and explained that the Newtonian laws remain invariant under these transformations. Next, we continue with the speed of light under Galilean transformations and then introduce the Michelson-Morley experiment as proof that Galilean velocity transformations must be incorrect. That led to the postulation of Einstein's special theory of relativity and the concept of space-time as a four-dimensional geometrical object, introduced in Section 3. In Section 4, Lorentz transformations are introduced for coordinates and velocities. Lorentz transformations (Section 4) can be introduced as an advanced-level physics course for the upper-secondary school students to gradually transition from secondary school to careers in science, technology, engineering, and mathematics.

Moreover, in Section 5, we describe the special theory of relativity and Lorentz transformations in the framework of the four-dimensional manifold space-time and its fibre-absolute time.

Additionally, we introduce the concept of the metric tensor of Minkowski space-time and the submanifold spaces of four-dimensional and covariant vectors. The focus is on establishing a framework for introducing the concepts of the special theory of relativity and Lorentz transformations suitable for the physics undergraduate curricula of universities, which are essential for describing relativistic mechanics and explaining Maxwell's equations as the law of mechanics [9, 10]. Section 6 introduces the Lorentz transformation tensor, and Section 7 discusses the conservation laws. Finally, Section 8 gives the concluding remarks of this study.

II. GALILEAN TRANSFORMATIONS

In Newtonian mechanics, the laws of physics apply the same to all observers in an inertial reference frame. Additionally, an inertial reference frame is defined in accordance with Newton's first law of motion. Moreover, any frame moving with constant velocity relative to an inertial reference frame is an inertial reference frame.

In this section, Galilean transformations are introduced, and then the speed of light under Galilean transformations and the Michelson-Morley experiment are discussed, as per the published literature [7, 11, 12].

Consider two inertial reference frames $S(t, x, y, z)$ and $S'(t', x', y', z')$, as presented in Figure 1. $S'(t', x', y', z')$ frame moves with constant velocity $\mathbf{V}$ along the x axis such that initially (at $t = 0$) the origins, O and O', and the axes of the two frames coincide. A *space-time* vector is defined by the coordinates $(t, \mathbf{r})$, where t is the time and $\mathbf{r}$ is the Cartesian coordinates vector, $\mathbf{r} = (x, y, z)$. Thus, the position of a particle at the point P can be described by the coordinates $(t, \mathbf{r})$, measured by an observer O at S, and coordinated $(t', \mathbf{r}')$, measured by another observer O' at S'.

A. Galilean Coordinates Transformation

The Galilean transformation of the space-time coordinates between S and S' inertial reference frames is defined as the following:

$$t' = t \tag{1}$$

$$x' = x - Vt$$

$$y' = y$$

$$z' = z$$

Alternatively,

$$t = t' \tag{2}$$

$$x = x' + Vt$$

$$y = y'$$

$$z = z'$$

In Eq. (1) or Eq. (2), (t, x, y, z) represent the space-time coordinates of a particle at time t in the inertial frame S measured by the observer O and (t', x', y', z') represent the space-time coordinates of the same particle in the moving frame S' measured by the observer O' (see also Figure 1).

For an arbitrary direction of $\mathbf{V}$, Eq. (2), representing the transformation from S' to S inertial reference frame (see Figure 2), can be written as

$$\mathbf{r} = \mathbf{r}' + \mathbf{V}t \tag{3}$$

where

$$\mathbf{r} = x\mathbf{i} + y\mathbf{j} + z\mathbf{k} \tag{4}$$

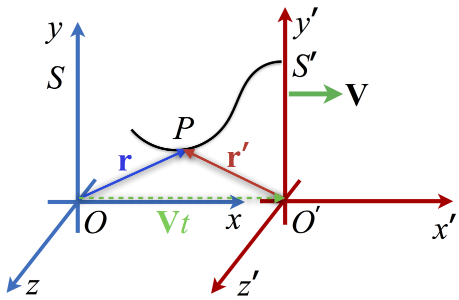

FIG. 1: $O(t, x, y, z)$ and $O'(t', x', y', z')$ represent the origins of two inertial reference frames, S and S'. S' frame moves with constant velocity $\mathbf{V}$ along the x axis such that initially (at $t = 0$) the origins and the axes of the two frames coincide. P represents the position of some particle at an instant.

$$\mathbf{r}' = x'\mathbf{i} + y'\mathbf{j} + z'\mathbf{k}$$

$$\mathbf{V} = V_x\mathbf{i} + V_y\mathbf{j} + V_z\mathbf{k}$$

Here, $\mathbf{i}$, $\mathbf{j}$ and $\mathbf{k}$ are orthogonal unit vectors along x, y, and z axes, respectively, such that:

$$\mathbf{i} \cdot \mathbf{i} = \mathbf{j} \cdot \mathbf{j} = \mathbf{k} \cdot \mathbf{k} = 1 \tag{5}$$

$$\mathbf{i} \cdot \mathbf{j} = \mathbf{i} \cdot \mathbf{k} = \mathbf{j} \cdot \mathbf{k} = 0$$

(Note that here it is assumed that always the axes of the inertial reference frames S and S' are aligned, and thus the same unit vectors $\mathbf{i}$, $\mathbf{j}$, and $\mathbf{k}$ are taken for both S and S'; otherwise, a relative rotation motion could have been introduced between the reference frames, which would have gained an acceleration, and thus S and S' would have been no longer inertial reference frames.)

The inverse transformations (that is, the transformation from S to S' inertial reference frame) are obtained by swapping the terms involving $(')$ with these without $(')$ and changing the sign of $\mathbf{V}$ to $-\mathbf{V}$:

$$\mathbf{r}' = \mathbf{r} - \mathbf{V}t \tag{6}$$

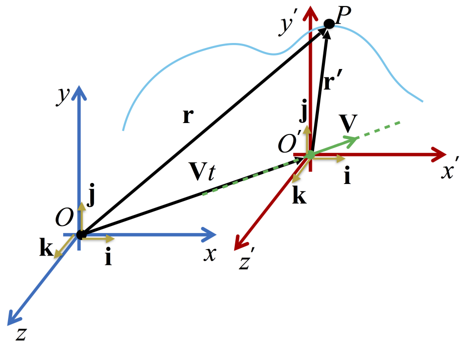

FIG. 2: $O(t, x, y, z)$ and $O'(t', x', y', z')$ represent the origins of two inertial reference frames, S and S'. S' frame moves with constant velocity $\mathbf{V}$ along any arbitrary direction such that initially (at $t = 0$) the origins and the axes of the two frames coincide. P represents the position of some particle at an instant.

Note that Eq. (3) and Eq. (6) can simply be obtained from the geometry (see also Figure 1 and Figure 2) using the rule of addition or subtraction of vectors in three-dimensions [7].

B. Galilean Velocity Transformation

Next, the transformations of particle velocity components between the two inertial frames are derived. The derivative for the time t' in Eq. (1) gives:

$$\frac{dx'}{dt'} = \frac{dx}{dt'} - \frac{d(Vt)}{dt'} \qquad (7)$$
$$\frac{dy'}{dt'} = \frac{dy}{dt'}$$
$$\frac{dz'}{dt'} = \frac{dz}{dt'}$$

Since $dt' = dt$, we obtain

$$v'_x = v_x - V \qquad (8)$$
$$v'_y = v_y$$
$$v'_z = v_z$$

where the velocities v_i (in S) and v'_i (in S'), for $i = x, y, z$, are defined as

$$v_x = \frac{dx}{dt}, \quad v_y = \frac{dy}{dt}, \quad v_z = \frac{dz}{dt} \tag{9}$$
$$v'_x = \frac{dx'}{dt'}, \quad v'_y = \frac{dy'}{dt'}, \quad v'_z = \frac{dz'}{dt'}$$

and $\mathbf{V} = (V, 0, 0)$ is assumed (see also Figure 1). Eq. (9) in vectorial form can be written as

$$\mathbf{v} = \frac{d\mathbf{r}}{dt} \tag{10}$$
$$\mathbf{v}' = \frac{d\mathbf{r}'}{dt'}$$

In general, for any direction of the constant velocity $\mathbf{V}$, it can written that

$$\mathbf{v}' = \mathbf{v} - \mathbf{V} \tag{11}$$

where

$$\mathbf{v} = v_x\mathbf{i} + v_y\mathbf{j} + v_z\mathbf{k} \tag{12}$$
$$\mathbf{v}' = v'_x\mathbf{i} + v'_y\mathbf{j} + v'_z\mathbf{k}$$

Similarly, the inverse transformation (that is, the transformation from S' to S inertial reference frame) is

$$v_x = v'_x + V \tag{13}$$
$$v_y = v'_y$$
$$v_z = v'_z$$

For an arbitrary direction of $\mathbf{V}$, we have

$$\mathbf{v} = \mathbf{v}' + \mathbf{V} \tag{14}$$

Eq. (11) and Eq. (14) are often used for different observations, and they are consistent with Newtonian mechanics [7, 11].

According to Eq. (1) or Eq. (2), the time measured by an observer in any two inertial reference frames, S and S', is the same (i.e., $t = t'$). That is, the time is absolute; hence, the time is one of the invariant quantities in the space-time. In addition, from Eq. (1), we obtain:

$$\begin{aligned} dx' &= dx \\ dy' &= dy \\ dz' &= dz \end{aligned} \tag{15}$$

Therefore, we get

$$\begin{aligned} (ds)^2 &\equiv (dx)^2 + (dy)^2 + (dz)^2 \\ &= (ds')^2 \equiv (dx')^2 + (dy')^2 + (dz')^2 \end{aligned} \tag{16}$$

where $(ds)^2$ is the square of an infinitesimal distance in Euclidean space. Eq. (16) indicates that $(ds)^2$ is an invariant measure; that is, it does not depend on the observer's inertial reference frame.

On the other hand, any velocity vector obeys the transformations given by either Eq. (8) or Eq. (13), and hence it is a relative measure (that is, it depends on the observer's inertial reference frame). That is the view of Newtonian mechanics.

Under Galilean transformations, the second law of Newton, relating the applied force to the rate of change of the momentum, can be written for an observer O in S frame as [7]:

$$\mathbf{F} = \frac{d\mathbf{p}}{dt} \tag{17}$$

and for an observer O' in S' frame as

$$\mathbf{F}' = \frac{d\mathbf{p}'}{dt'} = \frac{d\mathbf{p}}{dt} = \mathbf{F} \tag{18}$$

because $t = t'$. Therefore, the force $\mathbf{F}$ remains invariant under Galilean transformations.

Maxwell's equations of electromagnetism introduced a new universal constant, namely the speed of light c ($c = 3.00 \times 10^8$ m/s). That is inconsistent with Newtonian mechanics, which shows that velocity is a relative measure; its measured value depends on the observer's reference frame. This inconsistency suggests that either Newtonian or Maxwell equations should be modified. Albert Einstein solved the paradox in 1905 with his *special theory of relativity*(STR). He concluded that Galilean transformations are wrong and time is not absolute [13]. At that time, STR was necessary to solve the contradictions between the Newtonian and Maxwell equations, which were interpreted as mechanical laws.

C. The Speed of Light Under Galilean Velocity Transformations

Until the late 1800s, the value of the speed of light was a special measure relative to a fixed frame connected to *ether*, which was the medium through which the light signals were assumed to propagate. The frame connected to the ether was also called *absolute ether frame*. In these conditions, using the Galilean transformations, the speed of light can be lower or greater than c; that is, it is not an invariant measure, but it depends on the observer's reference frame. For example, consider the situation in Figure 3, as described in Ref. [11]. $S(t, x, y, z)$ and $S'(t', x', y', z')$ represent two inertial reference frames. S' moves with a constant velocity $\mathbf{V}$ along the x axis, and frame S is at rest (the so-called absolute ether frame). Initially (at $t = 0$), the two frames' origins and axes coincide. A light signal is sent from a source of light fixed at S' towards either the positive or negative x axis. v represents the speed of light for the absolute ether frame, S, and c represents the speed of light in the moving frame S'.

Using the Galilean velocity transformations, we will obtain:

$$v = c - V \tag{19}$$

if S' approaches S frame (or equivalently, towards the negative x axis), and

$$v = c + V \tag{20}$$

If the frame S' moves away from the S frame (or equivalently, towards the positive x axis). Note that Eq. (19) and Eq. (20) are obtained simply by projecting Eq. (14) along the x-axis for the two cases of the direction of the vector $\mathbf{V}$ shown in Figure 3.

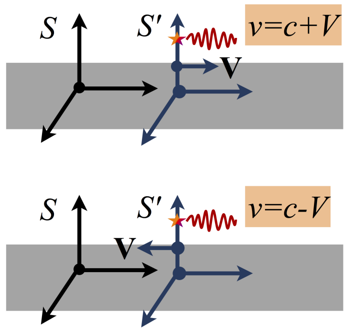

FIG. 3: $S(t, x, y, z)$ and $S'(t', x', y', z')$ represent two inertial reference frames. S' is a moving frame with constant velocity $\mathbf{V}$ along the x axis, and S is a frame at rest, the so-called absolute ether frame. Initially (at $t = 0$), the origins and the axes of the two frames coincide. A light signal is sent from a source of light fixed at S' towards either the positive or negative x axis. v represents the speed of light in the absolute ether frame S using the Galilean transformations, and c represents the speed of light in the moving frame S'.

Furthermore, when $\mathbf{v}$, which represents the velocity of light in the absolute ether frame S, is perpendicular to the relative velocity $\mathbf{V}$ of the absolute ether frame and moving frame, as shown in Figure 4, we will obtain:

$$v = \sqrt{c^2 - V^2} \tag{21}$$

That indicates that the speed of light is not an invariant measure. Its measured value depends on the observer's reference frame. Later, Maxwell, when deriving the electromagnetic equations [8], showed that the speed of light is a universal constant and that the electromagnetic wave does not need a medium to propagate; that is, it can also spread in an absolute vacuum.

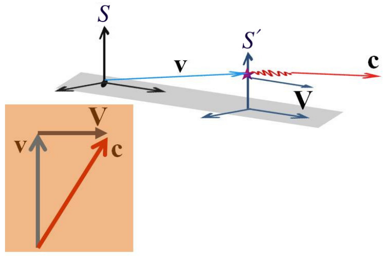

FIG. 4: Illustration of the case when $\mathbf{v}$, representing the velocity of light with respect to the absolute ether frame, is perpendicular to the relative velocity $\mathbf{V}$ of the absolute ether frame and moving frame.

D. The Speed of Light Under Michelson-Morley Experiment

Eq. (19) and Eq. (20) point out that small changes could exist in the measurement of the speed of light from what is predicted by Maxwell's equations.

The first experiment used to verify the existence of the ether medium is the Michelson-Morley experiment. A setup of that experiment is shown in Figure 5, as in Ref. [11]. M_0, M_1 and M_2 are three mirrors, where M_0 forms an angle of $45°$ with horizontal direction. Let S be the absolute frame of ether and S' the frame connected to the Earth. Suppose that the orientation of L_2 is such that Earth's velocity has an opposite direction to the positive x axis. Then, using Eq. (19) and Eq. (20), the speed of light along y-axis with respect to Earth (S' frame) for the incident light wave is:

$$v = c + V \tag{22}$$

and for the reflected light wave is

$$v = c - V \tag{23}$$

Note that in Eq. (22) and Eq. (23), c denotes the speed of light with respect to the absolute ether frame (S) and v the speed of light with respect to Earth (S').

