A Proof of the Rationality of the Definition of Momentum in Relativity

Chengshen Xu[a,*]

[a] *Autohome Inc., 10th Floor Tower B, No. 3 Dan Ling Street Haidian District, Beijing, China*

Abstract

In this paper we firstly use a new method–the invariance of space-time interval and some simple linear algebra knowledge to derive Lorentz transformations and four-dimensional vectors. Finally we discuss and prove how to define the force and the momentum in relativity which has not been discussed and proved in textbooks and scientific literature.

The first three dimensions of a four-dimensional momentum are defined as momentum and the derivative of momentum with respect to time is defined as force. But there is a problem that the rationality of the definition of momentum is not discussed and proved. Force and momentum cannot be arbitrarily defined. Because if our senses are sensitive and sophisticated enough, only a correct definition can guarantee that when we accelerate an object with a constant force, the momentum will increase at a constant rate. It is not necessary to be discussed in classical mechanics, because in classical mechanics the force is proportional to the acceleration and the force comes before the momentum. But it is just the opposite that the momentum comes before the force in relativistic mechanics, so it's important to discuss and prove how to define the force and the momentum in relativity.

In addition the fact that the same physical process does not depend on the space-time point means that the Lorentz transformations must be linear transformations, so we can derive Lorentz transformations and four-dimensional vectors by using the invariance of space-time interval and some simple linear algebra knowledge.

Keywords: Force, Monmentum, Relativity Theory, Space-time Interval Invariance, Energy, Mass-energy Equation

1. Introduction

In classical physics, time and space are independent of each other. In d-ifferent reference frames, space interval invariance is satisfied, and time does

*Corresponding author: xcssgzs@126.com

not matter. For example, the length of a rod is fixed no matter what reference frame you measure it in. However, due to the development of electrodynamics, Maxwell put forward four differential equations related to electromagnetism, forming a system of equations in which time and space are equally weighted. Therefore, from the view of the classical time-space, the system of equations is no longer satisfied with covariance from one reference frame to another. The so-called covariance is intuitively explained by the fact that the laws of physics will not change because of the switching of the reference frame. Mathematically, it means that the forms of kinematic equations are consistent in different reference frames. So the special and general relativity were proposed by Einstein in 1905 and 1915 respectively [1, 2].

As we all know, physics textbooks and scientific literature usually define forces in the following way

$$\boldsymbol{F} = \frac{d\boldsymbol{P}}{dt} = \frac{d(m_0 \frac{\boldsymbol{v}}{\sqrt{1-\frac{v^2}{c^2}}})}{dt} = \frac{m_0 d(\frac{\boldsymbol{v}}{\sqrt{1-\frac{v^2}{c^2}}})}{dt} \tag{1}$$

where $\boldsymbol{P}$ represents the first three dimensions of the four-dimensional momentum of an object; m_0 represents the proper mass of an object, namely the rest mass; $\boldsymbol{v}$ represents the velocity of the object; v represents the speed of the object, i.e. the magnitude of the velocity; c represents the speed of light.

But there is a problem that the rationality of the definition of momentum is not discussed and proved. Force and momentum cannot be arbitrarily defined. Because if our senses are sensitive and sophisticated enough, only a correct definition can guarantee that when we accelerate an object with a constant force, the momentum will increase at a constant rate. But historical literature and textbooks fail to justify the rationality of this definition.

It is not necessary to be discussed in classical mechanics, because in classical mechanics the force is proportional to the acceleration and the force comes before the momentum. But it is just the opposite that the momentum comes before the force in relativistic mechanics, so it's important to discuss and prove how to define the force and the momentum in relativity.

In addition, the derivation of Lorentz transformations and four-dimensional vectors can also have a simpler method, because the fact that the same physical process does not depend on the space-time point means that the Lorentz transformations must be linear transformations, so we can derive these conclusions by using the invariance of space-time interval and some simple linear algebra knowledge

In Section 2 and 3 we firstly use a new method–the invariance of space time interval to derive the conclusions of special relativity, which is much easier and more natural. In Section 4 we discuss and prove how to define the force and the momentum in relativity.

