Invariance of Maxwell's Equations: A Study of Their Properties Under Galilean and Lorentz Transformations

Jiayuan Fan

Received June 17, 2024
Accepted September 20, 2024
Electronic access October 15, 2024

We analyze the relationship between Maxwell's equations and two types of transformations of frame of reference: the Galilean transformation and the Lorentz transformation. To enhance our understanding of Maxwell's equations, we first introduce the del operator and its operations: gradient, divergence, and curl. Then, we explain Maxwell's equations in differential form and illustrate the equations' framework by presenting the derivation of the motion of a particle under the effect of a constant magnetic field in a stationary frame of reference. To evince the behavior of Maxwell's equations in moving frames of reference, we examine the concept of Galilean transformation and present the derivation Maxwell's equations in the Galilean framework to demonstrate that they are not invariant under this framework. We then proceed to evaluate a general solution to the wave equation under Maxwell's framework and use this solution as another demonstration of the incompatibility between Maxwell's equations and the Galilean transformation. To resolve this, we present the derivation of Lorentz transformation, another type of transformation that preserves both the general solution of the wave equation and Maxwell's equations, from the general solution of wave equation. To confirm this property, we confirm the constancy of the speed of light under the Lorentz transformation. Next, we examine the properties of the spacetime interval, a crucial concept in special relativity and the cornerstone of the Lorentz transformation. Finally, we present the implications of the Lorentz transformation, including its relationship with the Galilean transformation, time dilation, and length contraction.

1 Introduction

Before James C. Maxwell, electricity and magnetism were regarded as two subjects and studied separately. Maxwell's equations, first introduced in 1865, succinctly summarized the behavior of charged particles in electric and magnetic fields, related electricity with magnetism, and created the field of electromagnetism[1]. The set of equations also predicted the existence of electromagnetic waves, which expanded a long-lasting idea that views light as a beam of particles propagating through "ether," a matter filling the vacuum[2]. This prediction of electromagnetic waves allowed the later derivation of solutions of the wave equation in a coordinate plane, describing the behavior of waves in various mediums.

Galilean relativity, developed by Italian scientist Galileo Galilei, is a principle that states the law of motion remains the same in all inertial frames. A consequence of this concept is the *Galilean transformation*, where a frame of reference transforms linearly without accelerating. This idea significantly contributed to modern physics as this notion allowed for a consistent description of motion. However, Maxwell's equations were considered incompatible with Galilean relativity. If a frame of reference remains at rest in a coordinate system, Maxwell's equations hold invariant. Under Galilean transformation, the equations would not remain invariant[3]. Thus, it is crucial to find a way such that the transformation of the frame of reference does not alter the result of Maxwell's equations.

Later, in 1905, Albert Einstein developed the special theory of relativity. Before this theory, physicists, like Galilei, generally viewed time and space as separate variables, and changing one does not affect the other. Special relativity fundamentally challenges this framework. It views the time not as an absolute quantity but as a fourth dimension in addition to the three space dimensions x, y, and z and one dimension of time[4]. This concept built a capstone in special relativity and led to the development of Lorentz transformation, a type of transformation that is able to make the form of Maxwell's equations invariant. This type of transformation mixes space and time and keeps the speed of light constant.

2 Materials and Methods

As this paper focuses on theoretical results in electromagnetic theory, it is essential to establish a solid foundation in the mathematical tools that underpin these concepts. This section presents a concise overview of del operators and Maxwell's equations to elucidate the underlying mathematical processes and their physical implications that form the foundation of our subsequent analysis.

2.1 Del Operators

In a three-dimensional Cartesian coordinate system, the symbol ∇, called "del operator," is defined as

$$\nabla = \frac{\partial}{\partial x}\hat{i} + \frac{\partial}{\partial y}\hat{j} + \frac{\partial}{\partial z}\hat{k} \quad (1)$$

This notation signifies that the del operator is a vector in three dimensions. When applied to a function, each component of the del operator is a partial derivative of the function with respect to the variable. Here we write the partial derivative of a function instead of the total derivative $\mathrm{d}f$ because the function $f(x)$ is defined as a function of three variables: x, y, and z. Thus, we must take the partial derivative of the function with respect to every variable.

2.1.1 Gradient

In one-dimensional spaces, the gradient is synonymous with the derivative at a single point, defined as follows with respect to a continuous and differentiable function $f(x)$:

$$\frac{\mathrm{d}}{\mathrm{d}x} f(x) = \lim_{h \to 0} \frac{f(x+h) - f(x)}{h} \quad (2)$$

In a three-dimensional Cartesian coordinate system, consider an arbitrary scalar field $f(x,y,z)$. The gradient of the function f is defined as

$$\nabla f = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z} \right), \quad (3)$$

which is a vector that points in the direction of the greatest rate of change of f, and its magnitude is the rate of change in that direction.

2.1.2 Divergence

Divergence measures the density of a vector field entering or leaving a point. For a vector field $\vec{G} = (G_x, G_y, G_z)$, its divergence is denoted as $\nabla \cdot \vec{G}$ and defined as:

$$\begin{aligned} \nabla \cdot \vec{G} &= \nabla \cdot (G_x, G_y, G_z) \\ &= \left(\frac{\partial G_x}{\partial x} + \frac{\partial G_y}{\partial y} + \frac{\partial G_z}{\partial z} \right). \end{aligned} \quad (4)$$

The result of $\nabla \cdot \vec{G}$ would be a scalar since the divergence does not measure the direction of the density entering or leaving a point but only the net change.

Divergence measures the net outward flux passing through a volume in a vector field. To demonstrate this, we let an arbitrary function $g(x)$ represent the incoming flux entering a volume and another function $g(x+h)$ represent the flux leaving the volume. If $g(x) < g(x+h)$, the divergence is positive, indicating that there is a source in the given volume that is producing outward flux. If $g(x) > g(x+h)$, the divergence is positive, indicating that the flux is decreasing in magnitude as it travels through the volume.

Divergence is highly relevant to the concept of the electric field. A positively charged particle will induce an electric field that is radially outward, or pointing away from the particle, and a negatively charged particle will induce an electric field that is radially inward or pointing towards the particle. At a point, positive divergence of the electric field indicates the presence of a positive charge, and negative divergence indicates the presence of a negative charge.