The two reflected waves from M_1 and M_2 mirrors superimpose and an interference pattern is observed, using an interferometer at the position of the telescope in Figure 5, made up of dark and bright fringes [11]. First, the time for the light travelling the path L_2 is calculated as:

$$t_1 = \frac{L_2}{c+V} + \frac{L_2}{c-V} = \frac{2cL_2}{c^2 - V^2} \tag{24}$$

On the other hand, the time taken to travel the path L_1, which is perpendicular to the direction of the velocity $\mathbf{V}$, is given as

$$t_2 = \frac{2L_1}{\sqrt{c^2 - V^2}} \tag{25}$$

Therefore, Eq. (22) and Eq. (23) indicate that there is a time shift between the light waves following the paths L_1 and L_2 when they arrive at the telescope, given as

$$\begin{aligned} \delta t = t_1 - t_2 &= \frac{2cL_2}{c^2 - V^2} - \frac{2L_1}{\sqrt{c^2 - V^2}} \\ &= \frac{2L_2}{c\left(1 - \frac{V^2}{c^2}\right)} - \frac{2L_1}{c\sqrt{1 - \frac{V^2}{c^2}}} \end{aligned} \tag{26}$$

Because often $V/c \ll 1$, then the following approximation can be used from a Taylor expansion:

$$\begin{aligned} \frac{1}{\sqrt{1 - \frac{V^2}{c^2}}} &\approx 1 + \frac{V^2}{2c^2} \\ \frac{1}{1 - \frac{V^2}{c^2}} &\approx 1 + \frac{V^2}{c^2} \end{aligned} \tag{27}$$

Eq. (26) can then be written as

$$\delta t = \frac{2L_2}{c}\left(1 + \frac{V^2}{c^2}\right) - \frac{2L_1}{c}\left(1 + \frac{V^2}{2c^2}\right) \tag{28}$$

For $L_1 = L_2 = L$, we obtain

$$\delta t = \frac{LV^2}{c^3} \tag{29}$$

The difference in the travelling time δt will affect the interference pattern as seen from the telescope of the two reflected lights coming along the paths L_1 and L_2, respectively. If we rotate the interferometer by $90°$, then the direction of the velocity of Earth rotates with $90°$ such that $\mathbf{V}$ is along L_1 path. In such a case, the light speeds of the incident and reflected along L_1 will change according to Eq. (22) and Eq. (23), and hence the paths switch the roles, causing the time shift to double. Thus,

$$\Delta t = 2\delta t = 2\frac{LV^2}{c^3} \tag{30}$$

The time shift corresponds to a path shift given as

$$\Delta d = c\Delta t = 2\frac{LV^2}{c^2} \tag{31}$$

Furthermore, a shift in the path travelled by the two light waves with λ (wavelength) is equivalent to a shift with one fringe on the interference pattern; therefore, Δd will cause the interference pattern to slightly shift by

$$\text{shift} = \frac{\Delta d}{\lambda} = \frac{2LV^2}{\lambda c^2} \tag{32}$$

However, up to date, no one has managed to observe that shift in the interference pattern, and many have tried [11]. That suggests that the Galilean velocity transformation equations are incorrect. Hence, the argument that the speed of light is invariant remains, which was initially derived by Maxwell.

Albert Einstein, in 1905, decided that Maxwell's equations are correct and Newtonian mechanics, and hence Galilean transformations are incorrect [13]. Furthermore, he introduced the special theory of relativity, which Einstein considered necessary at that time to explain electromagnetism. More about the interpretation of Maxwell's equations as mechanical laws can be found in Refs. [9, 10].

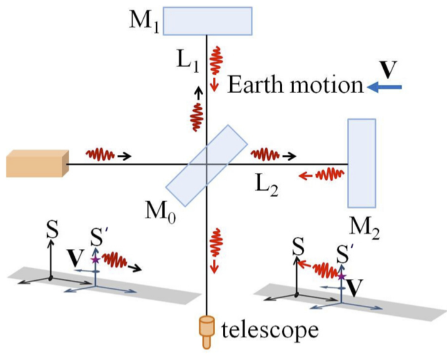

FIG. 5: Michelson-Morley experiment setup.

III. POSTULATES OF SPECIAL THEORY OF RELATIVITY

In this section, Einstein's special theory of relativity is introduced using the framework of Cartesian coordinates and space-time geometry [14]. Einstein's postulates of the special theory of relativity are also introduced and discussed [11–13, 15–18].

A. Postulates of Einstein's Special Theory of Relativity

Two postulates given by Einstein, known as the special theory of relativity, are:

1. **The laws of physics are equivalent in all inertial reference frames.**

   As an experiment, consider the free fall of a ball from the ceiling of a train moving with constant velocity **v** to the right, shown in Figure 6. Consider a reference frame $S(t, x, y, z)$ connected to the observer at rest in a laboratory, and $S'(t', x', y', z')$ connected to an observer sitting in the train, which moves with constant speed v. $\mathcal{L}'$ denotes the path of the free fall measured by the observer at the frame S' at rest relative to the train, and $\mathcal{L}$ is the path of the ball measured by the observer at the frame S. Since the train moves with constant velocity, both S and S' are inertial reference frames, and hence the laws of physics are the same in both frames. However, the measurements of the two observers are different. For the observer at S' at rest relative to the moving train, the motion of the ball represents a free fall. For the observer at frame S, the ball moves to the right with constant speed v along a parabolic path.

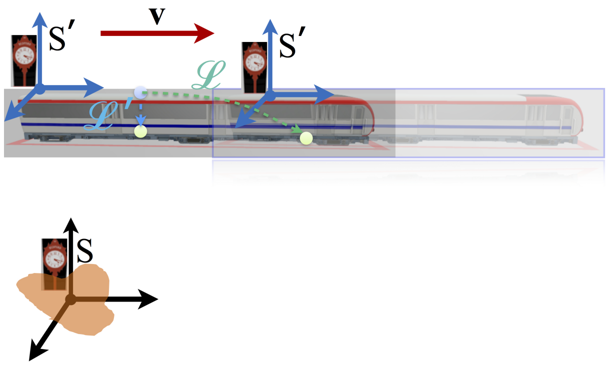

FIG. 6: A ball undertakes a free fall from the ceiling of a train moving with constant velocity **v** to the right. Consider the reference frame $S(t, x, y, z)$ connected to the object at rest in a laboratory, and $S'(t', x', y', z')$ connected to the train moving with constant speed v. $\mathcal{L}'$ denotes the path of the free fall measured by an observer at the frame S' at rest relative to the train, and $\mathcal{L}$ is the path of the ball measured by an observer at the frame S.

2. **The speed of light is the same in all inertial reference frames, *i.e.*, the speed of light in absolute vacuum is a universal constant,** $c = 3.00 \times 10^8$ **m/s.**

   Consider the experiment of a light signal sent from the point A to B, separated by a distance $l' = c\tau$, where c is the speed of light and τ is the time measured by the stationary clock (the so-called proper time). The reference frame $S'(t', x', y', z')$ is connected to an observer at rest relative to the observer moving with a significantly large constant speed v, $S(t, x, y, z)$. The axes of the two reference frames coincide. $d = vt$ denotes the distance travelled by the observer in S, t is the time measured by the running clock (the so-called laboratory time), and $l = ct$ is the distance between A and B measured by the running clock. Using

Pythagoras theorem: $(ct)^2 - (vt)^2 = (c\tau)^2$, or $t = \tau/\sqrt{1 - v^2/c^2}$ and $ct > c\tau$ ($l > l'$).

This experiment demonstrates that although the speed of light in a vacuum is constant (i.e., c), the time measured by two clocks (one at rest and the other moving with constant velocity) for the signal to travel between the same points is different because the measured distance between these two points depends on the reference frame. That is a consequence of the so-called time dilation and length contraction, which will be discussed in the following; however,

$$\frac{l'}{\tau} = \frac{l}{t} = c \tag{33}$$

Eq. (33) indicates that the time dilation factor should equal the length contraction factor.

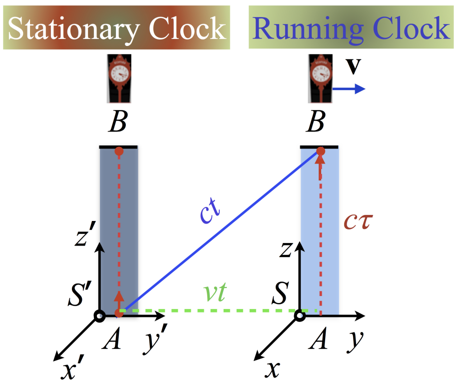

FIG. 7: A light signal is sent from A to B, separated by a distance $c\tau$, where c is the speed of light and τ is the time measured by the stationary clock (the so-called proper time). The reference frame $S'(t', x', y', z')$ is connected to an observer at rest relative to the observer moving with a significantly large constant speed v, $S(t, x, y, z)$. The axes of the two reference frames coincide. vt denotes the distance travelled by the observer in S, and t is the time measured by the running clock (the so-called laboratory time). Using Pythagoras theorem: $(ct)^2 - (vt)^2 = (c\tau)^2$, or $t = \tau/\sqrt{1 - v^2/c^2}$ and $ct > c\tau$.

B. Space-time

In the special theory of relativity, both time and space are unified in a single entity known as the *space-time*. That replaces Euclidean geometry of Newtonian mechanics, and hence, the space-time is the framework for describing the laws of physics in STR. Every observer can measure the space and time coordinates of the events; however, all inertial observers measure a unique separation of the events. That is:

$$(ds)^2 = c^2\,(dt)^2 - (d\mathbf{r})^2 \tag{34}$$

where $c^2\,(dt)^2$ is the square of the time interval and $(d\mathbf{r})^2 = (dx)^2 + (dy)^2 + (dz)^2$ is the square of the space interval between any two points in the space-time with an infinitesimal separation. A point in the space-time is characterised by four coordinates, $(ct,\, x,\, y,\, z)$, namely the time-like coordinate ct and the space coordinates $\mathbf{r} = (x,\, y,\, z)$. That point is also called *event*. The space-time, defined by Eq. (34), is also called Minkowski space.

Eq. (34) can also be written as:

$$(ds)^2 = c^2\,(dt)^2 \left(1 - \frac{v^2}{c^2}\right) \tag{35}$$

Eq. (35) indicates that $(ds)^2 = 0$ for $v = c$. That interval is called *light-like* interval. Furthermore, $(ds)^2 > 0$ for $v < c$, which is called *time-like* interval. Moreover, $(ds)^2 < 0$ for speeds of events greater than the speed of light $v > c$, which is called *space-like* interval.

Therefore, for any inertial reference frame, the objects travelling on the light-like paths (see also Figure 8) move with the speed of light and are called light-like objects, which have a rest mass equal to zero. The objects moving along the time-like paths have a speed less than the speed of light and are called *tardyons* (which are particles with non-zero mass); while those objects moving along the space-like paths have a speed higher than the speed of light and are called *tachyons* (which are hypothetical particles that travel backward in time, not discovered yet).

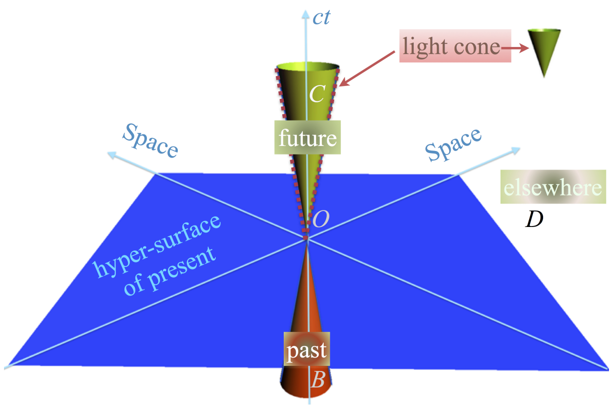

FIG. 8: The cone of light.

C. Invariant Space-time Interval

Note that all the inertial observers measure the same interval ds between any two infinitesimally separated events, and hence the space-time interval ds is an invariant measure (that is, its value does not depend on the inertial reference frame). Thus, if $S(t, x, y, z)$ and $S'(t', x', y', z')$ are two inertial reference frames, then we will have

$$(ds)^2 = (ds')^2 \tag{36}$$

where

$$(ds)^2 = c^2 (dt)^2 - \left((dx)^2 + (dy)^2 + (dz)^2\right) \tag{37}$$

$$(ds')^2 = c^2 (dt')^2 - \left((dx')^2 + (dy')^2 + (dz')^2\right)$$

For Eq. (36) to hold, when one goes from one inertial frame S to another inertial frame S', transformations of the coordinates between the two frames must involve the relative velocity between the two frames in both space and time parts. Therefore, t is no longer an invariant measure in two inertial frames (i.e., $t \neq t'$). This indicates that the relative partition of space-time into space and time components will be different for different inertial observers, depending on their relative velocity. In other words, the time measured in the laboratory frame (inertial frame at rest) is different from the time measured in a moving inertial reference frame with constant speed v relative to the inertial frame at rest, and hence $cdt \neq cdt'$. Furthermore, if we denote the length of any geometrical elementary object in three-dimensional space as

$$(dl)^2 = (dx)^2 + (dy)^2 + (dz)^2 \quad \text{(in } S\text{)} \tag{38}$$
$$(dl')^2 = (dx')^2 + (dy')^2 + (dz')^2 \quad \text{(in } S'\text{)}$$

which gives the distances between any two infinitesimally separated points in three-dimensional space, determined in Euclidean geometry, and measured in S and S'. Then, $dl \neq dl'$.

D. Time Dilation

Two types of clocks can be introduced: those at rest and those moving with constant speed v. The time measured by the clocks at rest is called *proper time*, and the time measured by all other inertial observer clocks is called *laboratory time*.

Consider the experiment with a light signal travelling from A to B, as shown in Figure 9. We also consider the reference frame $S(t, x, y, z)$ at rest (laboratory inertial frame) and $S'(t', x', y', z')$ moving with constant velocity $\mathbf{v}$ with the object where the events are observed along the y axis relative to S. The axes of the two reference frames coincide. As the object at rest relative to the inertial frame, S', moves to the right, the light signal follows the path shown in blue, as depicted in Figure 9. Here, dt equals the time interval measured by the observer's clock at reference frame S, as the light signal travels the distance AB. For the observer at S, the measured displacement in space between these two events is $dx = 0$, $dy = vdt$, and $dz = l$, where l is the distance between A and B measured along the direction of the z axis (see also Figure 9). The observer's clock at the reference frame S' measures a time interval dt' between these two events, and the displacements in space measured by this observer are $dx' = dy' = 0$, and $dz' = l$. From Eq. (37), we can write:

$$(ds')^2 = c^2 (dt')^2 - l^2 \tag{39}$$

$$(ds)^2 = c^2 (dt)^2 - l^2 - (vdt)^2$$

Because of the invariance of the space-time intervals (see Eq. (36)):

$$c^2 (dt')^2 - l^2 = c^2 (dt)^2 - l^2 - (vdt)^2 \tag{40}$$

or

$$dt = \frac{dt'}{\sqrt{1 - \dfrac{v^2}{c^2}}} \tag{41}$$

Eq. (41) gives the relationship between the time interval dt, measured by the clock of the observer at S, and the interval of time dt', measured by the clock of the observer at the moving frame S' with constant speed v relative to S. Here, dt is called the laboratory time; furthermore, the clock of the observer at S' is at rest relative to the moving object where the two events occur, and thus dt' is the so-called proper time interval $d\tau$.