2. Derive Lorentz Transformation by Using Space-time Interval Invariance

In this section use a new method–the invariance of space time interval and some simple linear algebra knowledge to derive the conclusions of special relativity.

In classical mechanics the space interval invariance is satisfied. Because time and space are equally weighted in relativity, the space-time interval invariance of two events should be satisfied–that is, the space-time interval of two events remains unchanged under different reference frames. In the inertial frame, the space-time interval is defined as follows

$$c^2(\Delta\tau)^2 = -[(\Delta x)^2 + (\Delta y)^2 + (\Delta z)^2 - c^2(\Delta t)^2] \tag{2}$$

where c is the speed of light. On the one hand, strictly speaking, space-time interval invariance already implies the principle of constancy of light velocity. Because when $\Delta\tau = 0$, no matter in which inertial frame, we have

$$v = \frac{\sqrt{(\Delta x)^2 + (\Delta y)^2 + (\Delta z)^2}}{\Delta t} = c \tag{3}$$

which implies the propagation speed of light does not change in any inertial frame.

It is precisely because of the invariance of space and time interval that the cause and effect structure of space-time can be guaranteed. Otherwise, in one reference frame, two things happen in different places at the same time, that is, the space-time interval is no more than zero, the photon cannot establish the connection between the two events. Originally, there is no causal relationship, but when it changes to another reference frame, the space-time interval is smaller than zero, the photon can establish the connection between the two things. That is, the photon starts from the place and at the time of occurrence of the one event and propagates to the place of occurrence before the time of occurrence of the occurrence of the other event, which makes the two things establish a causal relationship and destroys the causal structure of space-time, thus producing a contradiction.

In special relativity the space time interval can be represented by a matrix, namely

$$\begin{pmatrix} \Delta x & \Delta y & \Delta z & c\Delta t \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} \Delta x \\ \Delta y \\ \Delta z \\ c\Delta t \end{pmatrix} \tag{4}$$

where

$$G = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \tag{5}$$

is the metric matrix of the inertial space-time. This space is called Minkowski space. In Euclidean spaces, the orthogonal transformation keeps the length of the vector constant. Minkowski spaces are indefinite spaces, but there are corresponding pseudo-orthogonal transformations that keep the length of the vector constant, that is, the space-time interval constant. Similarly to the Euclidean spaces, in order to ensure that the length of the vector does not change after switching from one inertial frame to another, the metric matrix does not change after the transformation, which is obviously equivalent to the transformation matrix A satisfies

$$A^T G A = G \tag{6}$$

For simplicity, we assume that a moving reference frame is moving in a straight line with uniform velocity v along the x axis of the rest reference frame, then the y and z directions are not affected by anything, then we can get

$$A = \begin{cases} \begin{pmatrix} \cosh\theta & 0 & 0 & \sinh\theta \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ \sinh\theta & 0 & 0 & \cosh\theta \end{pmatrix} \\ \begin{pmatrix} \cosh\theta & 0 & 0 & \sinh\theta \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -\sinh\theta & 0 & 0 & -\cosh\theta \end{pmatrix} \\ \begin{pmatrix} -\cosh\theta & 0 & 0 & -\sinh\theta \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ \sinh\theta & 0 & 0 & \cosh\theta \end{pmatrix} \end{cases} \tag{7}$$

where

$$\cosh\theta = \gamma = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}} \geq 1, \sinh\theta = \pm\sqrt{\gamma^2 - 1} \tag{8}$$

The latter two forms are not feasible, because in combination with physical reality, only false rotation can occur in space-time, and no time or space inversion can occur. Therefore, only the first form can be taken. By writing the (7) as an equation we can get

$$\Delta x' = \gamma(\Delta x \pm \frac{v}{c} c\Delta t) \tag{9}$$

$$c\Delta t' = \gamma(\pm\frac{v}{c}\Delta x + c\Delta t) \tag{10}$$

which is the transformation between different inertial frames, known as the Lorentz transformation.

