A note on interpreting special relativity

A J Owen

April 12th 2022

Author's contact email address:
ajowen@physics.org

Abstract

In this note I try to assess Einstein's theory of special relativity, and decide whether or not it contains inconsistencies, as some authors have claimed. The factor ϕ in the original transforms, which Einstein sets as unity, is correct for the way coordinate time is defined via various clocks, and where relative velocity is then effectively orthogonal to the direction of observation. However, it can be reinterpreted outside SR as a scale factor that can be used to describe the Doppler effect when motion is parallel to the direction of observation. Einstein's use of simultaneity depicts a moving sphere as an oblate spheroid, whereas consideration of time delay in the coordinate frame shows it is a rotated sphere. The non-intuitive prediction of clock differences in SR is logically consistent within the theory, but the reasoning leading to a twin age difference does not recognize that clocks are not properly synchronized, nor does it include a general relativity spacetime curvature effect that negates the clock differences due to kinematical relativity alone. I have introduced the modification as a new postulate, but it may just be a consequence of the conservation of energy, or the equivalence principle.

1 Introduction

One of the most influential papers written in the history of mathematical physics was Albert Einstein's 1905 paper entitled "Zur Elektrodynamik bewegter Koerper" [1], in which he introduced a revolutionary theory, later to become known as the special theory of relativity (*SR*). However, over the century and more since its publication, numerous mathematicians and physicists have questioned its validity (see ref. [2] for a summary). A recent paper by Lev Verkhovsky [3] suggests there is a contradiction in the theory that invalidates it; Stephen Crothers also states that *SR* is logically inconsistent [4], while Laszlo Szabo claims [5] that special relativity theory tells us nothing new about the spatio-temporal features of the physical world, and that the longstanding belief that it does so is the result of a simple but subversive terminological confusion. In view of this level of scepticism, I decided to examine the situation myself, and try to come to some decision on whether or not *SR* should be regarded as correct, whether some of the objections are justified, and what the scope of the theory is. These issues have been discussed in the scientific literature countless times since *SR* was first introduced more than a century ago. Nevertheless, I believe this is still an intellectual challenge worthy of investigation.

Einstein's original paper [1] contains the following transformation of time and space coordinates between two inertial frames of reference, $O(x, y, z, t)$ and $O'(x', y', z', t')$, moving at a relative speed v in the x or x' direction (where I use my own notation):

$$x' = \phi(v)\gamma(v)\left[x - vt\right]; \quad y' = \phi(v)y; \quad z' = \phi(v)z;$$

$$t' = \phi(v)\gamma(v)\left[t - \frac{vx}{c^2}\right] \tag{1}$$

where $\gamma = 1/\sqrt{1 - v^2/c^2}$. In his paper, Einstein argued that the function $\phi(v)$ that appears as a multiplication factor affecting all terms was necessarily equal to unity, which meant that it subsequently disappeared from the transformation equations. In this note, I wish to reconsider the derivation of the transforms and the meaning of the function ϕ, as well as consider whether any other ingredients have been omitted from SR that would significantly alter the physics.

2 Deriving the transformation equations

In this section, I shall begin by trying to follow Einstein's reasoning. For the two inertial frames mentioned above, Einstein firstly considers a point X given by $X = x - vt$. I think it is initially quite difficult to understand what Einstein is aiming at here, but if you take $X = 0$, this could represent the origin of the O' frame which is advancing as $x = vt$, and where the frame origins overlap at time $t = 0$. The coordinate $x = vt + X$ then represents a fixed point in the O' frame a distance X further on, but measured in the O frame. X is thus essentially the distance between two fixed points in the moving frame, as measured from the O frame.