Eq. (41) points out that $dt > dt'$ because the factor $1/\sqrt{1 - v^2/c^2} > 1$ (since $v < c$). This is called *time dilation*; that is, the observers running faster through space run slower through time (the clock of the observer connected at S', who is running through space, runs slower through time $dt > dt'$).

E. Light Cone

Because the space-time interval ds between any two infinitesimally separated events is an invariant measure in Minkowski space, often the space-time can be imagined to be divided into four regions, as shown in Figure 8 with respect to an arbitrary event O at some time t_O located at the

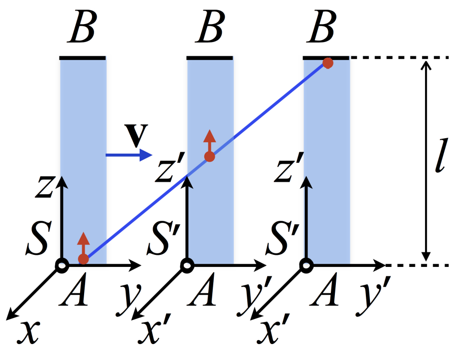

FIG. 9: A light signal is traveling from A to B. We consider the reference frame $S(t, x, y, z)$ at rest (laboratory inertial frame) and $S'(t', x', y', z')$ moving with constant velocity $\mathbf{v}$ with the object where the events are observed along the y axis. The axes of the two reference frames coincide.

origin (*i.e.*, $t = x = y = z = 0$). For any event B in the lower cone occurring in the past at the time t_B, we have

$$(ds_{OB})^2 > 0 \tag{42}$$

Furthermore, if $t_B < t_A$ in some inertial frame, then $t_B < t_A$ in any other chosen inertial frame. That region is called the *past* region. It is also possible to choose an inertial frame such that $(ds_{OB})^2 > 0$, where B has the same space coordinates as A.

The region C (see also Figure 8) is called the *future* region of events. Any event occurring at time $t_C > t_A$ at an inertial frame will have the same order of occurrence in any other chosen inertial frame. Again, it is also possible to choose an inertial frame such that $(ds_{OC})^2 > 0$, where C has the same space coordinates as the origin A.

On the other hand, if $(ds_{OD})^2 < 0$, then we have the *elsewhere* region. Here, there exist inertial frames for which the order of the occurrence of the events at t_A and t_D can be reversed (*i.e.*, either $t_A > t_D$ or $t_D > t_A$) or even made equal (*i.e.*, $t_A = t_D$) but at different space coordinates.

The fourth region is the so-called *light cone*, which separates the past-future and elsewhere regions. In this region, $(ds)^2 = 0$, and hence it characterises the space-time events moving with the speed of light; that is, the light emitted by these points reaches the event A, and vice versa, the light emitted from event A can reach the space-time points of the light cone.

F. Length Contraction

Consider a spaceship moving from the star A to the star B with constant speed v, as shown in Figure 10. Let S and S' be two inertial frame observers, respectively, at a laboratory on Earth and the spaceship. Furthermore, we consider two clocks connected to these two observers measuring the time interval for each observer between any two events in space-time. We consider an event A occurring when the spaceship is at the star A and an event B occurring when the spaceship reaches the star B.

The observer at rest on Earth, who is also at rest relative to the stars, measures the so-called proper distance between the stars, L_p. Furthermore, the clock of the observer at rest on Earth (at S) measures the time interval Δt that the spaceship needs to complete the trip:

$$\Delta t = \frac{L_p}{v} \tag{43}$$

On the other hand, the observer in the spaceship (at S') is at rest relative to the spaceship where the events occur; therefore, its clock measures the proper interval of time, $\Delta\tau$, between these two events. Since this observer (at rest relative to the spaceship) moves relative to Earth with speed v, the distance, L, measured between the stars is

$$L = v\Delta\tau \tag{44}$$

Here, L_p is the distance between the two events measured by the clock of the observer inertial frame, S, at rest on Earth and at rest relative to the stars A and B, and L is the distance between the two events measured by the clock of observer inertial frame S' on a spaceship who is moving relative to the Earth and the stars.

Combining Eq. (43) and Eq. (44), we obtain

$$\frac{L}{v} = \Delta\tau = \Delta t\sqrt{1-\frac{v^2}{c^2}} = \frac{L_p}{v}\sqrt{1-\frac{v^2}{c^2}} \tag{45}$$

or

$$L = L_p\sqrt{1-\frac{v^2}{c^2}} \tag{46}$$

From Eq. (46), we obtain

$$L_p > L \tag{47}$$

because $\sqrt{1-\frac{v^2}{c^2}} < 1$. The relationship given by Eq. (46) is called *length contraction* effect:

> When the observer is at rest relative to the object, then the object has a proper length L_p, and when the observer moves relative to the object with speed v in a direction parallel to its length, then the observer measures a shorter length of the object, L. Furthermore, no length contraction occurs in the directions perpendicular to the velocity of $\mathbf{v}$.

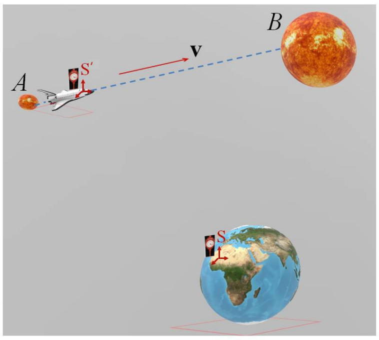

FIG. 10: The spaceship moving with constant velocity $\mathbf{v}$ along the direction that measures the distance from a star A to another star B.

G. Relativistic Momentum

Basic units of measurement in Physics are length, time, and mass. We have seen that length and time are relative measures dependent on the inertial reference frame. In analogy with that, in the special theory of relativity, the mass is also considered relativistic, defined as [11]

$$M = \frac{m}{\sqrt{1 - \frac{v^2}{c^2}}} \tag{48}$$

where m is the rest mass of a particle (or the mass measured in a stationary inertial reference frame in which the particle is at rest), M is the mass of that particle at an inertial reference frame relative to which the particle is moving with a constant speed v. Note that since $v < c$, then $M > m$.

Then, the relativistic three-dimensional momentum of a particle moving with constant velocity $\mathbf{v}$ and rest mass m is defined as

$$\begin{aligned} \mathbf{P} = M\mathbf{v} &= \frac{m\mathbf{v}}{\sqrt{1 - \frac{v^2}{c^2}}} \\ &\equiv \frac{\mathbf{p}}{\sqrt{1 - \frac{v^2}{c^2}}} \end{aligned} \tag{49}$$

Note that $\mathbf{P}$ preserves the conservation law of the total momentum in a system of two particles that collide; however, $\mathbf{p} = m\mathbf{v}$ is not conserved during the collision or for an isolated system of relativistic particles.

The relativistic three-dimensional force can also be defined as

$$\mathbf{F} = \frac{d\mathbf{P}}{dt} \tag{50}$$

H. Mass Energy Equivalence Principle

1. *Relativistic Work*

We will start using the definition of the relativistic force as in Eq. (50) to calculate the work done on a particle by that force:

$$W = \int_{x_1}^{x_2} F\, dx = \int_{x_1}^{x_2} \frac{dP}{dt}\, dx \tag{51}$$

Here, we have assumed that the force is acting along the x-axis, and P is the relativistic momentum. Therefore,

$$\begin{aligned}
\frac{dP}{dt} &= \frac{d}{dt}\left(\frac{mv}{\sqrt{1-\dfrac{v^2}{c^2}}}\right) \\
&= \frac{\dfrac{1}{\gamma} m \dfrac{dv}{dt} - mv\gamma\left(-\dfrac{v}{c^2}\dfrac{dv}{dt}\right)}{1-\dfrac{v^2}{c^2}} \\
&= \gamma m \frac{dv}{dt} + \gamma^3\beta^2 m \frac{dv}{dt} \\
&= m\gamma \frac{dv}{dt}\left(1+\gamma^2\beta^2\right) \\
&= \gamma^3 m \frac{dv}{dt}
\end{aligned} \tag{52}$$

where $\gamma = 1/\sqrt{1-v^2/c^2}$.

Substituting Eq. (52) into Eq. (51), we obtain:

$$W = \int_0^v \gamma^3 mv\, dv \tag{53}$$

assuming that the particle accelerates from speed 0 to some speed $v < c$. To integrate, we write $\gamma^3 = (1/x)^{3/2}$, where $x \equiv 1 - v^2/c^2 = 1/\gamma^2$. Then, the following change on the variable of integration takes place:

$$v\, dv = -\frac{c^2}{2} d\left(1 - \frac{v^2}{c^2}\right) = -\frac{c^2}{2} dx \tag{54}$$

With these changes, Eq. (53) becomes

$$W = -\frac{mc^2}{2} \int_1^{1/\gamma^2} \frac{dx}{x^{3/2}} \tag{55}$$

Integrating Eq. (55), we get

$$W = \frac{mc^2}{\sqrt{1 - \frac{v^2}{c^2}}} - mc^2 \tag{56}$$

Based on the work-kinetic energy theorem, work done by the force equals the change on the kinetic energy ΔK:

$$\Delta K = K_f - K_i \tag{57}$$

2. *Relativistic Kinetic Energy*

Since we assumed that the initial velocity of the particle is zero, then $K_i = 0$, and hence $\Delta K = K_f \equiv K$. Here, K is the relativistic kinetic energy, which becomes

$$K = \frac{mc^2}{\sqrt{1 - \frac{v^2}{c^2}}} - mc^2 = Mc^2 - mc^2 \tag{58}$$

Note that Eq. (58) is proved by the experimental results, in particular, high-energy particle physics. In Figure 11, the scaled relativistic and kinetic energy K/mc^2 is shown versus scaled velocity v/c. Additionally, the non-relativistic kinetic energy ($K = mv^2/2$) is also displayed in the plot. For small values of the particle's speed relative to the speed of light, both relativistic and non-relativistic kinetic energy agree. That can also be shown analytically, taking the limit when $v \ll c$ in Eq. (58). In particular, for $v \ll c$, or equivalently $v/c \ll 1$, we have

$$\gamma = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}} \approx 1 + \frac{v^2}{2c^2} \tag{59}$$

Substituting that expression into Eq. (145), we obtain

$$K = \frac{mv^2}{2} \tag{60}$$

which is the non-relativistic expression of the kinetic energy. That agreement is also shown graphically in the inset plot in Figure 11.

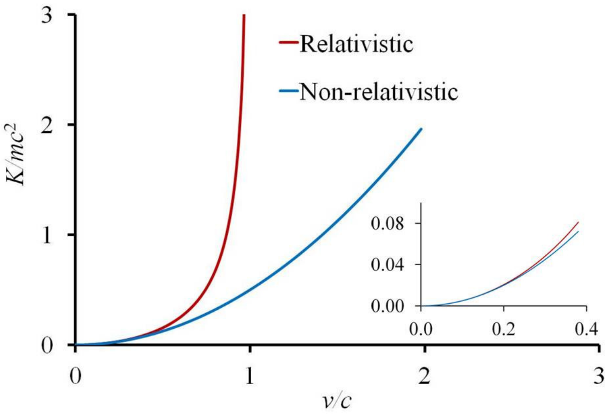

FIG. 11: The relativistic kinetic energy K/mc^2 versus v/c for the relativistic and non-relativistic mechanics.

3. *Rest Energy and Relativistic Total Energy*

In Eq. (58), the term mc^2 (which is a constant term) is the so-called *rest energy*, E_0:

$$E_0 = mc^2 \tag{61}$$

By definition, the sum of the rest energy with kinetic energy gives the total relativistic energy of the particle:

$$E = E_0 + K = \frac{mc^2}{\sqrt{1 - \frac{v^2}{c^2}}} = Mc^2 \tag{62}$$

Eq. (62) is known as Einstein's *mass-energy equivalence* for which he has become famous. Besides, the relationship in Eq. (62) indicates that mass is a form of energy, and c^2 is a constant used to convert the mass into units of energy. Since the factor c^2 is quite a large number:

$$c^2 = 9.0 \times 10^{16}\ \mathrm{m}^2/\mathrm{s}^2 \tag{63}$$

Then, just a small mass can represent a huge amount of energy, as is often the case in high-energy particle physics.

Taking the square of the three-dimensional relativistic momentum and multiplying by c^2, we get

$$\begin{aligned} P^2c^2 &= \frac{v^2}{c^2}\frac{m^2c^4}{1-\dfrac{v^2}{c^2}} \\ &= -\left(1-\frac{v^2}{c^2}\right)\frac{m^2c^4}{1-\dfrac{v^2}{c^2}} + \frac{m^2c^4}{1-\dfrac{v^2}{c^2}} \\ &= -m^2c^4 + E^2 \end{aligned} \tag{64}$$

which can be re-arranged as

$$E^2 = P^2c^2 + m^2c^4 \tag{65}$$

Thus, for a particle at rest, that is $P = 0$, then $E = E_0 = mc^2$. Here, E represents the total energy of a particle. Eq. (65) indicates that even when a particle is at rest relative to some reference frame, it possesses energy, which is the rest energy E_0.

Using Eq. (62) and Eq. (65), often the relativistic kinetic energy is also written as:

$$K = E - mc^2 = \sqrt{P^2c^2 + m^2c^4} - mc^2 \tag{66}$$

Note that the rest mass m does not depend on the reference frame; thus, from Eq. (65) the quantity $E^2 - P^2c^2$ is an invariant measure; that is, it is invariant under the Lorentz transformations.

For photons having the rest mass equal to zero, $m = 0$, then from Eq. (65), we obtain

$$E = Pc \tag{67}$$

To develop the equivalence of mass and energy formula, given by Eq. (62), Einstein thought about the following experiment, which is also described in Ref. [11].