Now we define τ as mentioned above as proper time which is the time interval between two events in the reference frame in which the two events occur at the same place, or it can be viewed as the time interval of a clock bound to a moving particle.

On the basis of

$$c^2(\Delta\tau)^2 = -[(\Delta x)^2 + (\Delta y)^2 + (\Delta z)^2 - c^2(\Delta t)^2] \tag{11}$$

we can get

$$\frac{\Delta\tau}{\Delta t} = \sqrt{1 - \frac{v^2}{c^2}} = \frac{1}{\gamma} \leq 1 \tag{12}$$

3. Derive Four-dimensional Invariants by Using Space-time Interval Invariance

With the Lorentz transformation, we find the relativity of time and space. For example, if a train is moving forward along the x axis at speed v, for two events happen at different times and at the same place on the train, the relationship between the two time intervals is

$$\Delta t' = \gamma\Delta t \tag{13}$$

where t' is the time in ground inertial reference frame and t is the time in the train inertial reference frame. The time interval measured on the ground is longer than that measured by people on the train. People on the ground will feel that all physical processes on the train, including people's mind metabolism, will slow down, which is the famous clock slowing effect. Of course, motion is relative, so people on the train also feel that physical processes on the ground slow down.

And for the same time on a train, measuring a bar relative rest to the ground on the ground put in the direction of motion, there is

$$\Delta x' = \gamma\Delta x \tag{14}$$

That is, the person on the train measure the length of the bar on the ground to be shorter than the person on the ground themselves, which is known as the scale effect. This is because the actions of the person on the train reading the two coordinates of the two ends of the bar at the same time are at different times for the person on the ground, which results in the larger space interval on the ground.

The second formula of the expression of the Lorentz transformation can be written as

$$c\Delta t' = \gamma(\pm\frac{v}{c}V\Delta t + c\Delta t) \tag{15}$$

where

$$V = \frac{\Delta x}{\Delta t} \tag{16}$$

is the velocity of the object relative to the train. We divide the first formula by (15)

$$v' = \frac{v + V}{1 + \frac{vV}{c^2}} \tag{17}$$

which is the relativistic velocity synthesis formula, that is, the velocity of the second object relative to the first object is v, the velocity of the third object relative to the second object is V, then the velocity of the third object relative to the first object is v'. Of course, this is the case in which the two velocities are collinear. The case in which two velocities are non-collinear can be solved by vector decomposition and synthesis.

Of course, there's another way to derive the relativistic velocity synthesis formula. Since the vector

$$(\Delta x, \Delta y, \Delta z, c\Delta t)^T \tag{18}$$

is a vector whose length does not change under the lorentz transformation and $\Delta\tau$ is proper time interval is invariant too,

$$(\frac{\Delta x}{\Delta\tau}, \frac{\Delta y}{\Delta\tau}, \frac{\Delta z}{\Delta\tau}, c\frac{\Delta t}{\Delta\tau})^T \tag{19}$$

is also a vector with constant modulus length, therefore according to

$$(\frac{\Delta x'}{\Delta\tau}, \frac{\Delta y'}{\Delta\tau}, \frac{\Delta z'}{\Delta\tau}, c\frac{\Delta t'}{\Delta\tau})^T = A(\frac{\Delta x}{\Delta\tau}, \frac{\Delta y}{\Delta\tau}, \frac{\Delta z}{\Delta\tau}, c\frac{\Delta t}{\Delta\tau})^T \tag{20}$$

the synthesis formula can also be derived.