A pulse of light is emitted from the origin of frame O' at time t'_0 towards this point P at $x = X + vt$ along the direction of relative motion (the x or x' axis), and from there at t'_1 it is reflected back to the origin of O' where it arrives at time t'_2. The transit time interval is (postulated to be) the same for both directions in the O' frame, so we have $t'_1 - t'_0 = t'_2 - t'_1$ or $\frac{1}{2}(t'_0 + t'_2) = t'_1$. Times t' in the O' frame depend on both x and t in the O frame, so the previous equation for $t'(x,t)$ (where we have $y = z = 0$) may be written:

$$\frac{1}{2}\left[t'(0,t) + t'\left(0, t + \frac{X}{c-v} + \frac{X}{c+v}\right)\right] = t'\left(X, t + \frac{X}{c-v}\right) \qquad (2)$$

Einstein next partially differentiates this equation with respect to t'

and X, which he now regards as infinitesimally small, and obtains:

$$\frac{1}{2}\left(\frac{1}{c-v}+\frac{1}{c+v}\right)\frac{\partial t'}{\partial t}=\frac{\partial t'}{\partial X}+\frac{1}{c-v}\frac{\partial t'}{\partial t} \tag{3}$$

which on rearranging gives:

$$\frac{\partial t'}{\partial X}+\frac{v}{c^2-v^2}\frac{\partial t'}{\partial t}=0 \tag{4}$$

Due to the assumed homogeneity of space the function $t'(x,t)$ is linear in x and t, so by integration one obtains:

$$t'=a\left(t-\frac{v}{c^2-v^2}X\right) \tag{5}$$

where a is an integration constant.

In the O' frame, for a pulse of light emitted at $t'=0$ from the origin along the direction x', we have $x'=ct'$ or

$$x'=ac\left(t-\frac{v}{c^2-v^2}X\right) \tag{6}$$

Einstein then states that this pulse moves relative to the origin of O' with a velocity $(c-v)$ when measured in the O frame, so that $X/(c-v)=t$. This then gives

$$x'=a\frac{c^2}{c^2-v^2}X \tag{7}$$

and with $X=(x-vt)$ we have:

$$x'=\frac{a}{(1-v^2/c^2)}\left[x-vt\right] \tag{8}$$

Comparing this with Equation 1 we see that $a=\phi/\gamma$. Using analogous considerations for the y and z axes, the equivalent transforms for y' and z' in Equation 1 are obtained (which I won't derive here).

3 Spherical light pulse

In his paper, Einstein actually talks about a ray or beam of light ("Lichtstrahl") - but this could be potentially confusing. I think everyone believes what he meant was essentially an infinitesimally short pulse of light or light signal, rather like a classical particle, with a well-defined position and velocity (not like in quantum theory). Furthermore, if this signal is propagated in all directions from a point source, it has a sharply defined "wavefront", but it is not a wave in the ordinary sense of having a wavelength and frequency, and it doesn't interfere with itself or anything else. It is in fact quite an abstract concept in Einstein's theory, and assigning it properties other than its velocity is not part of *SR*. The theory is essentially about symmetry, space and time, but not really about the properties of light, other than its invariant speed of propagation.

One of the theory's main predictions is that an abstract pulse of light emitted from a point source propagates as a spherical wavefront even in an inertial frame of reference moving relative to the frame in which the light source is at rest. This counterintuitive consequence (attributed to the lack of a background aether frame) can be seen to follow from the transforms of Equation 1. Imagine a pulse of light is emitted from the origin of frame O', just as it passes the origin of the frame O, so that we may write for the coordinates of the wavefront:

$$x'^2 + y'^2 + z'^2 - c^2t'^2 = 0 \tag{9}$$

Using the transforms in Equations 1 to convert to the frame labelled O, we then obtain:

$$x'^2 + y'^2 + z'^2 - c^2t'^2 = \phi^2[x^2 + y^2 + z^2 - c^2t^2] = 0 \tag{10}$$

We see that ϕ appears as a scale factor that could describe the relative scales of spacetime in the two frames.

These equations contain Einstein's postulate of the invariance of the speed of light, since c has been written as the same in both frames, and we see that the light pulse propagates in both frames as a spherical wavefront, satisfying his relativity principle, as well as the isotropy of space.