Suppose a box of mass M and length L at rest at $t = 0$, as shown in Figure 12. A light signal is sent from the left side of the box towards the right side of the box along the x axis. Let E be the energy of the light signal, then using Eq. (153), the momentum of the light pulse is

$$P = \frac{E}{c} \tag{68}$$

which is along the positive x axis. According to the conservation law of momentum, the box will gain the same momentum but in the opposite direction. Therefore, the box recoils, as indicated in Figure 12. The speed of the box is

$$v = \frac{P}{M} = \frac{E}{Mc} \tag{69}$$

Since the box is big, then $v \ll c$. The time needed for the light signal to strike the right end of the box is:

$$\Delta t = \frac{L}{c} \tag{70}$$

During that time, the box moves to the left by a small distance Δx:

$$\Delta x = v\Delta t = \frac{EL}{Mc^2} \tag{71}$$

When the light strikes the right end of the box, it transfers momentum P to the box, which is equal in magnitude but opposite in direction to the initial momentum, and hence the box stops moving. The box moves to the left by an amount Δx; thus, the centre of mass of the box shifts to the left by the same amount, Δx. However, the box-light pulse system is isolated, and hence the resultant force is zero; therefore, the centre of mass of the entire system must not have moved. To explain this, we must assume that light also carries a mass, let us say an effective mass M'. Thus, for the centre of mass of the entire system to remain fixed, we should write that

$$M'L = M\Delta x \tag{72}$$

or

$$M'L = M\frac{EL}{Mc^2} \tag{73}$$

From here, the effective mass of the light signal is

$$M' = \frac{E}{c^2} \tag{74}$$

or

$$E = M'c^2 \tag{75}$$

From that experiment, Einstein concluded that if a particular body gives off the energy E in the form of radiation, its mass reduces by E/c^2.

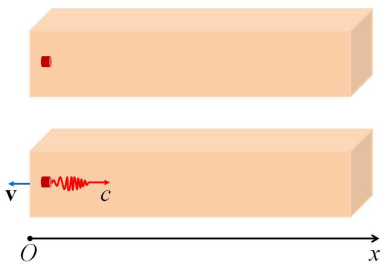

FIG. 12: Illustration of the light signal emitted at the left end of the box towards the right end of the box.

I. Consequences of Special Theory of Relativity

Consequences of the STR include [11, 16]:

1. In every inertial reference frame, the events in space-time are characterised by a four-dimensional vector (ct, x, y, z), as shown in Figure 13. That is equivalent to creating a

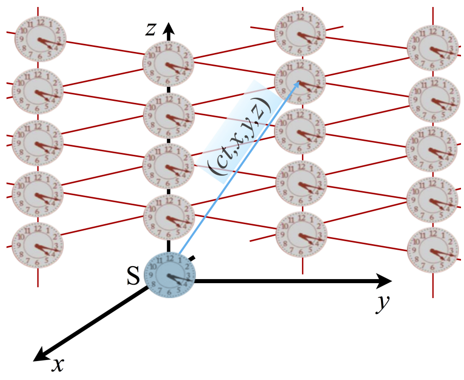

FIG. 13: An inertial reference frame observer in STR.

grid in three-dimensional space, where the intersections represent the spatial coordinates (x, y, z) [11]. Then, at every intersection grid point, synchronised clocks are set up to measure the time t. The synchronisation of the clocks can be done as follows. An observer at the origin at $t = 0$ and $x = y = z = 0$ measures the time using a master clock (shown in blue in Figure 13). This observer sends a light signal at a grid point at a distance $r = \sqrt{x^2 + y^2 + z^2}$, which takes $t = r/c$ time to reach that point, then the clock at this position denotes $t = r/c$ when the master clock is at $t = 0$, assuming that the speed of light is the same in all directions.

2. Not all the inertial reference frames are equivalent. Only the inertial reference frames at rest relative to each other are equivalent. Therefore, not all the inertial reference frames use the same grid of points and clocks.

3. Not all events are separated in space-time uniquely in all inertial reference frames as opposed to Newtonian mechanics. For example, in Newtonian mechanics, time and spatial distance are absolute. On the other hand, in relativistic mechanics, there is no absolute time and length. Indeed, the interval of time between any two events measured by the observer's clock at rest in an inertial reference frame is not the same as the time interval of the same

events measured by the observer's clock in an inertial reference frame moving relative to the frame at rest. The same holds for the length of the distance between any two events in space-time.

4. It is a consequence of the relativistic mechanics that two events that are simultaneous in an inertial reference frame may not be so in all other inertial reference frames. For example, consider the experiment shown in Figure 14. Two light sources are located at A and B, sending light towards each other. Consider two observers, one at the fixed laboratory inertial frame (location O) and the other at rest in a moving train with speed v (location O'). For the observer at O, the light signals sent from A and B will reach O simultaneously. However, the observer at O', as the train moves towards B with speed v, is approaching point B and moves away from the source at point A; hence the observer at O' will see first the light signal sent by the light source at B, then the light signal sent by the light source at A. Therefore, for the observer at O', these two events are not simultaneous.

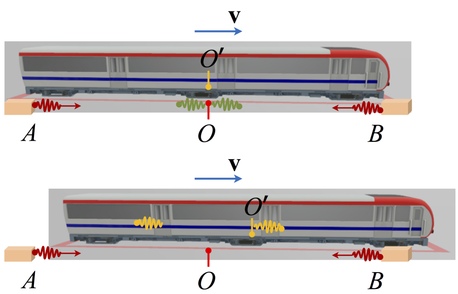

FIG. 14: Illustration of non-simultaneous events in STR.

5. In relativistic mechanics, there is no inertial reference frame that would be preferred. That is, every inertial reference frame can be employed to describe the laws of physics. In other words, the laws of physics will have the same form in every inertial reference frame; hence, the laws are the same for all observers moving at a uniform speed. For instance, Newton's second law in the frame S has the form $F = ma$, which will have the same form $F' = ma'$

in another frame S' that is moving at a uniform speed v relative to S. The relativity of time and space allows Newtonian and Maxwell mechanics to be the same for all observers in an inertial reference frame.

IV. LORENTZ TRANSFORMATIONS

In this section, we will introduce the Lorentz transformations in relativistic mechanics, as also discussed in Refs. [11, 12, 15, 16]. Lorentz velocity transformations using Cartesian coordinates, for the general case, are derived in this study.

A. Lorentz Transformations of Special Theory of Relativity

The Lorentz transformations preserve the invariant measure $(ds)^2$, representing the interval between two events in the space-time of the special theory of relativity.

The Lorentz transformations are linear with respect to the coordinates. Furthermore, these transformations converge to Galilean transformations in the limit when $v/c \to 0$. Often in the literature, the following notations are introduced:

$$\beta = \frac{v}{c} \qquad (76)$$
$$\gamma = \frac{1}{\sqrt{1-\beta^2}}$$

Consider $S(t, x, y, z)$ and $S'(t', x', y', z')$ represent two inertial reference frames (see Figure 15). S' is the moving frame with constant velocity $\mathbf{V} = (v, 0, 0)$ along the x axis and S is at rest. Initially (at $t = 0$), the origins and the axes of the two frames coincide. A and B are two events occurring at different spatial and temporal locations in space-time. A is the origin of S; therefore, initially at $t = 0$, the event A has the space-time coordinates $(0, 0, 0, 0)$, and the event B, occurring at some latter time, has the space-time coordinates $(ct, x, 0, 0)$ and $(ct', x', 0, 0)$, respectively, relative to S and S'.

For an observer at S, the x coordinate of the event B relative to origin of S is given as

$$x = L_{OO'} + L^{(p)}_{O'B} \tag{77}$$

In Eq. (77), $L_{OO'}$ represents the distance travelled by the observer at rest in S' during the time t measured by the observer's clock at rest in S. Therefore, the observer at rest in S measures the distance $L_{OO'} = vt$. Using the time dilation principle, $t = \gamma t'$, and we can write

$$L_{OO'} = \gamma v t' \tag{78}$$

where t' is the time measured by the observer's clock at rest in S'.

To measure the distance $L^{(p)}_{O'B}$, two events O' and B are considered, where O', initially, coincides with the origin of S', which is the moving inertial reference frame. For the observer at rest in S, moving relative to O' and B, the distance between the events B and O' is the proper length, $L^{(p)}_{O'B}$. On the other hand, for the observer at rest in S', moving relative to S, the length measurement is $L_{O'B} = x'$. Using the length contraction principle:

$$L^{(p)}_{O'B} = \gamma L_{O'B} = \gamma x' \tag{79}$$

Substituting Eq. (78) and Eq. (80) into Eq. (77), we obtain:

$$x = \gamma \left(x' + vt'\right) \tag{80}$$

which is often also found in the form

$$x = \gamma \left(x' + \beta c t'\right) \tag{81}$$

because $\beta c = v$.

Consider again two inertial reference frames $S(t, x, y, z)$ and $S'(t', x', y', z')$. S' is the moving frame with constant velocity $\mathbf{V} = (v, 0, 0)$ along the x axis and S is at rest. Initially (at $t = 0$), the origins and the axes of the two frames coincide (event O). Suppose that at $t = 0$ a light signal is sent from B (event B) along the x axis toward the origin O of the inertial frame at rest S. The event B has the space-time coordinates $(ct, x, 0, 0)$ relative to S, and it has the coordinates $(ct', x', 0, 0)$ relative to S'.

During the time interval t measured by the observer's clock at rest in S, the light signal sent from B meets the observer at rest in O' (event O'). During that time, the observer at rest in S' has travelled with speed $v < c$ a shorter distance relative to the distance travelled by the light sent from B. Therefore,

$$ct = L^{(p)}_{OO'} + L_{BO'} \tag{82}$$

where ct is the distance between the events O and B. $L_{BO'}$ is the distance travelled by the light signal between the events B and O' measured by the observer at rest in S:

$$L_{BO'} = ct' = \gamma ct \tag{83}$$

where t' is time measured by the moving clock in S', and t is the time measured by the observer's clock at rest in S; that is, t is the proper time. Using the time dilation principle: $t' = \gamma t$.

$L^{(p)}_{OO'}$ is the distance between the events O and O' measured by the observer at rest in S' during the time t' of the clock at rest in S':

$$L_{OO'} = vt' = v\frac{x'}{c} \tag{84}$$

where t' is the time that light travels the distance x' (which is the position of event B relative to S'), and thus, $t' = x'/c$. The observer at rest in S and relative to the events O and O' measures the proper distance $L(p)_{OO'} = \gamma L_{OO'}$, based on the length contraction principle. Therefore,

$$L^{(p)}_{OO'} = \gamma v\frac{x'}{c} \tag{85}$$

Substituting Eq. (83) and Eq. (85) into Eq. (82), we obtain

$$ct = \gamma ct' + \gamma v \frac{x'}{c} \tag{86}$$

or

$$ct = \gamma \left(ct' + \beta x' \right) \tag{87}$$

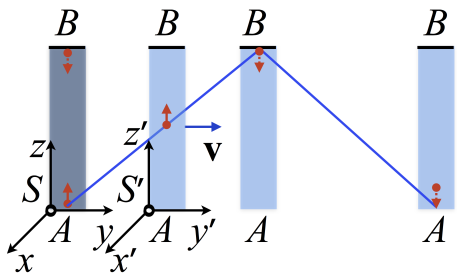

FIG. 15: $S(t, x, y, z)$ and $S'(t', x', y', z')$ represent two inertial reference frames. S' is the moving frame with constant velocity $\mathbf{V}$ along the x axis and S is at rest. Initially(at $t = 0$), the origins and the axes of the two frames coincide. A and B are two events occurring at different space-time locations, where A is the origin of S, therefore, initially at $t = 0$ A has the coordinates $(0, 0, 0, 0)$, and B at some arbitrary time has the space-time coordinates (t, x, y, z) and (t', x', y', z'), respectively, at S and S'.

B. Lorentz Coordinates Transformations

Since the inertial frame S' is not moving along y and z, there is no effect of the length contraction and time dilation along these directions. Therefore, the Lorentz transformations can be summarised as:

$$ct = \gamma \left(ct' + \beta x' \right) \tag{88}$$

$$x = \gamma \left(x' + \beta ct' \right)$$

$$y = y'$$

$$z = z'$$

Eq. (88) represents the Lorentz coordinate transformations from S' to S. To obtain the inverse Lorentz transformations, we can mutually switch the coordinates with and without prime, and change the sign before β (that is, $\beta \to -\beta$):

$$ct' = \gamma\,(ct - \beta x) \tag{89}$$

$$x' = \gamma\,(x - \beta ct)$$

$$y' = y$$

$$z' = z$$

The Lorentz transformations can be generalised for any arbitrary orientation of the velocity of the moving inertial frame with respect to the axes, as follows for the transformations from the moving frame S' to the frame at rest S:

$$ct = \gamma\,(ct' + \boldsymbol{\beta} \cdot \mathbf{r}') \tag{90}$$

$$\mathbf{r} = \mathbf{r}' + \frac{(\boldsymbol{\beta} \cdot \mathbf{r}')\,\boldsymbol{\beta}(\gamma - 1)}{\beta^2} + \boldsymbol{\beta}\gamma ct'$$

Here, the vector $\boldsymbol{\beta} = \mathbf{V}/c$ describes three independent motions of the inertial frame along the x, y and z axes, respectively:

$$\boldsymbol{\beta} = \beta_x \mathbf{i} + \beta_y \mathbf{j} + \beta_z \mathbf{k} \tag{91}$$

Here, $(\mathbf{i}, \mathbf{j}, \mathbf{k})$ denote the unit vectors along the x, y, and z axes, respectively. Therefore, Eq. (90) can be projected along each direction of the motion (namely x, y, and z) as the following, knowing that $\mathbf{r}' = x'\mathbf{i} + y'\mathbf{j} + z'\mathbf{k}$:

$$\begin{aligned}
ct &= \gamma\left(ct' + (\beta_x\mathbf{i} + \beta_y\mathbf{j} + \beta_z\mathbf{k}) \cdot (x'\mathbf{i} + y'\mathbf{j} + z'\mathbf{k})\right) \\
x &= x' + \frac{(\beta_x\mathbf{i} + \beta_y\mathbf{j} + \beta_z\mathbf{k}) \cdot (x'\mathbf{i} + y'\mathbf{j} + z'\mathbf{k})\,\beta_x(\gamma - 1)}{\beta^2} + \beta_x\gamma ct' \\
y &= y' + \frac{(\beta_x\mathbf{i} + \beta_y\mathbf{j} + \beta_z\mathbf{k}) \cdot (x'\mathbf{i} + y'\mathbf{j} + z'\mathbf{k})\,\beta_y(\gamma - 1)}{\beta^2} + \beta_y\gamma ct' \\
z &= z' + \frac{(\beta_x\mathbf{i} + \beta_y\mathbf{j} + \beta_z\mathbf{k}) \cdot (x'\mathbf{i} + y'\mathbf{j} + z'\mathbf{k})\,\beta_z(\gamma - 1)}{\beta^2} + \beta_z\gamma ct'
\end{aligned} \tag{92}$$

Assuming that the motions along x, y and z axes are independent, then

$$\begin{aligned}
&\text{along } x\text{-axis:} \quad \boldsymbol{\beta} = \beta_x\mathbf{i} \\
&\text{along } y\text{-axis:} \quad \boldsymbol{\beta} = \beta_y\mathbf{j} \\
&\text{along } z\text{-axis:} \quad \boldsymbol{\beta} = \beta_z\mathbf{k}
\end{aligned} \tag{93}$$