2.1.3 Curl

When ∇ is applied to a vector field $\vec{H} = (H_x, H_y, H_z)$ using a cross product, the result yielded is the curl of $\vec{H}$, or $\nabla \times \vec{H}$, that measures the rotation of $\vec{H}$ at a point, including the magnitude and direction of the rotation. The curl takes the form of

$$\begin{aligned} \nabla \times \vec{H} &= \nabla \times (H_x, H_y, H_z) \\ &= \left(\frac{\partial H_z}{\partial y} - \frac{\partial H_y}{\partial z}, \frac{\partial H_x}{\partial z} - \frac{\partial H_z}{\partial x}, \frac{\partial H_y}{\partial x} - \frac{\partial H_x}{\partial y} \right), \end{aligned} \quad (5)$$

describing both the direction in which the field points and the strength of the field at any given point. A positive curl corresponds to a counterclockwise rotation of the field lines, while a negative curl corresponds to a clockwise rotation, and the magnitude of the curl indicates the strength of this rotation. Physically, a non-zero curl of a magnetic field indicates the presence of a current or a changing electric field, and a non-zero curl of an electric field indicates the presence of a changing magnetic field.

2.2 Maxwell's equations

In free space, we have a set of four Maxwell's equations in differential forms that tell us the properties of and the relationship between electric and magnetic fields[1,5].

$$\nabla \cdot \vec{E} = \frac{\rho}{\varepsilon_0} \qquad \text{(Gauss's Law for Electricity)}$$

In Gauss's Law for Electricity, $\vec{E}$ denotes the electric field across the plane, ρ denotes the volume charge density in this region, and ε_0 is the permittivity of the free space which has a value of $8.854 \times 10^{-12} F \cdot m$. Mathematically, it represents that the electric flux Φ passing through the surface of an enclosed volume is equal to the total charge in that volume divided by the permittivity of the free space. We could also write the equation as $\Phi = \oint \vec{E} \mathrm{d}A = \frac{Q_{\mathrm{enc}}}{\varepsilon_0}$, where A denotes the surface area of the Gaussian surface, a closed, hypothetical surface in three-dimensional space, and Q_{enc} denotes the charge enclosed in the Gaussian surface[5,6].

$$\nabla \cdot \vec{B} = 0 \qquad \text{(Gauss's Law for Magnetism)}$$

In Gauss's Law for Magnetism, $\vec{B}$ denotes the magnetic field across the plane. The equation mathematically states that the divergence of the magnetic field is always zero. Physically, this means that magnetic field lines always form closed loops and do not originate or terminate at any point. This implies that magnetic monopoles do not exist in nature. The magnetic field line will leave the magnet from the north pole and re-enter the magnet from its south pole; therefore, if we create an arbitrary Gaussian surface that includes the entirety of the magnet, the change of the net flux Φ_B is 0[1].

An important application of Gauss's Law for Magnetism is the Helmholtz coil, which consists of two identical circular coils placed symmetrically along a common axis. The coils are separated by a distance equal to their radius. Helmholtz coils are specifically designed to create a uniform magnetic field in the region between the two coils, which functions akin to the north and south pole in a magnetic field: the magnetic field departs from one coil and re-enters the other coil from the opposite direction[7]. When an electric current flows through the coils in the same direction, the magnetic fields generated by each coil add together in the central region between the coils, which creates a uniform magnetic field along the axis of the coils. In this region, $\nabla \cdot \vec{B} = 0$ is satisfied because magnetic field is constant and has no divergence[8].

$$\nabla \times \vec{E} = -\frac{\partial \vec{B}}{\partial t} \qquad \text{(Faraday's Law of Induction)}$$

Faraday's Law of Induction relates magnetic field $\vec{B}$ with electric field $\vec{E}$. This equation indicates that the change of magnetic field will induce an electric field, and a rotating electric field always accompanies a time-varying magnetic field[1].

$$\nabla \times \vec{B} = \mu_0 \vec{J} + \mu_0 \varepsilon_0 \frac{\partial \vec{E}}{\partial t} \qquad \text{(Ampere's Law)}$$

In Ampère's Law with Maxwell's Addition, $\vec{J}$ denotes the current density at a point, and μ_0 denotes the permeability of free space with a value of $4\pi \times 10^{-7}\,\frac{kg\cdot m}{s^2\cdot A^2}$. Mathematically, the equation states that the magnetic field at a point is caused by a current $\mu_0\vec{J}$ and another changing current $\mu_0\varepsilon_0\frac{\partial \vec{E}}{\partial t}$, called "displace current," that is passing through the loop where the magnetic field is induced[9].

The four laws above together constitute Maxwell's equations in free space. These equations explain the nature of electromagnetism, governing electric and magnetic forces and the propagation of electromagnetic waves.

3 Motion of a Particle Under Electric and Magnetic Field

To better understand the properties of Maxwell's equations, we model the motion of a particle under their influence. Consider a particle q with an arbitrary positive charge traveling in a three-dimensional space that is parameterized by a Cartesian coordinate system with a uniform, time-independent magnetic field $\vec{B} = B_z\hat{k}$. We define the position of q with respect to time t as (x, y).

The Lorentz force equation in the absence of an electric field is given by $\vec{F} = q\vec{v} \times \vec{B}$. The cross product $\vec{v} \times \vec{B}$ yields a force vector in the plane perpendicular to the magnetic field, expressed as

$$\vec{F} = q(v_x\hat{\imath} + v_y\hat{\jmath}) \times (B_z\hat{k}) = qB_z(-v_x\hat{\jmath} + v_y\hat{\imath}), \tag{6}$$

where $\hat{\imath}, \hat{\jmath}$ and $\hat{k}$ are unit vectors in x, y, and z direction, respectively. We could rewrite the equation above to describe the relationship between x and y by applying Newton's second law, $\vec{F} = m\vec{a}$. For the component along the x-axis, the equation takes the form $ma_x = qv_yB_z$. This relationship can be written as a differential equation of x by recognizing that acceleration, a_x, is the second derivative of position x with respect to time and the velocity component v_y along the y-axis is the derivative of position x with respect to time. The equation takes the form of

$$m\frac{\mathrm{d}^2 x}{\mathrm{d}t^2} = q\frac{\mathrm{d}y}{\mathrm{d}t}B_z. \tag{7}$$

Conversely, for the y-component, by considering acceleration along the y-axis a_y as the second derivative of y with respect to time and the velocity component v_x along the x-axis as the derivative of position x with respect to time, we obtain the differential equation of y taking the form of

$$m\frac{\mathrm{d}^2 y}{\mathrm{d}t^2} = -q\frac{\mathrm{d}x}{\mathrm{d}t}B_z. \tag{8}$$