Einstein provided an argument [1] in his paper to show that ϕ is necessarily equal to unity, and subsequent textbooks (see, for example, [6]) have all followed suit, with the result that generations of physicists and mathematicians have all adopted that same understanding. However, I shall show below that ϕ is not necessarily required to be unity - in which case we might be obliged to regard it as a different theory from Einstein's, even though the essential postulates are the same.

4 Time dilation

An important prediction of *SR* is time dilation, which can be described, as follows, keeping the function ϕ in the analysis. Now write $y' = y = 0$; $z' = z = 0$ for relative motion of the frames in the x or x' direction, and we have:

$$x'^2 - c^2t'^2 = \phi(v)^2[x^2 - c^2t^2] \tag{11}$$

With $x' = 0$ and $x = vt$, where v is the relative speed of the frames, we have for time t' in frame O' compared to time t in frame O:

$$\frac{t'}{t} = \phi\sqrt{1 - \frac{v^2}{c^2}} = \frac{\phi}{\gamma} \tag{12}$$

With $\phi = 1$ this is the usual expression in *SR* for kinematical time dilation, interpreted to mean that a clock in a frame O' moving relative to an observer O runs at a slower rate ($t' < t$ for $v > 0$).

So, what does this factor $\phi(v)$ represent? Firstly, imagine you are an observer O standing on a railway platform, and a train passes through the station representing frame O' travelling at right angles to your line of sight. You also have many compatriots ("book-keepers") positioned at intervals along the very long platform, and between you you infer that there is time dilation given by Equation 12 with $\phi = 1$, viz. the time t' on the proper clock seen in the train seems to be ticking more slowly than the coordinate clocks t held by observers all along the platform. Importantly, in order to measure these coordinate time intervals, Einstein envisages more than one clock is required in the O frame, and that these clocks are all synchronized and pass on their information between each other without any time delay, or at least with any time delay due to the finite speed of light having been accounted for. If a train passes in the opposite direction, the same time dilation is measured, irrespective of the direction (because it enters the equation as v^2). This is the case where $\phi = 1$, and this is how Einstein envisaged and defined the situation.

However, if you look along the track at the clock in the receding train, you observe that the time dilation or clock speed has altered, due to the Doppler effect. This is not now part of Einstein's SR set-up, it seems, because you as an observer are essentially stationary and have a single clock (without relying on book-keeper information). There is an ever-increasing time delay as the light waves you are receiving are stretched out while the clock recedes, and given by the ratio $c/(c+v)$, where v is the velocity of the receding train. This phenomenon is telling us that the scale of space has increased for the receding frame O'. We then have

$$\frac{t'}{t} = \left(\frac{c}{c+v}\right)\sqrt{1 - \frac{v^2}{c^2}} = \sqrt{\frac{c-v}{c+v}} \qquad (13)$$

which is the well-known result for the relativistic Doppler effect, in this case a red-shift, From this we have an interpretation for the meaning of the function ϕ:

$$\phi = \left(\frac{c}{c+v}\right) \qquad [receding] \qquad (14)$$

It represents the way space has been dynamically scaled due to the relative motion of the frame O'. The equivalent expressions for an approaching frame, where a blue-shift occurs, are obtained by changing the sign before v, and then we have

$$\frac{t'}{t} = \left(\frac{c}{c-v}\right)\sqrt{1-\frac{v^2}{c^2}} = \sqrt{\frac{c+v}{c-v}};\ \ \phi = \left(\frac{c}{c-v}\right)\ \ [\mathit{approaching}] \tag{15}$$

We can easily calculate that this leads to an asymmetrical time difference between the two frames. For example, take the numerical example of a journey in frame O' to a distant place 3 light years away at a speed $v/c = 3/5$. We have $1/\gamma = 4/5$ and $t = 5$ years in frame O. This gives $t'/t = 1/2$, but the observation of this lasts for 8 years, since it takes 5 years according to observer O for O' to arrive at the destination plus 3 years for the light to travel back to the observer, i.e. the O' clock ticks at half rate for 8 years = 4 years, while the resting observer ages by 5 years. On a return journey, O observes the O' clock ticking at double rate for two years ($= 5 - 3$ years), again 4 years on the O' clock, instead of 5 years. The age or clock difference occurs on both receding and approaching legs of the journey, so it is theoretically not necessary for the traveller to return to prove he/she has aged differently. Ultimately, the overall ageing effect is given by the Lorentz factor alone, in this case 5/4, and the Doppler shift just gives us some ongoing information during the course of the thought experiment.