Hence, if the motion of the inertial frame is along the x-axis, $\beta_y = \beta_z = 0$, and Eq. (90) reduces to Eq. (88). If the motion of the inertial frame is along the y-axis, $\beta_x = \beta_z = 0$, and Eq. (90) reduces to

$$\begin{aligned}
ct &= \gamma\left(ct' + \beta_y y'\right) \\
x &= x' \\
y &= \gamma\left(y' + \beta_y ct'\right) \\
z &= z'
\end{aligned} \tag{94}$$

Furthermore, if the motion of the inertial frame is along the z-axis, $\beta_x = \beta_y = 0$, Eq. (90) can be written as

$$ct = \gamma\left(ct' + \beta_z z'\right) \tag{95}$$

$$x = x'$$

$$y = y'$$

$$z = \gamma\left(z' + \beta_z ct'\right)$$

Similarly, the inverse Lorentz transformations from the frame at rest S to the moving frame S' are given as:

$$\begin{aligned} ct' &= \gamma\left(ct - \boldsymbol{\beta}\cdot\mathbf{r}\right) \\ \mathbf{r}' &= \mathbf{r} + \frac{(\boldsymbol{\beta}\cdot\mathbf{r})\,\boldsymbol{\beta}(\gamma-1)}{\beta^2} - \boldsymbol{\beta}\gamma ct \end{aligned} \tag{96}$$

C. Lorentz Velocity Transformations

It is straightforward to obtain the Lorentz velocity transformations. For that, we can take the derivative for the time t in Eq. (90):

$$\frac{d\mathbf{r}}{dt} = \frac{d\mathbf{r}'}{dt} + \frac{\left(\boldsymbol{\beta}\cdot\dfrac{d\mathbf{r}'}{dt}\right)\boldsymbol{\beta}(\gamma-1)}{\beta^2} + \boldsymbol{\beta}\gamma c\frac{dt'}{dt} \tag{97}$$

From Eq. (96), if we take the derivative for t, we will get

$$c\frac{dt'}{dt} = \gamma\left(c - \boldsymbol{\beta}\cdot\frac{d\mathbf{r}}{dt}\right) \tag{98}$$

Combining Eq. (97) with Eq. (98), we find

$$\begin{aligned} \mathbf{v} &= \frac{d\mathbf{r}'}{dt'}\frac{dt'}{dt} + \frac{\left(\boldsymbol{\beta}\cdot\dfrac{d\mathbf{r}'}{dt'}\dfrac{dt'}{dt}\right)\boldsymbol{\beta}(\gamma-1)}{\beta^2} + \boldsymbol{\beta}\gamma^2\left(c - \boldsymbol{\beta}\cdot\mathbf{v}\right) \\ &= \mathbf{v}'\frac{\gamma}{c}\left(c - \boldsymbol{\beta}\cdot\mathbf{v}\right) + \frac{\left(\boldsymbol{\beta}\cdot\mathbf{v}'\dfrac{\gamma}{c}\left(c-\boldsymbol{\beta}\cdot\mathbf{v}\right)\right)\boldsymbol{\beta}(\gamma-1)}{\beta^2} \\ &\quad + \boldsymbol{\beta}\gamma^2\left(c - \boldsymbol{\beta}\cdot\mathbf{v}\right) \\ &= \gamma\mathbf{v}' - \frac{\gamma}{c}(\boldsymbol{\beta}\cdot\mathbf{v})\mathbf{v}' + \frac{\gamma(\gamma-1)\left(\boldsymbol{\beta}\cdot\mathbf{v}'\right)}{\beta^2}\boldsymbol{\beta} \end{aligned} \tag{99}$$

$$
\begin{aligned}
&- \frac{\gamma(\gamma-1)\,(\boldsymbol{\beta}\cdot\mathbf{v}')\,(\boldsymbol{\beta}\cdot\mathbf{v})}{c\beta^2}\boldsymbol{\beta} \\
&+ \boldsymbol{\beta}\gamma^2\,(c - \boldsymbol{\beta}\cdot\mathbf{v})
\end{aligned}
$$

where $\mathbf{v}$ is the velocity measured in S and $\mathbf{v}'$ is the velocity measured in S'.

Multiplying both sides of Eq. (99) by $\boldsymbol{\beta}$, we get

$$
\begin{aligned}
\boldsymbol{\beta}\cdot\mathbf{v} = \gamma(\boldsymbol{\beta}\cdot\mathbf{v}') - \frac{\gamma}{c}(\boldsymbol{\beta}\cdot\mathbf{v})(\boldsymbol{\beta}\cdot\mathbf{v}') + \gamma(\gamma-1)\,(\boldsymbol{\beta}\cdot\mathbf{v}') \\
- \frac{\gamma(\gamma-1)\,(\boldsymbol{\beta}\cdot\mathbf{v}')\,(\boldsymbol{\beta}\cdot\mathbf{v})}{c} + \beta^2\gamma^2 c - \beta^2\gamma^2(\boldsymbol{\beta}\cdot\mathbf{v})
\end{aligned}
\tag{100}
$$

Or,

$$
(\boldsymbol{\beta}\cdot\mathbf{v})\left(1 + \frac{\gamma^2}{c}\,(\boldsymbol{\beta}\cdot\mathbf{v}') + \beta^2\gamma^2\right) = \gamma^2(\boldsymbol{\beta}\cdot\mathbf{v}') + \beta^2\gamma^2 c \tag{101}
$$

which is further written as

$$
\boldsymbol{\beta}\cdot\mathbf{v} = \frac{(\boldsymbol{\beta}\cdot\mathbf{v}') + (\boldsymbol{\beta}\cdot\boldsymbol{\beta})c}{1 + \frac{1}{c}\,(\boldsymbol{\beta}\cdot\mathbf{v}')} \tag{102}
$$

or,

$$
\mathbf{v} = \frac{\mathbf{v}' + \boldsymbol{\beta}c}{1 + \frac{1}{c}\,(\boldsymbol{\beta}\cdot\mathbf{v}')} \tag{103}
$$

For example, for the transformations from the moving frame S' to the frame at rest S such that $\boldsymbol{\beta} = (\beta, 0, 0)$, we obtain:

$$
\begin{aligned}
v_x &= \frac{v'_x + \beta c}{1 + \frac{\beta}{c} v'_x} \\
v_y &= v'_y \\
v_z &= v'_z
\end{aligned}
\tag{104}
$$

Similarly, the Lorentz velocity transformations for the inverse transformation from the frame at rest S to the moving frame S' are given as:

$$\mathbf{v}' = \frac{\mathbf{v} - \boldsymbol{\beta} c}{1 - \frac{1}{c}\left(\boldsymbol{\beta} \cdot \mathbf{v}\right)} \tag{105}$$

For $\boldsymbol{\beta} = (\beta, 0, 0)$, we obtain

$$v'_x = \frac{v_x - \beta c}{1 - \frac{\beta}{c} v_x} \tag{106}$$
$$v'_y = v_y$$
$$v'_z = v_z$$

Note that for non-relativistic motion (that is, $v \ll c$) we have

$$\frac{\beta}{c} \to 0, \quad \gamma \to 1 \tag{107}$$

Therefore, for this limit it is straightforward to show that the Lorentz transformations converge to Galilean transformations:

$$\mathbf{v} = \mathbf{v}' + \boldsymbol{\beta} c \tag{108}$$
$$\mathbf{v}' = \mathbf{v} - \boldsymbol{\beta} c$$

where $\boldsymbol{\beta} = \mathbf{V}/c$ and $\mathbf{V}$ is the velocity of inertial reference frame S' relative to S.

V. FOUR-DIMENSIONAL VECTORS AND METRIC TENSOR

The following framework is established in this study for the first time (to the best of our knowledge).

A. Space-time

The basic framework of the special theory of relativity is the *space-time* and its fibres over absolute time [14]. The time is an affine space $\mathbf{T}$ associated with the vector space $\mathbb{T} \otimes \mathbb{R}$, where $\mathbb{T}$ is the space of future-oriented time intervals and $\mathbb{R}$ is the real 1-dimensional space. (Here, $\otimes$ denotes the tensor product defined in the following.)

In general, a *manifold* $\mathbf{E}$ of the four-dimensional space is a smooth space with local regions defining a $\mathbb{R}^4$ real space for 4-dimensional vectors.

The space-time is an oriented 4-dimensional manifold $\mathbf{E}$ fibred over time by the absolute time map $t: \ \mathbf{E} \to \mathbf{T}$.

Furthermore, we consider the so-called *tangent space* $\mathbf{E}_t$ of space-time, consisting of the vectors tangent to the fibres, which are called *space-like*. Moreover, we also consider the so-called *cotangent space* (or the *dual space*) $\mathbf{E}_t^*$ of space-time, which is called *time-like*, consisting of forms vanishing on tangent vectors.

The local coordinated bases are denoted $\boldsymbol{\omega}^\mu$ and $\mathbf{e}_\mu$ (for $\mu = 0, 1, 2, 3$), respectively, for the cotangent and tangent space.

B. Four-dimensional Vectors

Often, the notion of the coordinates in four dimensions is introduced to describe the STR, which may not be all Cartesian coordinates. As such, the so-called *4-vector* is introduced, x^μ for $\mu = 0, 1, 2, 3$, where $x^0 = ct$ is the time coordinate and x^1, x^2, and x^3 are the spatial coordinates. $x^\mu = (x^0, x^1, x^2, x^3)$ is, in the older notation, the vector itself, also called a *contravariant vector*:

$$x^\mu = \left(x^0, x^1, x^2, x^3, \right) = (ct, x, y, z,) \tag{109}$$

On the other hand, x_μ is called the *covariant vector*:

$$x_\mu = (x_0, x_1, x_2, x_3,) = (ct, -x, -y, -z,) \tag{110}$$

Note that x^μ is an element of tangent space, $x^\mu \in \mathbf{E}_t$, and x_μ is an element of cotangent space, $x_\mu \in \mathbf{E}_t^*$. In some literature, x^μ is called a vector and x_μ is called a covector or 1-form. In this paper, we will use the term contravariant (or sometimes simply three-vector and four-vector) for any geometrical object $\mathbf{a} \in \mathbf{E}_t$ with its components labeled with superscripts (such as x^μ, $T^{\mu\nu}$) and covariant for any geometrical object $\mathbf{a}^* \in \mathbf{E}_t^*$ with its components labeled with subscripts (such as x_μ, $T_{\mu\nu}$).

Then, the Lorentz coordinates transformations given by Eq. (90) from S' to S can be written in more compact form:

$$\begin{aligned} x^0 &= \gamma\left((x^0)' - \beta_j (x^j)'\right) \\ x^i &= (x^i)' - \frac{\gamma - 1}{\beta^2}\left(\beta_j (x^j)'\right)\beta^i + \gamma\beta^i (x^0)', \quad i = 1, 2, 3 \end{aligned} \tag{111}$$

where

$$\beta_i = \frac{V_i}{c} = -\frac{V^i}{c}, \quad i = 1, 2, 3 \tag{112}$$

are the covariant spatial components and

$$\beta^i = \frac{V^i}{c}, \quad i = 1, 2, 3 \tag{113}$$

are the contravariant spatial components of three-dimensional vector $\boldsymbol{\beta}$. In Eq. (111), note that when Einstein's summation notation is used, the repeating indices, one raised and one lowered, are summed:

$$a_i b^i = a_1 b^1 + a_2 b^2 + a_3 b^3 \quad \text{(Latin indices)} \tag{114}$$

$$a_\mu b^\mu = a_0 b^0 + a_1 b^1 + a_2 b^2 + a_3 b^3 \quad \text{(Greek indices)}$$

In this paper, the Latin indices (such as $i, j, k, l, \cdots$) are taking the values $1, 2, 3$ and indicate three-dimensional space; the Greek indices (such as $\mu, \nu, \sigma, \lambda, \cdots$) are taking the values $0, 1, 2, 3$ and indicate four-dimensional space-time. Besides, Einstein's summation notation is used throughout this paper; the repeating indices, one raised and one lowered, are summed as in Eq. (114).

Eq. (111) represents the Lorentz transformations of the 4-vector coordinates from S' to S. Similarly, the Lorentz transformations given by Eq. (96) can be written in the following form to give the Lorentz transformations of the 4-vector coordinates from S to S':

$$(x^0)' = \gamma \left(x^0 + \beta_j x^j\right) \tag{115}$$
$$(x^i)' = x^i - \frac{\gamma - 1}{\beta^2} \left(\beta_j x^j\right) \beta^i - \gamma \beta^i x^0, \quad i = 1, 2, 3$$

Often it is required to create a so-called *space-time graph*, where time ($x^0 = ct$) is one coordinate and the displacement other coordinates (x^i for $i = 1, 2, 3$); often, the space-time graph is also called the *chart* (x^0, x^i), for $i = 1, 2, 3$.

A path in that space-time graph is called a *world-line*. The world-line is a one-dimensional curve in the four-dimensional space-time (see Figure 16), which can be described by a parameter, namely λ. The world-line represents the motion on the manifold $\mathbf{E}$, which is defined as a one-dimensional time-like sub-manifold $l \subset \mathbf{E}$. The phase space of a classical particle is defined as the subspace $\mathbf{E}_u \in J(\mathbf{E}, 1)$ of all 1-jet motions. Here, $J(\mathbf{E}, 1)$ represents the manifold of 1-jets, where the 1-jet of one-dimensional sub-manifolds in $\mathbf{E}$ at $x^\mu \in \mathbf{E}$ is the equivalence class of one-dimensional sub-manifolds intersecting at x^μ with a contact of order one. Projection of $J(M, 1)$ in manifold $\mathbf{E}$ is defined as

$$P_1 : \; J(\mathbf{E}, 1) \to \mathbf{E} \tag{116}$$

A natural-fibered isomorphism over $\mathbf{E}$ of the 1-jets bundle with the Grassmannian bundle of 1-dimension is generated as

$$G(\mathbf{E},1):\ \psi \to L_\psi \tag{117}$$

where $\psi \in J(\mathbf{E},1)$ and $L_\psi \subset \mathbf{E}_t(\bar{\psi})$, where $\mathbf{E}_t(\bar{\psi})$ is the tangent space at $\bar{\psi} = P_1(\psi)$ of 1-dimensional sub-manifolds generating ψ.

Moreover, $\psi = j_1 l(\psi) \in J(\mathbf{E},1)$ is in $\mathbf{E}_u$ if and only if $L_\psi = \mathbf{E}_{t,\bar{\psi}} l$ lies inside the light cone. A motion line l can be expressed locally in terms of a space-time chart x^μ as

$$x^i = l^i(x^0) \tag{118}$$

$$l^i:\ \mathbb{R} \to \mathbb{R}$$

Every space-time chart on manifold $\mathbf{E}$ is a local fibre chart of the form $(x^\mu; x_0^i) \in \mathbf{E}_u$ (for $\mu = 0,1,2,3$, $i = 1,2,3$), where x_0^i is the initial point in the spatial space of the space-time.