For the convenience of solving these equations, we could take the derivative of both equations with respect to time again. Now, we have a set of equations that describes the relationship between x and y:

$$m\frac{\mathrm{d}^2 x}{\mathrm{d}t^2} = q\frac{\mathrm{d}y}{\mathrm{d}t}B_z, \quad m\frac{\mathrm{d}^2 y}{\mathrm{d}t^2} = -q\frac{\mathrm{d}x}{\mathrm{d}t}B_z, \tag{9a}$$

$$m\frac{\mathrm{d}^3 x}{\mathrm{d}t^3} = q\frac{\mathrm{d}^2 y}{\mathrm{d}t^2}B_z, \quad m\frac{\mathrm{d}^3 y}{\mathrm{d}t^3} = -q\frac{\mathrm{d}^2 x}{\mathrm{d}t^2}B_z. \tag{9b}$$

We begin by solving the x-component of the position of the particle. Taking the derivative of both sides with respect to time again and with some algebraic manipulation, we obtain $-\left(\frac{B_z q}{m}\right)^2 x - \frac{\mathrm{d}^2 x}{\mathrm{d}t^2} = 0$, which has a solution taking the form of

$$x = C\cos\left(\frac{Bq}{m}t\right) + iD\sin\left(\frac{Bq}{m}t\right),$$

where C and D are constants that are dependent on the initial condition of the particle q. With a similar approach, we find the y-component of particle q's position to be

$$y = M\cos\left(\frac{Bq}{m}t\right) + iN\sin\left(\frac{Bq}{m}t\right),$$

where M and N are arbitrary constants whose values are dependent on the initial condition of the particle q. Substituting the coordinates into equations (7) and (8). The expressions then take the form of $C = -iN$ and $D = -iM$. With this relationship, the coordinate of the particle is:

$$(x,y) = \begin{pmatrix} M\sin\left(\frac{Bq}{m}t\right) - iN\cos\left(\frac{Bq}{m}t\right), \\ M\cos\left(\frac{Bq}{m}t\right) + iN\sin\left(\frac{Bq}{m}t\right) \end{pmatrix}$$

The equation describes the motion of the particle q in a complex plane. If we only look at the real components of the equation, it has the form $\mathrm{Re}(x,y) = \left(M\sin\left(\frac{Bq}{m}t\right), M\cos\left(\frac{Bq}{m}t\right)\right)$, which suggests that the particle in a constant magnetic field would undergo a circular motion with radius M and angular frequency $\frac{Bq}{m}$.

However, this equation for motion is only valid in a stationary frame of reference. It is unclear how this result will change when viewed from a frame in relativistic motion relative to another. To generalize this relationship to transformed frames, we must first understand the properties of these frame transformations.

4 Relativity and Frames of Reference

A frame of reference is defined as a system of coordinates that measures the position of an object and its trajectory with respect to space and time. There are two types of frames of reference: inertial and non-inertial. Specifically, an inertial frame of reference is defined as a frame of reference that does not undergo acceleration or rotation. In such a frame of reference, Newton's first law strictly applies: unless an external force is applied to an object, it will remain in uniform motion or a motionless state[10].

The *Galilean transformation* describes a special type of transformation of an inertial frame of reference, where the two frames are moving at a constant velocity relative to each other. For the transformation of frames of reference to be Galilean, the following assumptions must be true:

1. If we apply a force F, the force measured in one frame has the same magnitude and direction as the force measured in the other frame.
2. The measurement of mass is the same in both frames of reference.
3. The measurement of time is the same in both frames of reference[11].

We could then set up a Galilean transformation in a three-dimensional coordinate system (x,y,z). First, we set up an arbitrary two-dimensional Cartesian coordinate system, and define a frame of reference in the coordinate system as R and another frame of reference as R', such that the R' is a Galilean-transformed product of R. The origin of R is O and the origin of R' is O'. We apply an arbitrary force in the coordinate system. In the R frame, the observers measure a force F, while in the R' frame, the observers measure a force F'. Force F and F', as measured, must have the same magnitude. We define a time period t in R and another time period t' in R', which is homogeneous. We place an arbitrary mass m in R and another arbitrary mass m' in R'. Express the assumption in mathematical expressions, we have a set of equations such that

$$F = F', \quad m = m', \quad t = t'.$$

Under Newton's second law, which states $F = ma$, we could rewrite the equations as functions of acceleration:

$$a = \frac{F}{m}, \quad a' = \frac{F'}{m'}.$$

Since $F = F'$ and $m = m'$, $a = a'$ must be true. Given that the measurement of time remains invariant under both frames of reference, $\frac{\mathrm{d}^2x}{\mathrm{d}t^2} = \frac{\mathrm{d}^2x'}{\mathrm{d}t'^2}$ must also be true. Thus, the relationship between x and x' must be such that it yields a constant acceleration. Thus, we define the relationship between x and x' to be

$$x' = x + vt + x_0, \tag{10}$$

where x' is the position of the R' frame in the coordinate system, x is the position of R frame, v is the velocity of R' traveling with respect to R frame, t is the time span of both frames of reference, and x_0 is the initial position of R' frame in the coordinate system. The set of equations that describes a Galilean transformation could be summarized as:

$$x' = x - \vec{v}t + x_0 \tag{11a}$$

$$y' = y \tag{11b}$$

$$z' = z \tag{11c}$$

$$t' = t. \tag{11d}$$

4.1 Incompatibility Between Maxwell's Equations and Galilean Relativity

With this property of Galilean relativity in hand, we could use Ampère's Law to demonstrate that Maxwell's equations are not invariant under Galilean transformation. Consider Ampère's Law in free space:

$$\nabla \times \vec{B} = \mu_0\vec{J} + \mu_0\varepsilon_0\frac{\partial\vec{E}}{\partial t}.$$

There are three variables in this equation: the magnetic field $\vec{B}$, the electric field $\vec{E}$, and the current density $\vec{J}$. To determine whether Ampère's Law remains invariant after the transformation, we need to understand the behavior of the variables. We define an arbitrary electric field E and an arbitrary magnetic field B in R frame. After transformation, the electric field becomes E', and the magnetic field becomes B'. Consider a particle with charge q traveling with a speed $\vec{u}$ in the R frame. In the perspective of the R' frame, the particle is traveling with a speed $\vec{u}-\vec{v}$. We have the Lorentz force equation

$$F = q(\vec{E}+\vec{u}\times\vec{B}) \tag{12}$$

According to the assumption in Galilean transformation, $F = F'$. If $u = 0$, the equation becomes

$$\vec{E} = \vec{E'} - \vec{v}\times\vec{B'}, \tag{13}$$

We substitute equation (13) back into equation (11) to solve for an expression of u, which returns $\vec{u}\times\vec{B} = \vec{u}\times\vec{B'}$. Since $\vec{u}\times\vec{u} = 0$, the variables B and B' must satisfy a relationship such that $\vec{B'} = \vec{B}+k\vec{u}$, where k is an arbitrary constant. Now, we substitute this expression into equation (13). Since $\vec{u}\times k\vec{u} = 0$, the expression simplifies to