Because

$$(\frac{\Delta x}{\Delta\tau}, \frac{\Delta y}{\Delta\tau}, \frac{\Delta z}{\Delta\tau}, c\frac{\Delta t}{\Delta\tau})^T \tag{21}$$

has a constant length under the Lorentz transformation. Therefore, for a particle whose proper mass is m_0, its four momentum is defined as

$$P = (m_0\frac{\Delta x}{\Delta\tau}, m_0\frac{\Delta y}{\Delta\tau}, m_0\frac{\Delta z}{\Delta\tau}, m_0 c\frac{\Delta t}{\Delta\tau})^T \tag{22}$$

$$= (m_0\gamma\frac{\Delta x}{\Delta t}, m_0\gamma\frac{\Delta y}{\Delta t}, m_0\gamma\frac{\Delta z}{\Delta t}, m_0\gamma c\frac{\Delta t}{\Delta t})^T \tag{23}$$

$$= (\frac{m_0 v_x}{\sqrt{1 - \frac{v^2}{c^2}}}, \frac{m_0 v_y}{\sqrt{1 - \frac{v^2}{c^2}}}, \frac{m_0 v_z}{\sqrt{1 - \frac{v^2}{c^2}}}, \frac{m_0 c}{\sqrt{1 - \frac{v^2}{c^2}}})^T \tag{24}$$

Since m_0 is invariant, the length of the modulus of the four-momentum is also constant under the Lorentz transformation, and this invariance is called the four-momentum conservation.

4. The Derivation of Force and Momentum

We can define the force as the accelerated velocity of the object measured in the inertial reference frame which is instantaneous relative rest to the object times the proper mass of the object. And the reason why it's defined this way is very simple, because this definition is equal for every inertial reference frame, and no inertial reference frame is special. We can then restore this expression to the general inertial reference frame to obtain a reasonable definition of the force.

When we accelerate objects from velocity v to v', for the purpose of simplicity, the v and v' are assumed to be collinear. According to the relativistic velocity synthesis formula, the observed velocity change in the inertia reference frame which is instantaneous relative rest to the object is

$$\Delta\bar{\boldsymbol{v}} = \frac{\boldsymbol{v}' - \boldsymbol{v}}{1 - \frac{\boldsymbol{v}}{c} \cdot \frac{\boldsymbol{v}'}{c}} = \frac{\Delta\boldsymbol{v}}{1 - \frac{\boldsymbol{v}}{c} \cdot \frac{\boldsymbol{v}'}{c}} \Longrightarrow d\bar{\boldsymbol{v}} = \frac{d\boldsymbol{v}}{1 - \frac{\boldsymbol{v}^2}{c^2}} \tag{25}$$

And according to the Lorentz transformation expression, there is

$$\Delta t = \frac{v\Delta x'}{c^2\sqrt{1 - \frac{v^2}{c^2}}} + \frac{\Delta t'}{\sqrt{1 - \frac{v^2}{c^2}}} = \frac{\Delta t'}{\sqrt{1 - \frac{v^2}{c^2}}} \tag{26}$$

where t is the time in the rest reference and t' is the time in the inertial reference frame with v velocity. Therefore, in the inertial system with the object moving at the instantaneous speed, the object has zero velocity, and the acceleration is

$$\boldsymbol{a} = \frac{d\bar{\boldsymbol{v}}}{dt'} = \frac{d\boldsymbol{v}}{\sqrt{(1 - \frac{v^2}{c^2})^3}dt} = \frac{d\frac{\boldsymbol{v}}{\sqrt{1 - \frac{v^2}{c^2}}}}{dt} \tag{27}$$

So we can get the force in general coordinates

$$\boldsymbol{F} = m_0\boldsymbol{a} = m_0\frac{d\bar{\boldsymbol{v}}}{dt'} = \frac{m_0 d\frac{\boldsymbol{v}}{\sqrt{1 - \frac{v^2}{c^2}}}}{dt} = \frac{d\frac{m_0\boldsymbol{v}}{\sqrt{1 - \frac{v^2}{c^2}}}}{dt} = \frac{d\boldsymbol{P}}{dt} \tag{28}$$

So far we have proved the first three dimensions of four-dimensional momentum

$$\boldsymbol{P} = \frac{m_0\boldsymbol{v}}{\sqrt{1 - \frac{v^2}{c^2}}} \tag{29}$$

is really a reasonable definition of momentum.