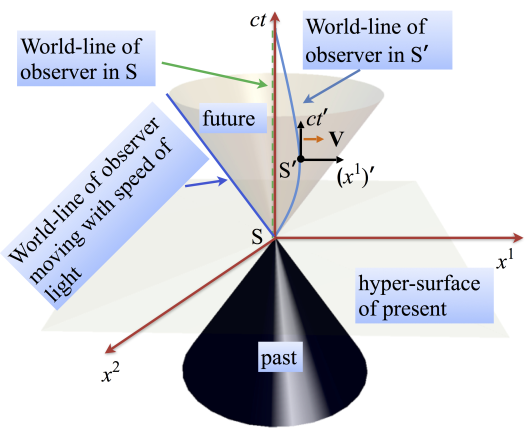

FIG. 16: A space-time graph of the observer at rest in S; observer at rest in S' moving with constant speed $v < c$ relative to S; world-line of a light ray.

Every point of the world-line can be represented by four coordinates

$$x^\mu(\lambda) \equiv \left(x^0(\lambda),\, x^1(\lambda),\, x^2(\lambda),\, x^3(\lambda)\right) \tag{119}$$

for a given value of λ.

At every point of the world-line, we can draw a tangent vector characterised by a magnitude and direction, as illustrated in Figure 17. We can determine an event A on the curve, where the tangent vector is required. Then, we can take any other close-by event B_1 on the curve. The arc length along the curve from A to B_1 defines λ. Besides, $\mathbf{v}_{AB_1}$ denotes the vector from A to B_1. As the event B_1 moves towards A through events B_2 and so on, the parameter λ goes to zero; hence the vector $\mathbf{v}_{AB_i}$, by definition, converges to the tangent vector at event point A to the curve:

$$\mathbf{u} = \lim_{\lambda \to 0} \mathbf{v}_{AB_i} = \left(\frac{dP}{d\lambda} \right)_{\lambda=0} \tag{120}$$

with components u^μ for $\mu = 0, 1, 2, 3$. Here, $v_{AB_i} = P_{B_i} - P_A$ is a 4-vector in a one-dimensional curve $P(\lambda)$ in 4-dimensional space-time with tail at the event A and tip at the event B_i on the curve. Note that the parameter λ is a measure of the length along the curve from A to B_i; hence, as point B_i moves toward A, the length of the arc and thus λ goes to zero.

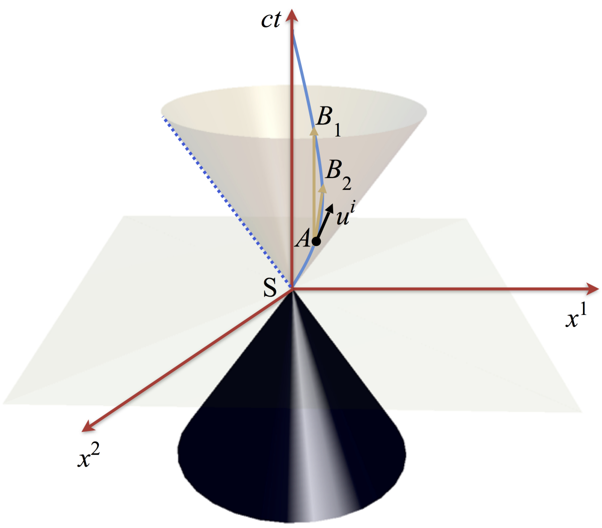

FIG. 17: Illustration of the tangent vector to the world-line.

In general, a 4-vector $\mathbf{a} \in \mathbf{E}_t$ is defined as a vector a^μ (for $\mu = 0, 1, 2, 3$) that has a *time-like* component, a^0, and three *space-like* components, a^i (for $i = 1, 2, 3$):

$$a^\mu = \left(a^0,\, a^i\right), \quad i = 1, 2, 3 \tag{121}$$

which is the so-called contravariant vector. The covariant form of that vector, $\mathbf{a}^* \in \mathbf{E}_t^*$, is written as

$$a_\mu = \left(a_0,\, a_i\right), \quad i = 1, 2, 3 \tag{122}$$

C. Tangent 4-vector

By definition, Eq. (95) determines the tangent 4-vector at any event of the world-line with components (u^0, u^1, u^2, u^3). Note that the tangent 4-vector of the world-line of a light ray is such that its square is equal to zero. Furthermore, the tangent 4-vector has no extension in the Minkowski space-time. The drawn arrows are for illustration. For a time-like curve, the proper time τ is the parameter λ, and the coordinates in the laboratory frame S are expressed as:

$$x^0 = ct(\tau),\ x^1 = x(\tau),\ x^2 = y(\tau),\ x^3 = z(\tau) \tag{123}$$

D. Basis Vectors

The set of basis vectors can be defined in the four-dimensional space-time given by 4-vectors $\mathbf{e}_\mu$ ($\mu = 0, 1, 2, 3$) with components such that:

$$\mathbf{e}_\mu \cdot \mathbf{e}_\nu = \begin{cases} 0, & \mu \neq \nu \\ 1, & \mu = \nu = 0 \\ -1, & \mu = \nu = 1, 2, 3 \end{cases} \tag{124}$$

Then, any 4-vector can be expressed in terms of its projections along the set of basis vectors $\mathbf{e}_\mu$ (for $\mu = 0, 1, 2, 3$). For example, the position ($\mathbf{x} \in \mathbf{E}_t$) of an event on the world-line curve $P(\tau)$

is defined as

$$\mathbf{x} = x^{\mu}(\tau)\mathbf{e}_{\mu} = (ct, \mathbf{r}) \tag{125}$$

where Einstein's notation is used that the repeated indices, one raised and one lowered, are summed. That is:

$$\mathbf{x} = x^{\mu}(\tau)e_{\mu} = x^{0}(\tau)\mathbf{e}_{0} + x^{1}(\tau)\mathbf{e}_{1} + x^{2}(\tau)\mathbf{e}_{2} + x^{3}(\tau)\mathbf{e}_{3} \tag{126}$$

That system of the 4-vector basis is called *coordinate basis*.

E. Four-velocity

The tangent to the curve is the so-called *four-velocity*, $\mathbf{u}$, of the particle moving along the curve. Using Eq. (120), we obtain

$$\begin{aligned} u^{0} &= \frac{d(ct)}{d\tau} = \gamma c \\ u^{i} &= \frac{dx^{i}}{d\tau} = \gamma v^{i}, \quad i = 1, 2, 3 \end{aligned} \tag{127}$$

In Eq. (127), v^{i} is the standard three-velocity in the laboratory frame defined as

$$v^{i} = \frac{dx^{i}}{dt}, \quad i = 1, 2, 3 \tag{128}$$

and $\mathbf{v}^2 = (v^1)^2 + (v^2)^2 + (v^3)^2$.

The particle's 4-velocity (see Eq. (127)) is a function of the parameter λ, and hence a continuous set of the 4-velocity can be defined once for each value of λ, forming in that way a so-called *vector field*.

Also, the 4-velocity given by Eq. (127) can be defined in terms of the coordinate basis vectors as

$$\mathbf{u} = \frac{dP(\tau)}{d\tau} = \frac{dx^{\mu}(\tau)}{d\tau}\mathbf{e}_{\mu} = u^{\mu}\mathbf{e}_{\mu} \tag{129}$$

The 4-velocity's magnitude is a scalar (similar to the three-velocity) with value varying as a function of parameter λ. The set of magnitude values forms the so-called *scalar field*.

F. Scalar Product

In general, combination of the covariant vector field $\mathbf{a}^* \in \mathbf{E}_t^*$ with components a_μ and a contravariant vector field $\mathbf{b} \in \mathbf{E}_t$ (a 4-vector) with components b^μ gives the quantity $\langle \mathbf{a}^*, \mathbf{b} \rangle$, which is a number giving the scalar product.

Furthermore, it gives the number of surfaces of $\mathbf{a}^*$ that are pierced by $\mathbf{b}$, and it is given as:

$$\langle \mathbf{a}^*, \mathbf{b} \rangle \equiv \mathbf{a}^* \cdot \mathbf{b} = a_\mu b^\mu \tag{130}$$

where the summation of the repeating indices, one raised and one lowered, is assumed.

G. Tensor Product

Scalars, contravariant vectors, and covariant vectors are examples of geometric objects called *tensors*. In a tensor, we insert p contravariant vectors and n covariant vectors to make a mapping onto a scalar. To describe the tensor, we say it has rank $\begin{pmatrix} n \\ p \end{pmatrix}$ given by the numbers p and n, where n is the number of possible covariant forms insertions $(\sigma, \lambda, \cdots, \beta)$ and p is the number of possible vector insertions $(u, v, \cdots, w)$. For example, the 4-velocity vector $u^\mu = (\gamma c, \gamma \mathbf{v})$ is a tensor of rank $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$, since contracting it with a covariant form, $u^\mu u_\mu$, gives a scalar, as shown in the following discussion.

For any two vectors, $\mathbf{a}$ and $\mathbf{b}$, a second-rank tensor can be constructed by the operation called *tensor product*, $\mathbf{T} = \mathbf{a} \otimes \mathbf{b}$ as the following:

$$T^{\mu\nu} = a^\mu b^\nu \tag{131}$$

$$T_{\mu\nu} = a_\mu b_\nu$$

Note that the tensor product has as output a number when the two vectors $(\mathbf{a}, \mathbf{b})$ and two covariant forms $(\boldsymbol{\alpha}, \boldsymbol{\beta})$ are involved:

$$(\mathbf{a} \otimes \mathbf{b})(\boldsymbol{\alpha}, \boldsymbol{\beta}) = \langle \mathbf{a}, \boldsymbol{\alpha} \rangle \langle \mathbf{b}, \boldsymbol{\beta} \rangle \tag{132}$$

H. Wedge Product

For any two four-vectors $\mathbf{a}$ and $\mathbf{b}$ ($\mathbf{a}, \mathbf{b} \in \mathbf{E}_t$), or two covariant vectors $\mathbf{a}^*$ and $\mathbf{b}^*$ ($\mathbf{a}^*, \mathbf{b}^* \in \mathbf{E}_t^*$), the so-called *wedge product* can be constructed as the following:

$$\mathbf{T} \equiv \mathbf{a} \wedge \mathbf{b} = \mathbf{a} \otimes \mathbf{b} - \mathbf{b} \otimes \mathbf{a} \tag{133}$$

$$\mathbf{T}^* \equiv \mathbf{a}^* \wedge \mathbf{b}^* = \mathbf{a}^* \otimes \mathbf{b}^* - \mathbf{b}^* \otimes \mathbf{a}^*$$

where $\otimes$ defines a tensor product as given by Eq. (131) and $\mathbf{T} \in \mathbf{E}_t \otimes \mathbf{E}_t$, $\mathbf{T}^* \in \mathbf{E}_t^* \otimes \mathbf{E}_t^*$. The components of $\mathbf{T}$ and $\mathbf{T}^*$ are given as

$$(a \wedge b)^{\mu\nu} = a^\mu b^\nu - b^\mu a^\nu \tag{134}$$

$$(a \wedge b)_{\mu\nu} = a_\mu b_\nu - b_\mu a_\nu$$

which indicates that the wedge product is an asymmetric tensor of second-rank.

I. Four-gradient

If we have to deal with scalar fields, for instance, let say $\phi(x^0, \mathbf{r})$ and $\phi'((x^0)', \mathbf{r}')$ are two 4-scalars measured, respectively, in the inertial frames S and S', then $\phi(x^0, \mathbf{r}) = \phi'((x^0)', \mathbf{r}')$ (based on Lorentz invariant principle). That is, the scalar is an invariant quantity. However, its 4-gradient, called the *4-gradient* of the scalar field, gives a covariant vector defined as:

$$A_\mu = \frac{\partial \phi(x^\mu)}{\partial x^\mu} \equiv \partial_\mu \phi(x^\mu) \, , \quad (\mu = 0, 1, 2, 3) \tag{135}$$

which transforms as a vector when going from the inertial frame S to S', or vice-versa, as described in the following discussion.

The gradient is an example of the covariant field vector (element of $\mathbf{E}_t^*$ space), defined as:

$$\partial_\mu = \partial_{\mathbf{e}_\mu} \equiv \frac{\partial}{\partial x^\mu} \tag{136}$$

where ∂_μ indicates the gradient of a scalar.

J. Dual Coordinate Basis

The gradient of the coordinates, $\boldsymbol{\omega}^\mu$, is defined as:

$$\boldsymbol{\omega}^\mu = \partial_{\mathbf{e}_\nu} x^\mu = \left(\frac{\partial x^\mu}{\partial x^0}, \frac{\partial x^\mu}{\partial x^1}, \frac{\partial x^\mu}{\partial x^2}, \frac{\partial x^\mu}{\partial x^3} \right) \tag{137}$$

We obtain that

$$\langle \mathbf{e}_\nu, \boldsymbol{\omega}^\mu \rangle = \mathbf{e}_\nu \cdot \boldsymbol{\omega}^\mu = \frac{dx^\mu}{dx^\nu} = \delta_\nu^{\ \mu} \tag{138}$$

which provides the dual coordinate basis of the dual space. In Eq. (138), $\delta_\nu^{\ \mu}$ is Kronecker's number: $\delta_\nu^{\ \mu} = 1$, if $\mu = \nu$, and $\delta_\nu^{\ \mu} = 0$, if $\mu \neq \nu$. Eq. (138) indicates that the vector basis $\mathbf{e}_\mu$ and the dual coordinate basis $\boldsymbol{\omega}^\mu$ are orthogonal. Therefore, any covariant vector $\mathbf{a}^* \in \mathbf{E}_t^*$ can be written as

$$\mathbf{a}^* = a_\mu \boldsymbol{\omega}^\mu \tag{139}$$

where

$$a_\mu = \mathbf{a}^* \cdot \mathbf{e}_\mu \tag{140}$$

K. Metric Tensor

A functional form is defined, allowing the conversion of a pair of 4-vectors into a scalar at every point in space-time. Hence, it determines the magnitude of a 4-vector. Thus, we have to determine the scalar product of two 4-vectors or the square of the length of a 4-vector. That procedure is called *mapping*, and the functional form is called *metric tensor*, denoted by $\mathbf{g}^*$. The metric tensor function is the basis vector $\mathbf{e}_\mu$. For that, assume a four-dimensional space-time oriented manifold $\mathbf{E}$ equipped with a metric tensor, $\mathbf{g}^*$, with signature $(+,-,-,-)$ of the following mapping:

$$\mathbf{g}^* : \ \mathbf{E} \to \mathbb{L}^2 \otimes (\mathbf{E}_t^* \otimes \mathbf{E}_t^*) \tag{141}$$

where $\mathbb{L}$ is the space of lengths. In coordinates,

$$g_{\mu\nu} = \mathbf{g}^*(\mathbf{e}_\mu, \mathbf{e}_\nu) = \mathbf{e}_\mu \cdot \mathbf{e}_\nu \tag{142}$$

In the case of the four-dimensional Minkowski space-time with coordinates of a point $x^\mu = (ct, \mathbf{r})$, where $\mathbf{r} = (x, y, z)$, consider an infinitesimal displacement, $dx^\mu = (cdt, d\mathbf{r})$. Then, the relative displacement between the two end points is given as:

$$ds = dx^\mu \mathbf{e}_\mu \tag{143}$$

The invariant interval in the four-dimensional Minkowski space-time is defined as:

$$\begin{aligned}(ds)^2 = ds \cdot ds &= (dx^\mu \mathbf{e}_\mu) \cdot (dx^\nu \mathbf{e}_\nu) \\ &= dx^\mu dx^\nu (\mathbf{e}_\mu \cdot \mathbf{e}_\nu) \\ &= g_{\mu\nu} dx^\mu dx^\nu \end{aligned} \tag{144}$$

Comparing the expression in Eq. (144) with

$$(ds)^2 = (cdt)^2 - (dx)^2 - (dy)^2 - (dz)^2 = dx^\mu dx_\mu \tag{145}$$

we obtain the metric tensor given as:

$$\mathbf{g}^* = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \tag{146}$$

with components given as:

$$g_{\mu\nu} = \begin{cases} 1, & \text{if } \mu = \nu = 0 \\ -1, & \text{if } \mu = \nu = 1, 2, 3 \\ 0, & \text{if } \mu \neq \nu \end{cases} \tag{147}$$

Here, $\mathbf{g}^* \in \mathbf{E}_t^*$. It is important to notice that the space-time geometry is equipped with a metric tensor given by Eq. (146). Besides, the metric corresponds to a flat geometry, and hence, for the spatial dimensions of Minkowski space-time, we use a Galilean coordinate system to represent the three-dimensional geometry, given by Eq. (145); that is, the spatial dimensions of the space-time obey the Pythagorean theorem. (Note that Eq. (146) or Eq. (147) can also be obtained by combining Eq. (142) and Eq. (124).)