$$\vec{E'} = \vec{E}+\vec{v}\times\vec{B}. \tag{14}$$

As per the definition of current density, $\vec{J} = \rho\vec{u}$. Thus, after transformation, the current density $\vec{J'} = \rho(\vec{u}-\vec{v}) = \vec{J}-\rho\vec{v}$. Substituting them into Ampère's law after transformation, we get

$$\begin{aligned}\nabla\times\vec{B} = \mu_0(\vec{J}-\rho\vec{v}) - k(\nabla\times\vec{u}) + \mu_0\varepsilon_0\frac{\partial\vec{E}}{\partial t} \\ +\mu_0\varepsilon_0\left(\frac{\partial\vec{v}}{\partial t}\times\vec{B}+\vec{v}\times\frac{\partial\vec{B}}{\partial t}\right)\end{aligned} \tag{15}$$

Comparing with the original Ampère's law equation, we notice that there are additional terms $-\mu_0\rho\vec{v}$ and $k(\nabla\times\vec{u})$ in the transformed Ampère's law that is not present in the original expression. Additionally, the transformed equation introduces a dependency on the relative velocity of the frames $\vec{v}$, which is a relationship not present in the original expression. These differences conclude that Ampère's law is not invariant under Galilean transformation.

5 Plane Wave Equation

In section 4, we solved the position of a particle as a function of time. Since the electric and magnetic fields drive the motion of the particle, it is crucial to understand these fields in detail. A broader and more general implication of Maxwell's equations regards fields, either electric or magnetic, as a function of space and time. In other words, the change in the spatial quantity and time of a field may lead to a change in the field itself.

To further our understanding, we need to study electromagnetic waves. These are waves of electric and magnetic fields that are able to propagate through free space at the speed of light. Notably, when a charged particle, such as an electron, accelerates, it generates an electromagnetic wave. The acceleration causes a changing electric field, which then induces a varying magnetic field. The electric field $\vec{E}$ and magnetic field $\vec{B}$ oscillate perpendicular to each other and to the direction of wave propagation, which forms a transverse wave called electromagnetic waves.

We can derive a general solution of Maxwell's equations in waveform. In a three-dimensional free space (x,y,z), we consider an electric field defined as $\vec{E} = E_x\hat{\imath}$ and a magnetic field defined as $\vec{B} = B_y\hat{\jmath}$. This is a special case; in general, both fields have components in all three directions: $\hat{\imath}$, $\hat{\jmath}$, and $\hat{k}$. Maxwell's equations relate the electric field and the magnetic field together:

$$\begin{aligned}\nabla\times(E_x\hat{\imath}) &= \left(0\hat{\imath}, \frac{\partial E_x}{\partial z}\hat{\jmath}, -\frac{\partial E_x}{\partial y}\hat{k}\right) \\ &\Rightarrow -\frac{\partial B_y}{\partial z}\hat{\imath} = \mu_0\varepsilon_0\frac{\partial E_x}{\partial t}\hat{\imath},\end{aligned} \tag{16a}$$

$$\begin{aligned}\nabla\times(B_y\hat{\jmath}) &= \left(-\frac{\partial B_y}{\partial z}\hat{\imath}, 0\hat{\jmath}, \frac{\partial B_y}{\partial x}\hat{k}\right) \\ &\Rightarrow \frac{\partial B_y}{\partial t}\hat{\jmath} = \frac{\partial E_x}{\partial z}\hat{\jmath}.\end{aligned} \tag{16b}$$

Taking the partial derivatives of equations (16a) with respect to z and (16b) with respect to t and combining similar terms, we obtain a general relationship of the magnetic field $\vec{B}$ with respect to space and time:

$$\frac{\partial^2\vec{B}}{\partial z^2} - \mu_0\varepsilon_0\frac{\partial^2\vec{B}}{\partial t^2} = 0. \tag{17}$$

This equation shows the wave equation for the magnetic field and relates the second partial derivative of B with respect to z and the second partial derivative of B with respect to t, which indicates that the magnetic field is a function of both quantities. To gain a deeper understanding of its behavior, we examine the general solution of the wave equation:

$$\frac{\partial^2 f}{\partial z^2} - \frac{1}{c^2}\frac{\partial^2 f}{\partial t^2} = 0, \tag{18}$$

where f can represent any wave-like quantity. In this case, we see that f is equivalent to $\vec{B}$. It is clear that $\frac{1}{c^2} = \mu_0\varepsilon_0$, which implies $c = \frac{1}{\sqrt{\mu_0\varepsilon_0}}$. As per definition, c is the expression of the speed of light with a value of 3×10^8 m/s.

A solution to equation (17) is:

$$B(z,t) = M\cos[k(z\pm vt)] + N\sin[k(z\pm vt)]. \tag{19}$$

We define a new variable ω such that $\omega = kv$, and expressions $A\cos(\delta) = M$ and $A\sin(\delta) = N$. Upon simplification, the equation becomes

$$B(z,t) = A\cos[(kz - \omega t) - \delta], \tag{20}$$

where A is the amplitude of the wave, ω is the angular frequency of the wave equation, measuring how many radians the phase of the wave changes per second, k is the spatial frequency of the wave, measuring the number of radians the wave phase changes per unit of space, and δ is a phase shift, which adjusts where the wave starts in its cycle at $t = 0$ and $z = 0$. If the time t is set to be an arbitrary instant at any moment, the magnitude of $\vec{B}$ oscillates as a function of z. If the spatial quantity z is set to an arbitrary value, the value of $\vec{B}$ oscillates as a function of t.

We can derive the relationship between z, k, ω, and t by solving a particular solution of the plane wave equation. Setting $A = 1$ and $\delta = 0$, we define a point C_1 on the wave function that maintains a constant phase of 2π. As the wave oscillates, C_1 undergoes linear translational motion with a velocity of ω along $y = 2\pi$. At any point during the oscillation, the function is given by $B(z,t) = \cos(kz - \omega t) = 1$. Since $\cos(2\pi a) = 1$ for $a \in \{x \in \mathbb{Z} \mid x \geq 0\}$, it follows that $kz - \omega t = 2\pi a$. Thus, the relationship $z = \frac{2\pi a}{k} - \frac{\omega}{k}t$ indicates that any point on the constant phase moves at a speed of $\frac{\omega}{k}$.