5. The Derivation of Energy and Mass-energy Equation

The kinetic energy

$$dE_k = \boldsymbol{F} \cdot d\boldsymbol{s} = \frac{d\boldsymbol{P}}{dt} \cdot \boldsymbol{v}dt = \boldsymbol{v}d\boldsymbol{P} = m_0\boldsymbol{v}d\frac{\boldsymbol{v}}{\sqrt{1 - \frac{\boldsymbol{v}^2}{c^2}}} \tag{30}$$

Thus, the kinetic energy of an object with velocity $\boldsymbol{v}$

$$E_k = \int_0^{E_k} dE_k = \int_0^{v} m_0 \boldsymbol{v} d\frac{\boldsymbol{v}}{\sqrt{1-\frac{\boldsymbol{v}^2}{c^2}}} \tag{31}$$

$$= m_0 \frac{\boldsymbol{v}^2}{\sqrt{1-\frac{\boldsymbol{v}^2}{c^2}}}|_0^v - \int_0^v m_0 \frac{\boldsymbol{v}}{\sqrt{1-\frac{\boldsymbol{v}^2}{c^2}}} d\boldsymbol{v} \tag{32}$$

$$= m_0 \frac{\boldsymbol{v}^2}{\sqrt{1-\frac{\boldsymbol{v}^2}{c^2}}}|_0^v - \frac{1}{2}\int_0^v m_0 \frac{d\boldsymbol{v}^2}{\sqrt{1-\frac{\boldsymbol{v}^2}{c^2}}} \tag{33}$$

$$= m_0 \frac{\boldsymbol{v}^2}{\sqrt{1-\frac{\boldsymbol{v}^2}{c^2}}}|_0^v + \frac{1}{2}\int_0^v m_0 c^2 \frac{d(-\frac{\boldsymbol{v}^2}{c^2})}{\sqrt{1-\frac{\boldsymbol{v}^2}{c^2}}} \tag{34}$$

$$= m_0 \frac{\boldsymbol{v}^2}{\sqrt{1-\frac{\boldsymbol{v}^2}{c^2}}}|_0^v + m_0 c^2 \sqrt{1-\frac{\boldsymbol{v}^2}{c^2}}|_0^v \tag{35}$$

$$= \frac{m_0 c^2}{\sqrt{1-\frac{v^2}{c^2}}} - m_0 c^2 \tag{36}$$

Usually we define the motion quality as

$$m = \frac{m_0}{\sqrt{1-\frac{v^2}{c^2}}}, \tag{37}$$

so we can rewrite the kinetic energy as

$$E_k = mc^2 - m_0 c^2 \tag{38}$$

When the velocity is much less than the speed of light, i.e., $v << c$, we can obtain according to Taylor expansion

$$E_k = \frac{1}{2} m_0 v^2 \tag{39}$$

Naturally the total energy of the object is

$$E = mc^2 \tag{40}$$

This is the famous mass-energy equation. The energy is proportional to its mass, and the ratio is very large. This is why the nuclear fusion and fission reactions of atomic and hydrogen bombs can release huge amounts of energy with even a small loss of mass.

6. Conclusions

In this paper, the reasonable definitions of momentum and force in relativity theory are derived, which has not been discussed and proved in textbooks and scientific literature. It is a huge contribution to special relativity. In addition, the Lorentz transformation and four-dimensional invariants are derived by using the invariance of space-time interval and some simple linear algebra knowledge, which is a simpler method.

We have come to the conclusion that the first three dimensions of a four-dimensional momentum are indeed a reasonable definition of momentum. It satisfies invariance in different inertial reference frames.

[1] A. Einstein, Zur elektrodynamik bewegter körper, Annalen der physik 4.

[2] D. G. der Allgemeinen Relativitätstheorie, Leipzig 1916; ferner in, Annalen der Physik 49 (1916) 769–822.