Furthermore, comparing Eq. (144) and Eq. (145), we get

$$dx_\mu = g_{\mu\nu} dx^\nu \tag{148}$$

In general, the scalar product of any two vector fields (4-vectors) $\mathbf{a} \in \mathbf{E}_t$ and $\mathbf{b} \in \mathbf{E}_t$ is defined as

$$\begin{aligned} \langle \mathbf{a}, \mathbf{b} \rangle = \mathbf{a} \cdot \mathbf{b} &= g_{\mu\nu} a^\mu b^\nu \\ &= a_\nu b^\nu \equiv \langle \mathbf{a}^*, \mathbf{b} \rangle \\ &= a^\mu b_\mu \equiv \langle \mathbf{a}, \mathbf{b}^* \rangle \end{aligned} \tag{149}$$

where $\mathbf{a}^* \in \mathbf{E}_t^*$ and $\mathbf{b}^* \in \mathbf{E}_t^*$.

If $\mathbf{a}$ and $\mathbf{b}$ are the same vector fields, then the square of the length of the 4-vector is obtained as:

$$\mathbf{a} \cdot \mathbf{a} = a^2 = g_{\mu\nu} a^\mu a^\nu = a_\nu a^\nu = a^\mu a_\mu \tag{150}$$

In Eq. (149) and Eq. (150), the following identities are used for any arbitrary 4-vector $a^\mu \in \mathbf{E}_t$ for $\mu = 0, 1, 2, 3$:

$$a_\nu = g_{\mu\nu} a^\mu, \quad a_\mu = g_{\mu\nu} a^\nu \tag{151}$$

where a_μ are the components of the covariant vector $\mathbf{a}^* \in \mathbf{E}_t^*$ and the symmetry of metric tensor is used $g_{\mu\nu} = g_{\nu\mu}$.

The inverse metric can be defined as

$$\mathbf{g}^{-1}\mathbf{g} \equiv \mathbf{g} \cdot \mathbf{g}^* = \mathbf{I} \tag{152}$$

or in terms of the components as

$$g^{\mu\nu} g_{\nu\sigma} = \delta^\mu_\sigma \tag{153}$$

That is,

$$g^{\mu\nu} = g_{\mu\nu} \tag{154}$$

and g, with signature $(+, -, -, -)$ and components $g^{\mu\nu}$, represents the following mapping:

$$\mathbf{g} : \ \mathbf{E} \to \mathbb{L}^2 \otimes (\mathbf{E}_t \otimes \mathbf{E}_t) \tag{155}$$

Here, $\mathbf{I}$ is the diagonal identity matrix. Besides, the components $g_{\mu\nu}$ and $g^{\mu\nu}$ are dimensionless.

The contravariant vector u^μ can be converted into a covariant vector and vice versa as follows:

$$u^\mu = g^{\mu\nu} u_\nu, \quad u_\mu = g_{\mu\nu} u^\nu \tag{156}$$

Therefore, for any two 4-vectors $\mathbf{a}$ and $\mathbf{b}$ ($\mathbf{a}, \mathbf{b} \in \mathbf{E}_t$) or covariant vectors $\mathbf{a}^*$ and $\mathbf{b}^*$ ($\mathbf{a}^*, \mathbf{b}^* \in \mathbf{E}_t^*$), we get

$$\begin{aligned} \mathbf{a} \cdot \mathbf{b} &= g_{\mu\nu} a^\mu b^\nu = a_\nu b^\nu \\ &= a^\mu b_\mu = a_\mu b_\nu g^{\mu\nu} \end{aligned} \tag{157}$$

Furthermore, combining Eq. (146) with Eq. (157), we obtain that

$$\begin{aligned} \mathbf{a} \cdot \mathbf{b} &= a^0 b^0 - a^1 b^1 - a^2 b^2 - a^3 b^3 \\ &= a_0 b_0 + a_1 b_1 + a_2 b_2 + a_3 b_3 \end{aligned} \tag{158}$$

or

$$\begin{aligned} \mathbf{a} \cdot \mathbf{b} &= a^0 b_0 + a^1 b_1 + a^2 b_2 + a^3 b_3 \\ &= a_0 b^0 + a_1 b^1 + a_2 b^2 + a_3 b^3 \end{aligned} \tag{159}$$

Using Eq. (158), the square of the magnitude of a vector can be determined as

$$a^\mu a_\mu \equiv a^2 = (a^0)^2 - a^i a_i \tag{160}$$

where the time components are related as $a^0 = a_0$ and the space components as $a^i = -a_i$:

$$\begin{aligned} a^i &= g^{ij} a_j = -a_i \\ a_i &= g_{ij} a^j = -a^i \end{aligned} \tag{161}$$

If the square of the magnitude of a vector is positive, then the vector is called a *time-like vector*; if the square of the magnitude of a vector is negative, it is called *space-like vector*, and if the square of the magnitude of the vector is zero, then the vector is called a *light-like vector*.

The space components of a 4-vector, concerning the rotations in Euclidean space that keep the time constant, form a 3-vector, denoted as $\mathbf{a}$. The time component, concerning the rotations in

Euclidean space, forms a three-scalar. Thus, the vectors, $\mathbf{a} \in \mathbf{E}_t$ and $\mathbf{a}^* \in \mathbf{E}_t^*$, are also written as:

$$a^\mu = (a^0, \mathbf{a}) \tag{162}$$

$$a_\mu = (a_0, -\mathbf{a})$$

In general, any 4-vector $\mathbf{a} \in \mathbf{E}_t$ can be written in terms of the *basis vector*, $\mathbf{e}_\mu$ (for $\mu = 0, 1, 2, 3$), as shown in the following discussion, as:

$$\begin{aligned} \mathbf{a} &= (a^0, a^1, a^2, a^3) \\ &= a^0\mathbf{e}_0 + a^1\mathbf{e}_1 + a^2\mathbf{e}_2 + a^3\mathbf{e}_3 \\ &= a^\mu \mathbf{e}_\mu \end{aligned} \tag{163}$$

where the summation of repeating indices, one raised and one lowered, is assumed.

The square of the magnitude of the four-velocity in any coordinate system is obtained as:

$$\begin{aligned} \mathbf{u} \cdot \mathbf{u} = u^2 &= g_{\mu\nu} u^\mu u^\nu \\ &= (u^0)^2 - (u^1)^2 - (u^2)^2 - (u^3)^2 \end{aligned} \tag{164}$$

where the four-velocity is $\mathbf{u} = (\gamma c, \gamma \mathbf{v})$ (see Eq. (127)). Therefore, we get

$$\mathbf{u} \cdot \mathbf{u} = \gamma^2 c^2 - \gamma^2 v^2 = c^2 \tag{165}$$

which gives the square of the 4-velocity magnitude of a particle moving with speed v in an inertial reference frame at rest relative to the particle. Eq. (165) indicates that the four-velocity is a time-like type of vector field and that the square of the 4-velocity magnitude is a Lorentz invariant (i.e., it is independent of the Lorentz system).

Any tensor of rank $\begin{pmatrix} n \\ p \end{pmatrix}$ can be converted using the metric tensor $g_{\mu\nu}$ as:

$$T^\mu_{\sigma\nu} = g_{\gamma\nu} T^{\mu\gamma}_\sigma \tag{166}$$

L. The 4-divergence

In Minkowski space-time, the 4-divergence operator, often denoted by either $\boldsymbol{\nabla}$ or $\Box$, has the following components:

$$\Box_\mu \equiv \nabla_\mu = \frac{\partial}{\partial x^\mu}\boldsymbol{\omega}^\mu = \left(\frac{\partial}{\partial x^0}, \boldsymbol{\nabla}\right) \tag{167}$$

where $\boldsymbol{\nabla}$ is three-dimensional divergence operator:

$$\begin{aligned} \boldsymbol{\nabla} &= \left(\frac{\partial}{\partial x^1}, \frac{\partial}{\partial x^2}, \frac{\partial}{\partial x^3}\right) \\ &\equiv \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}\right) \end{aligned} \tag{168}$$

$\boldsymbol{\omega}^\mu$ is the set of dual coordinate basis.

For example, if we consider a contravariant vector $\mathbf{A} \in \mathbf{E}_t$, which can be expressed as

$$\mathbf{A} = A^\nu \mathbf{e}_\nu \tag{169}$$

Then, the 4-divergence of the 4-vector $\mathbf{A}$ is given as

$$\begin{aligned} \boldsymbol{\nabla} \cdot \mathbf{A} = \frac{\partial A^\nu}{\partial x^\mu}\left(\boldsymbol{\omega}^\mu \cdot \mathbf{e}_\nu\right) &= \frac{\partial A^\mu}{\partial x^\mu} \equiv \nabla_\mu A^\mu \\ &= \frac{\partial A^0}{\partial x^0} + \frac{\partial A^1}{\partial x^1} + \frac{\partial A^2}{\partial x^2} + \frac{\partial A^3}{\partial x^3} \\ &= \frac{1}{c}\frac{\partial A^0}{\partial t} + \frac{\partial A^1}{\partial x^1} + \frac{\partial A^2}{\partial x^2} + \frac{\partial A^3}{\partial x^3} \end{aligned} \tag{170}$$

where the orthogonality of $\boldsymbol{\omega}^\mu$ and $\mathbf{e}_\nu$ given by Eq. (138) is used.

In contravariant form, the divergence can be obtained using the metric tensor as follows:

$$\nabla^\mu = g^{\mu\sigma}\nabla_\sigma = \frac{\partial}{\partial x_\mu} = \left(\frac{\partial}{\partial x^0}, -\boldsymbol{\nabla}\right) \tag{171}$$

The contravariant divergence of the 4-vector $\mathbf{A}$ is given as

$$\nabla^\mu A^\mu = \frac{\partial A^\mu}{\partial x_\mu} \tag{172}$$

$$= \frac{\partial A^0}{\partial x_0} + \frac{\partial A^1}{\partial x_1} + \frac{\partial A^2}{\partial x_2} + \frac{\partial A^3}{\partial x_3}$$

$$= \frac{1}{c}\frac{\partial A^0}{\partial t} - \left(\frac{\partial A^1}{\partial x^1} + \frac{\partial A^2}{\partial x^2} + \frac{\partial A^3}{\partial x^3}\right)$$

The operator $\boldsymbol{\nabla}^2$, or $\Box^2$, is called the d'Alembertian operator defined as:

$$\Box^2 = \boldsymbol{\nabla}^2 = \boldsymbol{\nabla} \cdot \boldsymbol{\nabla} = g^{\mu\nu}\frac{\partial}{\partial x^\mu}\frac{\partial}{\partial x^\nu} \tag{173}$$

$$= \frac{\partial^2}{\partial (x^0)^2} - \left(\frac{\partial^2}{\partial (x^1)^2} + \frac{\partial^2}{\partial (x^2)^2} + \frac{\partial^2}{\partial (x^3)^2}\right)$$

$$\equiv \frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}\right)$$

Note that the $\boldsymbol{\nabla}^2$ operator applies to a four-dimensional scalar field.

M. Four-momentum

By definition, the linear four-momentum in any Lorentz coordinate system is defined by

$$\mathbf{p} = m\mathbf{u} \tag{174}$$

where m is the scalar mass, and $\mathbf{u}$ is the 4-velocity given by Eq. (127). Therefore, the time and space components of the linear 4-momentum are as follows:

$$\mathbf{p} = (\gamma mc, \gamma m\mathbf{v}) = \left(\frac{E}{c}, \mathbf{P}\right) \tag{175}$$

where $\mathbf{P}$ is the three-dimensional momentum and E is the total relativistic energy. Hence,

$$p^0 = \gamma mc, \quad p^i = \gamma mv^i, \quad i = 1, 2, 3 \tag{176}$$

The square of the magnitude of the linear 4-momentum is

$$\mathbf{p} \cdot \mathbf{p} = (m\mathbf{u}) \cdot (m\mathbf{u}) = m^2c^2 \tag{177}$$

Eq. (177) indicates that the square of the linear 4-momentum magnitude is a Lorentz invariant (*i.e.*, it is independent of the Lorentz system).

It can be seen that for particle velocity much less than the speed of light c, that is $v \ll c$, then

$$\gamma = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}} \to 1 \tag{178}$$

and hence the space component of the 4-momentum reduces to the non-relativistic linear three-momentum $m\mathbf{v}$.

The covariant vector of the four-momentum is given as:

$$p_0 = \gamma mc, \quad p_i = -\gamma m v_i, \quad i = 1, 2, 3 \tag{179}$$

The square of the magnitude of the 4-momentum can also be written as

$$\mathbf{p} \cdot \mathbf{p} = g_{\mu\nu} p^\mu p^\nu = p_\nu p^\nu = p^\mu p_\mu \tag{180}$$

N. Forces in the Special Theory of Relativity

The 4-velocity and 4-momentum are related to the kinematics of the special theory of relativity. In the following, we will describe the dynamics of the STR. In particular, we will assume that Newton's laws are valid for the objects at rest in the rest observer frame, and they are only approximately correct for the objects moving with speed v such that $v \ll c$. We introduce the four-force law in analogy to the generalisation of the four-velocity from the three-velocity.