5.1 General Solution of the Wave Equation

To derive the general solution of the wave equation, we can start by defining an arbitrary function $f_{w_1}(z,t) = A_1\cos(kz - \omega t)$. Now, consider the function $f(u)$, where $u = z - ct$. By rearranging this relationship, we get $z = u + ct$, which shows that the variable u represents a wave traveling in the positive z-direction with a constant speed c. Assuming that the wave number k and angular frequency ω are related by $k = \frac{u}{\omega}$, the wave function can be written in a more general form:

$$f_{w_1}(u) = A_1\cos(\omega_1 u). \tag{21}$$

If we define another function $f_{w_2}(u) = A_2\cos(\omega_2 u)$ in the plane, it is another solution to the wave equation with a different frequency ω_2 and amplitude A_2. Thus, according to the superposition principle, the linear combination $c_1 f_{w_1}(u) + c_2 f_{w_2}(u)$ is still a solution. To generalize the result, we could write the solution of the wave equation as a sum of u:

$$S(u) = \sum_{\omega=-\infty}^{\infty} A(\omega)\cos(\omega u). \tag{22}$$

Next, we could use the Fourier transform, an operation that transforms functions about space and time into functions of frequency. The general form of the Fourier transform is given by

$$\widehat{g}(\xi) = \int_{-\infty}^{\infty} G(u)\, e^{-i2\pi\xi x}\,\mathrm{d}u, \tag{23}$$

where G is the Fourier-transformed product, and ξ is the linear frequency of the wave. Notice that in our solutions to the wave equation, we have the term ω, the angular frequency, and we can transform it to linear frequency using the expression $\xi = 2\pi\omega$. The Fourier transform now has the expression

$$g(u) = \int_{-\infty}^{\infty} G(\omega)e^{-i2\omega u}\,\mathrm{d}u, \tag{24}$$

where $G(\omega)$ is the sum of all amplitudes of the wave equation and the quantity $e^{-i2\pi\omega u}$ represents a rotating vector in the complex plane. With Euler's formula $e^{i\theta} = \cos(\theta) + i\sin(\theta)$, we can expression the Fourier-transformed products as:

$$g(u) = \int_{-\infty}^{\infty} G(\omega)\,[\cos(2\omega u) - i\sin(2\omega u)]\,\mathrm{d}u. \tag{25}$$

We are interested in the behavior of $g(u)$ in the real number set. Looking at the components, it is clear that $G(\omega)$ is equivalent to A, the amplitude. Putting the complex terms aside, the equation is mathematically equivalent to equation (20) in domain $(-\infty, \infty)$. From this property of the wave equation, we can conclude that the general solution of the wave equation, equation (20), is an infinite sum of all the wave equations with different frequencies and amplitudes in the plane.

To better understand the behavior of the plane wave equation, we introduce another variable v, defined as $v = z + ct$, alongside u. While u represents the wave equation propagating through positive z direction, v represents the wave equation propagating through negative z direction. We define the wave function $f(u,v)$. Taking the partial derivatives of f with respect to z, we have

$$\frac{\partial f}{\partial z} = \frac{\partial f}{\partial u}\frac{\partial u}{\partial z} + \frac{\partial f}{\partial v}\frac{\partial v}{\partial z},$$

which equals to $\frac{\partial f}{\partial u} + \frac{\partial f}{\partial v}$ since u and v are both linear equations of z. Similarly, we could take the partial derivatives of f with respect to t and obtain

$$\frac{\partial f}{\partial t} = c\frac{\partial f}{\partial u} - c\frac{\partial f}{\partial v}. \tag{26}$$

After solving the partial derivatives, we could further derive equation (18):

$$\frac{\partial^2 f}{\partial z^2} - \frac{1}{c^2}\frac{\partial^2 f}{\partial t^2} = 0 \quad \Rightarrow \quad \frac{\partial^2 f}{\partial u \partial v} = 0. \tag{27}$$

This expression conveys the idea that there is no interaction between the waves moving in the $-z$ and $+z$ direction, but both waves are one of the infinitely many solutions to the plane wave equations. Combining this result with the conclusion driven by the Fourier transform, we obtain an important property of the wave equation: the general solution of the wave equation is a superposition of infinitely many independent traveling waves[12].

5.2 Wave Equations in Different Reference Frames

We define two reference frames, R and R', where the R frame is an arbitrary inertial frame and the R' frame is a Galilean-transformed product of R. The two frames of reference have a relationship taking the form of

$$\begin{aligned} F &= F' \\ z' &= z + vt + z_0 \\ t' &= t \\ y' &= y. \end{aligned}$$

In examining the differential operator $\frac{\partial}{\partial t'}$ under Galilean transformation, where the transformation is given by $z' = z + vt + z_0$, we derive the following general expression:

$$\frac{\partial}{\partial t'} = \frac{\partial t}{\partial t'}\frac{\partial}{\partial t} + \frac{\partial z}{\partial t'}\frac{\partial}{\partial z} + \frac{\partial y}{\partial t'}\frac{\partial}{\partial y}. \tag{28}$$

This equation can be simplified by considering the specific relationships between the variables. Given that the time transformation under Galilean relativity asserts $t' = t$, it follows that $\frac{\partial t}{\partial t'} = 1$. Furthermore, since the transformation happens only on the z coordinate, it does not affect the y coordinate ($y' = y$). Therefore, it is logical to conclude that the value of y does not depend on the change of t', so we have $\frac{\partial y}{\partial t'} = 0$. The differential equation becomes:

$$\frac{\partial}{\partial t'} = \frac{\partial}{\partial t} + v\frac{\partial}{\partial z}. \tag{29}$$

Next, we could derive the equation of $\frac{\partial}{\partial z'}$ under the Galilean transformation of $z' = z + vt + z_0$. As per the definition,

$$\frac{\partial}{\partial z'} = \frac{\partial z}{\partial z'}\frac{\partial}{\partial z} + \frac{\partial t}{\partial z'}\frac{\partial}{\partial t} + \frac{\partial y}{\partial z'}\frac{\partial}{\partial y} \tag{30}$$

Same as before, we could simplify this equation by using relationships between the variables. The variable y is independent of z; thus, $\frac{\partial y}{\partial z'}\frac{\partial}{\partial y} = 0$. The variable t is also independent of z; thus, the term $\frac{\partial t}{\partial z'}\frac{\partial}{\partial t} = 0$. We could rewrite $z' = z + vt + z_0$ as $z = z' - vt - z_0$, where v, t, and z_0 are all constants, which means $\frac{\partial}{\partial z'}(z' - vt - z_0) = 1$. Substitute this equation into $\frac{\partial z}{\partial z'}$, the equation becomes

$$\frac{\partial}{\partial z'} = \frac{\partial}{\partial z}. \tag{31}$$

Now, we define a wave equation in the R' frame that has the form

$$\frac{\partial^2}{\partial z'^2} - \frac{1}{c}\frac{\partial^2}{\partial t'^2} = 0. \tag{32}$$

Substituting equation (29) and (31) into the equation above, we obtain the expression

$$\frac{\partial}{\partial z} - \frac{1}{c}\left(\frac{\partial^2}{\partial t^2} + 2v\frac{\partial^2}{\partial t \partial z} + v^2\frac{\partial^2}{\partial z^2}\right) = 0.$$

It is clear that, under Galilean transformation, the wave equation is not invariant because the equation has a partial derivative changing with respect to t and z, which is a relationship that is not shown in the original expression.

6 Special Relativity and Lorentz Transformation

6.1 Derivation

Recognizing the inability of the Galilean transformation to make the solution of the wave equations invariant, it is crucial for us to find a different way of transforming the frames of reference such that under the transformation, the wave equation would remain invariant[13]. To understand such a transformation, we present the idea of special relativity, a scientific theory about the relationship between space and time proposed by Albert Einstein. There are two postulates in special relativity:

1. The law of physics is invariant and identical in all inertial frames of reference[14].
2. The speed of light c in a vacuum is constant for all observers regardless of their motion with respect to the light source.

The transformation we intend to find is built upon these two postulates. We set up two frames of reference S and S', where S' is a transformed product of S. In S frame, we define an arbitrary function $f(z,t)$ on the z-direction and is changing as a function of time and another function $f'(z',t')$ in S' frame. In the S frame, we have the wave equation $\frac{\partial^2 f}{\partial z^2} - \frac{1}{c^2}\frac{\partial^2 f}{\partial t^2} = 0$, and an ideal transformation should ensure that the wave equation in the S' frame has a solution of the form $\frac{\partial^2 f'}{\partial z'^2} - \frac{1}{c^2}\frac{\partial^2 f'}{\partial t'^2} = 0$. To achieve this, we propose a set of expressions that relates z and t in f with z' and t' in f' such that

$$z' = az + bt, \tag{33a}$$
$$t' = mz + nt. \tag{33b}$$

The relationship of z and t with respect to z' and t' must be linear because this set of equations must satisfy the first postulate of special relativity, which explicitly requires the frames to be inertial and the law of physics to be the same everywhere in the frames. If an object in one frame of reference does not experience acceleration, it must not experience acceleration in the transformed frame of reference. Thus, the relationship between the variables can only be linear.

We take the partial derivative of f' with respect to z' and t'. Substituting them into the wave equation of f', which takes the form of $\frac{\partial^2 f'}{\partial z'^2} - \frac{1}{c^2}\frac{\partial^2 f'}{\partial t'^2} = 0$, we obtain a relationship between all the variables such that $a^2 - \frac{b^2}{c^2} = 1, \quad 2bn - c^2am = 0, \quad n^2 - c^2m^2 = 1$.

Recalling the hyperbolic trigonometry property that $\cosh^2(\phi) - \sinh^2(\phi) = 1$, it is clear that $a = n = \cosh(\phi)$, $b = c\sinh(\phi)$, and $m = \frac{\sinh(\phi)}{c}$. Substituting this property into the expression of z' and t', we obtain

$$z = \cosh(\phi)z' + c\sinh(\phi)t', \tag{34a}$$

$$t = \frac{\sinh(\phi)}{c}z' + \cosh(\phi)t'. \tag{34b}$$

Consider a particle at rest at $z' = 0$ in the S' frame and moving at a velocity v in the S frame. We can express the velocity of the particle as $v = \frac{z}{t}$. Since $z' = 0$, we can express z' and t' as $z = c\sinh(\phi)t'$, $t = \cosh(\phi)t'$. Substituting these expressions into the particle's velocity v, we obtain the expression taking the form of $\tanh(\phi) = \frac{v}{c}$. Using relevant hyperbolic trigonometry identities, we obtain the expressions

$$\cosh(\phi) = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}}, \quad \sinh(\phi) = \frac{\frac{v}{v_0}}{\sqrt{1 - \frac{v^2}{c^2}}},$$

where $\frac{1}{\sqrt{1-\frac{v^2}{c^2}}}$ is the Lorentz factor γ. Substituting these expressions back into the set of equations (34), we obtain the standard form of the transformation that would preserve the wave equation, called *Lorentz transformation*:

$$z' = \gamma(z - vt), \tag{35a}$$

$$t' = \gamma\left(t - \frac{vz}{c^2}\right). \tag{35b}$$

6.2 Proof of Constant Speed of Light

The second postulate of special relativity states that the speed of light must be constant in all inertial frames regardless of their state of motion[1]. Since Lorentz transformation does not involve accelerating frames, the speed of light should also be constant and have the same value in the initial frame and the transformed frame. We now show that this postulate is satisfied by the Lorentz transformation.

Proof. Consider a light wave propagating in the z-direction. In the original frame, the equation for the propagation of light is given by

$$z = ct.$$

We will apply the Lorentz transformation to the coordinates of the light:

$$t' = \gamma\left(t - \frac{vz}{c^2}\right),$$
$$z' = \gamma(z - vt).$$

Substitute $z = ct$ into the transformation equations, we obtain:

$$t' = \gamma\left(t - \frac{v(ct)}{c^2}\right),$$
$$z' = \gamma(ct - vt).$$

If the speed of light c is indeed preserved, the expressions $z' = ct'$ should be true. Substituting the expressions from above, we are seeking to show that whether $\gamma\left(t - \frac{v(ct)}{c^2}\right) = c\gamma(ct - vt)$. Simplifying the expression, it can be confirmed that both sides of the equation are equal to each other. Therefore, the speed of light is preserved under the Lorentz transformation. $\square$

6.3 Spacetime Interval

To enhance our understanding of Lorentz transformations, a concept to examine is the *spacetime interval*, denoted as s, which indicates the way two events are related in both space and time[15]. The spacetime interval between two discrete events with coordinates (t_1, x_1, y_1, z_1) and (t_2, x_2, y_2, z_2) is

$$\begin{aligned} s^2 &= c^2(t_2 - t_1)^2 \\ &-[(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2]. \end{aligned} \tag{36}$$

Just like a rotation in space preserves the Euclidean distance $r = \sqrt{(\Delta x)^2 + (\Delta y)^2 + (\Delta z)^2}$, a hyperbolic rotation preserves the spacetime interval[16]. It is essential to note that we use the term ct rather than t alone in the spacetime interval formula. The reason for this is to ensure dimensional consistency across the terms of the equation. By multiplying the time variable t by the speed of light c, we convert time into units of distance, which allows for direct computation with spatial measurements.