Thus, by definition, the proper generalisation of the second law of Newton is given as follows [16]:

$$\frac{dp^\mu}{d\tau} = K^\mu, \quad \mu = 0, 1, 2, 3 \tag{181}$$

In Eq. (181), K^μ is the 4-vector force, also known as the *Minkowski force*. Note that the spatial components of the 4-vector force components (*i.e.*, K^i, for $i = 1, 2, 3$) are not the components of the three-vector force in the form

$$F^i = \frac{d(mv^i)}{dt}, \quad i = 1, 2, 3 \tag{182}$$

The exact form can be derived using the Lorentz transformations of the forces present in the system. However, they reduce to the force F^i in the limit of the non-relativistic particles, that is, for $v \ll c$ (or equivalently $\beta \to 0$). The time component of the 4-vector force relates to the rate of change of the particle's energy for any reference frame as follows:

$$K^0 = \frac{\gamma}{c}\frac{dE}{dt} \tag{183}$$

and the spatial components obey Newton's second law in the form:

$$\gamma \frac{dp^i}{dt} = \gamma F^i \tag{184}$$

Moreover, the magnitude of the 4-vector force is determined by its square, given as:

$$\mathbf{K} \cdot \mathbf{K} = K^2 = g_{\mu\nu} K^\mu K^\nu = -F^2_{\text{Newtonian}} \tag{185}$$

where $F_{\text{Newtonian}}$ is the magnitude of the three-vector Newtonian force.

There is no unique way to determine the expression of the proper relativistic force. Electromagnetism can justify the special theory of relativity, and it is possible to relate the electromagnetism forces to relativistic forces [15, 19].

Relating other forces in nature to Minkowski's relativistic force remains a problem. Often, two methods are argued to provide acceptable transformation of the properties of forces and hence give a correct relativistic form of the forces. The first one consists of accepting that there are only four fundamental forces in nature, namely *gravitational*, *weak nuclear*, *electromagnetic*, and *strong nuclear* forces. Then, a relativistic theory should be able to derive the correct expressions for each of these forces in a covariant form such that the properties of these forces are transferred correctly. The expressions of the forces given in their covariant form, valid in one inertial frame, can be assumed to be correct in all inertial frames. As such, they do not need to involve terms of the form $(v/c)^3$ that can vanish in some frame. The efforts continue [16].

The second method on determining the correct expression for relativistic force is to assume that the force can be defined as the rate of change of the momentum:

$$\frac{dp^i}{dt} = F^i, \quad i = 1, 2, 3 \tag{186}$$

In Eq. (186), p^i is the relativistic generalisation of the Newtonian non-relativistic momentum, given as

$$p^i = \gamma m v^i \tag{187}$$

Thus, it reduces to mv^i in the limit of $v \ll c$. Note that this approach has failed to give any results other than those predicted using the first method so far.

VI. LORENTZ TRANSFORMATIONS TENSOR

The Lorentz transformations of the coordinates can be given in terms of the basis vectors. For that, we consider two frames of coordinates, namely the frame S with coordinates x^μ ($\mu = 0, 1, 2, 3$) and the frame S' with coordinates $(x^\mu)'$ ($\mu = 0, 1, 2, 3$). Then, the Lorentz transformations from S to S' can be written as:

$$(x^\mu)' = L^\mu_{\ \nu} x^\nu \tag{188}$$

where

$$\mathbf{L} = \begin{pmatrix} \gamma & -\gamma\beta_x & -\gamma\beta_y & -\gamma\beta_z \\ -\gamma\beta_x & 1+(\gamma-1)\dfrac{\beta_x^2}{\beta^2} & (\gamma-1)\dfrac{\beta_x\beta_y}{\beta^2} & (\gamma-1)\dfrac{\beta_x\beta_z}{\beta^2} \\ -\gamma\beta_y & (\gamma-1)\dfrac{\beta_y\beta_x}{\beta^2} & 1+(\gamma-1)\dfrac{\beta_y^2}{\beta^2} & (\gamma-1)\dfrac{\beta_y\beta_z}{\beta^2} \\ -\gamma\beta_z & (\gamma-1)\dfrac{\beta_z\beta_x}{\beta^2} & (\gamma-1)\dfrac{\beta_z\beta_y}{\beta^2} & 1+(\gamma-1)\dfrac{\beta_z^2}{\beta^2} \end{pmatrix} \tag{189}$$

is the so-called *Lorentz transformation tensor* with components L^{μ}_{ν}. It can be seen that $\mathbf{L}$ is a symmetric matrix.

On the other hand, the Lorentz transformations from the coordinates in the frame S' to those in S are given as:

$$x^{\mu} = (L^{\mu}_{\ \nu})' \, (x^{\nu})' \tag{190}$$

where

$$\mathbf{L}' = \begin{pmatrix} \gamma & \gamma\beta_x & \gamma\beta_y & \gamma\beta_z \\ \gamma\beta_x & 1+(\gamma-1)\dfrac{\beta_x^2}{\beta^2} & (\gamma-1)\dfrac{\beta_x\beta_y}{\beta^2} & (\gamma-1)\dfrac{\beta_x\beta_z}{\beta^2} \\ \gamma\beta_y & (\gamma-1)\dfrac{\beta_y\beta_x}{\beta^2} & 1+(\gamma-1)\dfrac{\beta_y^2}{\beta^2} & (\gamma-1)\dfrac{\beta_y\beta_z}{\beta^2} \\ \gamma\beta_z & (\gamma-1)\dfrac{\beta_z\beta_x}{\beta^2} & (\gamma-1)\dfrac{\beta_z\beta_y}{\beta^2} & 1+(\gamma-1)\dfrac{\beta_z^2}{\beta^2} \end{pmatrix} \tag{191}$$

is the so-called *inverse Lorentz transformation tensor*, which can be obtained using Eq. (189), by replacing $\boldsymbol{\beta}$ with $-\boldsymbol{\beta}$.

It is easy to show that

$$\mathbf{L}\mathbf{L}' = \mathbf{L}'\mathbf{L} = \mathbf{I} \tag{192}$$

where $\mathbf{I}$ is a unitary diagonal matrix.

Following [12], a physical quantity can be defined in terms of A^{μ} and $A^{\mu\prime}$, respectively, in the inertial frame S and S', if between them there exists a relation as follows:

$$A^{\mu\prime} = L^{\mu}_{\ \nu} A^{\nu} \tag{193}$$

which is similar to Eq. (188). That is, $L^{\mu}_{\ \nu}$ is given as

$$L^{\mu}_{\ \nu} = \frac{\partial x^{\mu\prime}}{\partial x^{\nu}} \tag{194}$$

Then, the quantity A^{μ} is referred to as a contravariant vector in four-dimensional space. Furthermore, if $A^{\mu}(x)$ is a function of the point x in four-dimensional space, then $A^{\mu}(x)$ is called a contravariant field vector.

Moreover, if the quantity B_{μ} is transferred from the inertial frame S to S' according to

$$B'_{\mu} = \frac{\partial x^{\nu}}{\partial x^{\mu\prime}} B_{\nu} = L_{\mu}^{\ \nu} B_{\nu} \tag{195}$$

then B_{μ} is called a covariant vector, and if $B_{\mu}(x)$ is a function of the point x in four-dimensional space, B_{μ} is called a covariant field vector. Note that

$$L_{\mu}^{\ \nu} = g_{\mu\alpha} L^{\alpha}_{\ \beta} g^{\beta\nu} \tag{196}$$

Now, if we consider a second rank tensor, $C^{\mu\nu} = A^{\mu}B^{\nu}$, that transforms from inertial frame S to S' according to

$$\begin{aligned} C^{\mu\nu\prime} &= A^{\mu\prime}B^{\nu\prime} = \frac{\partial x^{\mu\prime}}{\partial x^{\alpha}} A^{\alpha} \frac{\partial x^{\nu\prime}}{\partial x^{\beta}} B^{\beta} = \frac{\partial x^{\mu\prime}}{\partial x^{\alpha}} \frac{\partial x^{\nu\prime}}{\partial x^{\beta}} C^{\alpha\beta} \\ &= L^{\mu}_{\ \alpha} L^{\nu}_{\ \beta} C^{\alpha\beta} \end{aligned} \tag{197}$$

Then, $C^{\mu\nu}$ is referred to as a contravariant second-rank tensor. Similarly, a second rank tensor, $C_{\mu\nu} = A_{\mu}B_{\nu}$, that transforms as follows, is called a covariant second-rank tensor:

$$\begin{aligned} C'_{\mu\nu} &= A'_{\mu}B'_{\nu} = \frac{\partial x^{\alpha}}{\partial x^{\mu\prime}} A_{\alpha} \frac{\partial x^{\beta}}{\partial x^{\nu\prime}} B_{\beta} = \frac{\partial x^{\alpha}}{\partial x^{\mu\prime}} \frac{\partial x^{\beta}}{\partial x^{\nu\prime}} C_{\alpha\beta} \\ &= L_{\mu}^{\ \alpha} L_{\nu}^{\ \beta} C_{\alpha\beta} \end{aligned} \tag{198}$$

VII. CONSERVATION LAWS

The expression given by Eq. (186) is a reasonable relativistic form of Newton's second law because it preserves the classical mechanics in the limit of non-relativistic particles, that is $v \ll c$, and furthermore, it requires the conservation of the linear momentum for an isolated system of a relativistic system of particles. Since the system is isolated, the total force is zero. Consider a system of two particles, namely A and B. Suppose that F_{AB} is the relativistic force exerted by particle A on particle B, and F_{BA} is the force exerted by particle B on particle A. Using Eq. (186), we get

$$\frac{d\mathbf{p}_A}{dt} = \mathbf{F}_{BA} \qquad (199)$$
$$\frac{d\mathbf{p}_B}{dt} = \mathbf{F}_{AB}$$

where $\mathbf{p}_A$ and $\mathbf{p}_B$ are the relativistic three-momenta of the particle A and B, respectively. Since, $\mathbf{F}_{AB} + \mathbf{F}_{BA} = 0$, then

$$\frac{d\mathbf{p}_A}{dt} + \frac{d\mathbf{p}_B}{dt} = 0 \qquad (200)$$

or equivalently

$$\frac{d(\mathbf{p}_A + \mathbf{p}_B)}{dt} = 0 \qquad (201)$$

Eq. (201) indicates that

$$\mathbf{p}_A + \mathbf{p}_B = \text{constant} \qquad (202)$$

Therefore, we can write that the linear momentum 4-vectors are equal before and after a physical process (such as a collision). That is, in terms of the components of the 4-momentum, we have:

$$(p^{\mu})_{\text{before}} = (p^{\mu})_{\text{after}}, \quad \mu = 0, 1, 2, 3 \qquad (203)$$

That is known as the principle of the *conservation law of the linear momentum* 4-vector.

It is important to note that the momentum 4-vector is not Lorentz transformation invariant; hence, it depends on the Lorentz system. Therefore, it transfers as a common 4-vector from a

Lorentz system (such as S) to another one (such as S') according to:

$$(p^\mu)' = L^\mu_\nu p^\nu, \quad \mu = 0, 1, 2, 3 \tag{204}$$

However, it is possible to find a measure that involves p^μ, which is invariant under the Lorentz transformation, and hence it is independent of the Lorentz system. That is, the square of the momentum 4-vector magnitude, $\mathbf{p} \cdot \mathbf{p}$ given as

$$\mathbf{p} \cdot \mathbf{p} = g_{\mu\nu} p^\mu p^\nu = p^\mu p_\mu = p_\nu p^\nu \tag{205}$$

The conservation of the square of the momentum 4-vector magnitude under the Lorentz transformation is written as:

$$p^\mu p_\mu = (p^\mu)'(p_\mu)' \tag{206}$$

where $p^\mu p_\mu$ is measured in the S frame and $(p^\mu)'(p_\mu)'$ is measured in the frame S'.

Note that, as we will show when introducing the total relativistic energy, the conservation law of the momentum 4-vector is equivalent to two conservation principles in non-relativistic mechanics, namely the conservation of mechanical energy and linear momentum. That is, since the time component of the 4-momentum (p^0) is related to the system's total energy and spatial components with the three-momentum.

VIII. CONCLUSIONS

This study presents a pedagogical framework that can be incorporated into the syllabi of physics courses for teaching Galilean and Lorentz transformations, as well as the special theory of relativity at various levels of high school and undergraduate university curricula. This study aims to facilitate a gradual transition from secondary school to careers in science, technology, engineering, and mathematics.

In particular, it aims to create an in-class material by balancing the mathematical skills necessary to gain an understanding of modern physics concepts, hence creating a teaching material that focuses on coaching a self-directed curiosity on real-world issues and, at the same time, encouraging social interaction in the process of learning [20, 21].

Acknowledgments

The author is grateful for the support of Eco/Logical Learning and Simulation Environments in Higher Education (ELSE), IT02-KA203-048006 (2018-2021). He is also thankful for supporting the Transversal Skills in Applied Artificial Intelligence (TSAAI) - Curriculum Development, ES01-KA220-HED-000030125 (2021-present) project.

[1] I. V. S. Mullis, M. O. Martin, P. Foy, and M. Hooper. *Advanced 2015 International Results in Advanced Mathematics and Physics*. Number 978-1-889938-31-8. TIMMS & PIRLS International Study Center, Boston College, Chestnut Hill, MA, (2016).

[2] L. Moran. *New Directions in the Teaching of the Physical Sciences* 6, 68–71 (2010).

[3] O. Goldstein. *Cogent Education* 3, 1200833 (2016).

[4] W. Muliawan, W. S. Nahar, C. E. Sebastian, E. Yuliza, and Khairurrijal. *Journal of Physics: Conference Series* 739, 012139 (2016).

[5] S. T. Kanyesigye, J. Uwamahoro, and I. Kemeza. *Physical Review Physics Education Research* 18, 010140 (2022).

[6] L. Moran and H. Vaughan. *New Directions in the Teaching of the Physical Sciences* 8, 17–21 (2016).

[7] H. Kamberaj. *Classical Mechanics*. De Gruyter, (2021).

[8] H. Kamberaj. *Electromagnetism. With Solved Problems*. Springer Nature, (2022).

[9] M. A. Jaswon. *Nature* 224, 1303–1304 (1969).

[10] U. J. Nöckel. *Eur. J. Phys.* 43(4), 045202 (2022).

[11] D. Holliday, R. Resnick, and J. Walker. *Fundamentals of Physics*. John Wiley & Sons, (2011).

[12] H. Sykja. *Bazat e Elektrodinamikës*. SHBUT, (2006).

[13] M. J. Klein, A. J. Kox, and R. Schulman, editors. *The Collected Papers of Albert Einstein: The Berlin Years: Writings, 1914-1917*, volume 6. Princeton University Press, (2016).

[14] R. Meidani. *Mbi Hapsirë-Kohën*. Toena, (2022).

[15] D. J. Griffiths. *Introduction to Electrodynamics*. Prentice Hall, third edition, (1999).

[16] H. Goldstein, C. Poole, and J. Safko. *Classical Mechanics*. Addison Wesley, (2002).

[17] R. P. Feynman. *Feynman Lectures on Gravity*. Addison Wesley, (1995).

[18] M. Nakahara. *Geometry, Topology, and Physics*. Institute of Physics Publishing, (1996).

[19] J. D. Jackson. *Classical Electrodynamics*. John Wiley & Sons, third edition, (1999).

[20] V. K. LaBoskey, J. Loughran, and M. L. Hamilton. *International Handbook of Self-study of Teaching and Teacher Education Practices* 2, 817–869 (2004).

[21] L. S. Vygotsky. *Mind in society: The development of higher psychological process*. Cambridge, MA, Harvard, (1978).