Now, we use the rotation of matrices to demonstrate the invariance of the spacetime interval under Lorentz transformation. In a two-dimensional Cartesian plane, we have a vector $\vec{r} = \begin{bmatrix} x \\ y \end{bmatrix}$. If we want this vector to rotate around the origin by an angle θ, a rotation matrix $B(\theta) = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$ could be applied on $\vec{r}$:

$$\vec{r'} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x\cos\theta - y\sin\theta \\ x\sin\theta + y\cos\theta \end{bmatrix}. \tag{37}$$

Similarly, we can construct a spacetime plane and define a vector $\vec{v}$ of time and space. The vector that defines the property of an object in the plane is $\vec{v} = \begin{bmatrix} ct \\ x \end{bmatrix}$, while y and z remain unchanged. To transform the spacetime plane, We could apply a hyperbolic rotation matrix $B(\phi)$ to the vector $\vec{v}$:

$$\begin{aligned} \vec{v'} = B_x(\phi)\vec{v} &= \begin{bmatrix} \cosh(\phi) & -\sinh(\phi) \\ -\sinh(\phi) & \cosh(\phi) \end{bmatrix} \begin{bmatrix} ct \\ x \end{bmatrix} \\ &= \begin{bmatrix} ct\cosh(\phi) - x\sinh(\phi) \\ -ct\sinh(\phi) + x\cosh(\phi) \end{bmatrix}, \end{aligned} \tag{38}$$

which changes the location and the time of the vector. The term ϕ is the *rapidity* of the Lorentz transformation, a parameter used in special relativity to describe the relative velocity between two inertial frames[17]. Denoting $ct\cosh(\phi) - x\sinh(\phi)$ as t' and $-ct\sinh(\phi) + x\cosh(\phi)$ as x', we can prove a fundamental property of the Lorentz Boost that for all values of ϕ, $t^2 - x^2 = t'^2 - x'^2$:

$$\begin{aligned}(ct')^2 - x'^2 &= [ct\cosh(\phi) - x\sinh(\phi)]^2 \\ &\quad - [-ct\sinh(\phi) + x\cosh(\phi)]^2 \\ &= ct^2\left[\cosh^2(\phi) - \sinh^2(\phi)\right] \\ &\quad + x^2\left[\sinh^2(\phi) - \cosh^2(\phi)\right].\end{aligned} \tag{39}$$

Since $\cosh^2(\phi) - \sinh^2(\phi) = 1$, the expression can be simplified to $(ct)^2 - x^2$, which is equal to $(ct')^2 - x'^2$. Given that the Lorentz transformation occurs only in the x direction, the spacetime interval for the object $\vec{v}$ is $s^2 = (ct)^2 - x^2$, and the spacetime interval for the transformed object $\vec{v'}$ is $s'^2 = (ct')^2 - x'^2$. Therefore, $s^2 = s'^2$. This expression reflects the invariance of the spacetime interval under Lorentz transformation, as discussed in equation (36).

6.4 Implications of Lorentz Transformation

6.4.1 Relationship Between Lorentz Transformation and Galilean Transformation

Having established the spacetime interval as a cornerstone of Lorentz transformations, we now explore some key implications. First, we examine the relationship between Lorentz and Galilean transformations. Given the Lorentz transformation of space,

$$x' = \gamma(x - vt) = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}}(x - vt), \tag{40}$$

it is evident that when $v \ll c$, $\gamma \approx 1$, resulting in $x' \approx x - vt$. This is the standard form of Galilean transformation. Thus, we can conclude that Galilean transformation is a special type of Lorentz transformation at low speed.

6.4.2 Time Dilation

The constancy of the speed of light under Lorentz transformation brings some interesting phenomena, one of which is *time dilation*. When two observers in different frames have a relative motion with respect to each other, they will measure the other observer's time as passing slower than their own. We could consider an intuitive example of time dilation. Two mirrors A and B, as in Figure 1, are placed at rest in the coordinate system with a distance L apart. At the center of A, there is a laser beam shooting towards the mirror B. The laser beam is traveling at the speed of light; thus, if it takes the laser a time Δt to reach mirror B, the separation distance $L = c\Delta t$. Define an observer at an arbitrary point in the plane. If the mirror A starts to move to the right relative to the mirror B with a constant velocity v and the mirror B stays at rest relative to the observer, the laser beam takes the path D as denoted in Figure 2. The total distance x traveled by mirror A is $v\Delta t'$, where $\Delta t'$ is the time interval perceived by the observer. According to the second postulate of special relativity, the speed of light is constant in all inertial frames; thus, $D = c\Delta t'$. We now have constructed a triangle with the segments D, L, and x. According to the Pythagorean Theorem, $D = \sqrt{x^2 + L^2}$. Expressing this equation in terms of $v, \Delta t$, and $\Delta t'$, we obtain the expression $c\Delta t' = \sqrt{(v\Delta t')^2 + (c\Delta t)^2}$. Through algebraic manipulation, we obtain the relationship between $\Delta t'$ and Δt as

$$\Delta t' = \frac{\Delta t}{\sqrt{1 - \frac{v^2}{c^2}}} = \gamma\Delta t. \tag{41}$$

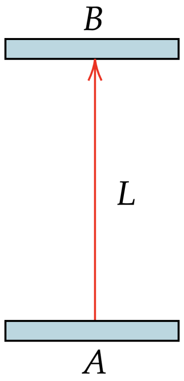

Fig. 1 Laser Beam Model before Mirror A's Movement. The laser shoots from Mirror A to Mirror B, with L representing the separation distance.

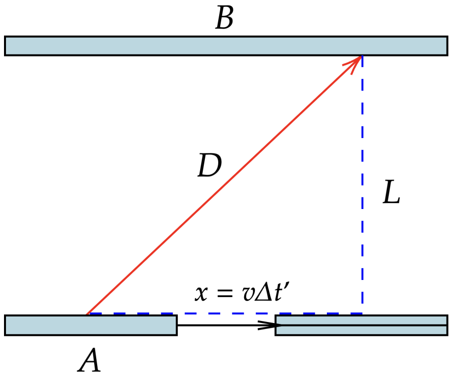

Fig. 2 Laser Beam Model After Mirror A's Movement. Mirror A has moved a distance x, and the laser now follows the path D.

For this equation to hold, the speed of light c must be a constant value in all frames, regardless of their motion. The time taken for the laser beam to travel between two mirrors relies on this postulate; otherwise, the expression $\Delta t = \frac{2L}{c}$ would not be true. In Galilean relativity, velocity is additive, and the speed of

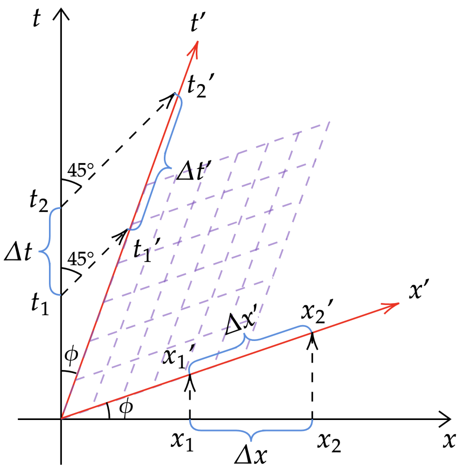

Fig. 3 Demonstration of Lorentz Transformation. The xt plane is transformed with a rapidity ϕ, and the purple lines represent the transformed coordinate system $x't'$. Events E_1 and E_2 occur at (x_1, t_1) and (x_2, t_2), respectively. The time differences are Δt in frame R and $\Delta t'$ in frame R'. The spatial separations are Δx in R.

light is not constant. When the mirror B is moving to the right, from its frame of reference, the speed of light travels slower because the light and the mirror B are in relative motion away from each other.

The second reason time dilation would not happen in Galilean relativity is because time is considered to be universal for all observers regardless of their state of motion, a concept called *absolute time*. If two events happen 1 second apart from each other in a certain frame, the two events would happen 1 second apart from each other in all frames. Thus, in Figure 2's scenario, $\Delta t' = \Delta t$, even when mirror B is moving to the right.

We could demonstrate the concept of time dilation with a coordinate system under Lorentz transformation. Figure 3 demonstrates the Lorentz transformation of the frame R, where the red line t' demonstrates the transformed time in R'. We let an arbitrary event E_1 happen at time t_1 at location $x = 0$. Expressed in coordinates, the event could be described with coordinates $(0, t_1)$. We let another event E_2 happen with coordinates $(0, t_2)$. We can describe the difference in time $t_1 - t_2$ as Δt. For an observer in the R' frame, the time of the two events happening is perceived at t'_1 and t'_2 respectively, where $t'_1 = \gamma t_1$ and $t'_2 = \gamma t_2$. We could then express the difference between t'_1 and t'_2, denoted as $\Delta t'$ in Figure 1, as

$$\Delta t' = \gamma \Delta t = \cosh(\phi)\Delta t. \tag{42}$$

6.4.3 Length Contraction

Length contraction is another consequence under special relativity. When an observer is moving relative to an object, the measured length of the object will be shorter than its length when the observer is stationary relative to the object. To derive the length contraction formula, we consider an object in a stationary frame of reference with a length L_0. L_0 is called the proper length, defined as the length of an object in its rest frame. Define a light signal traveling at the speed c along the object. The time it takes for the light to pass the entirety of the object $\Delta t = \frac{L_0}{c}$.

Now, consider an observer moving at a speed v relative to the frame where the object is located. In the observer's view, the length of the object is L, and the time it takes for the light to pass the object is $\Delta t' = \frac{L}{c}$ since the speed of light is preserved in inertial frames. Recall the time dilation formula and combine the two expressions above, we obtain

$$\Delta t' = \gamma \Delta t, \quad \frac{L}{c} = \gamma \frac{L_0}{c}.$$

Upon simplification, we obtain the equation of length contraction as

$$L = \frac{L_0}{\gamma} = L_0 \sqrt{1 - \frac{v^2}{c^2}}. \tag{43}$$

Similar to time dilation, length contraction will not happen under the framework of Galilean relativity, mainly because the speed of light is not preserved. The spatial quantity is universal across all inertial frames in Galilean relativity; thus, there is no need to account for the difference in length across various frames of reference.

7 Results and Discussion

This paper provides a foundational review of electromagnetic theory. It presented Maxwell's equations and the proof that they are not invariant in frames of references under Galilean transformation using the general solution of the plane wave equation. We presented the derivation of Lorentz transformation that maintains the property of Maxwell's equations under transformations and makes the equations invariant. Finally, we used the properties of the Lorentz transformation to present spacetime interval, time dilation, and length contraction.

This study offers a solid foundation for those seeking a deeper understanding of the relationship between electromagnetism and special relativity. It also serves as a foundation for those who wish to further engage with more advanced concepts and provides the necessary context for future studies in related fields.

8 Acknowledgment

My mentor, Ms. Bilyana Tomova from the University of Cambridge, has provided comprehensive and detailed guidance for the completion of my paper, from the selection of the topic to the completion of the paper. Although she was busy with her work, she often found time to provide me with attentive guidance and support by email. She helped me solve the problems and difficulties I encountered in my research, which made me take fewer detours in my paper. Thank you for making this research an enriching experience.

References

1 J. C. Maxwell, *Philosophical Transactions of the Royal Society of London*, 1865, **155**, 459–512.

2 O. Lodge, *The Ether of Space*, Harper, United Kingdom, 1909, pp. 1–12.

3 Y.-D. Tan, 2022.

4 H. P. Robertson, *Rev. Mod. Phys.*, 1949, **21**, 378–382.

5 M. Heald and J. Marion, *Classical Electromagnetic Radiation*, Dover Publications, Incorporated, 2012.

6 D. J. Griffiths, *Introduction to electrodynamics*, Pearson, 2013.

7 M. Saqib, F. S. N. and F. J. N., 2020 International Youth Conference on Radio Electronics, Electrical and Power Engineering (REEPE), 2020, pp. 1–5.

8 A. Kustepeli, *Electromagnetics*, 2016, **36**, 135–148.

9 S. E. Hill, *The Physics Teacher*, 2010, **48**, 410–412.

10 A. Einstein, *Annalen der Physik*, 1905, **322**, 891–921.

11 A. Chalmers, *Correspondence, Invariance and Heuristics: Essays in Honour of Heinz Post*, Springer, 1993, pp. 189–205.

12 R. P. Feynman, *Feynman Lectures on Physics: Electrical and Magnetic Behavior. Volume 4*, Perseus Books, 1999.

13 J. M. Aguirregabiria, A. Hernández and M. Rivas, *European Journal of Physics*, 2022, **43**, 035.

14 A. R. Lee and T. M. Kalotas, *American Journal of Physics*, 1975, **43**, 434–437.

15 R. Ferraro, 2007.

16 J. Barrett, 2015.

17 G. Goldoni, *Resonance*, 2007, **12**, 86–91.