In his original paper, Einstein [1] argued that $\phi(v) = \phi(-v)$ and $\phi(v)\phi(-v) = 1$, which gives $\phi = 1$. In other words, the dynamical scale of space was not considered, and we can see that Einstein was in effect concerning himself only with relative motion at right angles to the line of sight, i.e. for a passing frame, and then it is correct to write $\phi(v) = 1$, since at that particular point, $v = 0$ in the expression for ϕ in Equation 14 (as it suddenly changes from plus to minus, from approaching to receding).

To repeat myself, the whole essence of SR is that the coordinate frame is equipped with clocks throughout the frame in different spatial locations, which are all synchronized with each other, and then the motion is always transverse to the observer at the point of measurement, and time delay in receiving this information is not considered.

Thus, the relativistic Doppler effect is treated as an "add-on" effect, but I have shown that SR can be augmented to incorporate this effect by recognizing that it is described by the function ϕ in the transformation equations. This has also been suggested by Verkhovsky [3].

5 Discussion

As explained above, Einstein took great trouble in his paper to define all the quantities he used rigorously, but it does seem strange to me, and potentially confusing, that coordinate time is a quantity that requires many clocks for its measurement, spread out in the observer frame and synchronized with each other. Whereas one of the postulates of SR is that the speed of light is invariant in relatively moving inertial reference frames, the use of these "book-keeper" clocks is tantamount to taking the speed of light to be infinite for the purpose of recording the time an event occurred. This procedure is not inconsistent when defined correctly, as Einstein has done. Nevertheless, I shall now obtain a coordinate transformation that averts this exchange of information between coordinate clocks and uses just a single clock at the origin of each frame of reference. The reader will probably already have guessed what the outcome will be.

Previously we established that

$$t = \frac{t'}{\sqrt{1 - v^2/c^2}} \tag{16}$$

for $x' = 0$ and $x = vt$. In the equation, the time interval t is essentially determined via two coordinate clocks separated by a distance $x = vt$, and the time it takes to relay the signal back to the origin for comparison is ignored. Now add this time to t, and we have for the total (return) time t^*:

$$t^* = t + \frac{vt}{c} = t\left(\frac{c}{c+v}\right) = t'\sqrt{\frac{c+v}{c-v}} \qquad [x' = 0; x = vt] \tag{17}$$

which is the same as Equation 13. It incorporates the Doppler effect, but the definition of a time interval is different from Einstein's. Now, a time interval is measured in the same location, rather than being a recording of simultaneous events. Both interpretations are equivalent and consistent. Recognizing the difference clears up any confusion.

6 Addition of velocities

Now consider the addition of velocities in a straight line, where observer O sees frame O' receding at speed u and frame O' observes frame O'' receding at speed v. The usual analysis (with $\phi = 1$) gives for the relativistic sum of the velocities $w = (u + v)/(1 + uv/c^2)$. Instead of using $\phi = 1$, however, by retaining ϕ in the analysis we obtain the following simultaneous equations from the transforms:

$$\phi(u)\gamma(u)\phi(v)\gamma(v)[1 + uv/c^2] = \phi(w)\gamma(w) \tag{18}$$

$$\phi(u)\gamma(u)\phi(v)\gamma(v)[u + v] = \phi(w)\gamma(w)w \tag{19}$$

These equations can be rearranged as:

$$\frac{\phi(u)\gamma(u)\phi(v)\gamma(v)}{\phi(w)\gamma(w)} = \frac{w}{u+v} = \frac{1}{[1 + uv/c^2]} \tag{20}$$

It can be seen, irrespective of the values of ϕ, that the usual formula for the addition of velocities is obtained, and after some mathematical gymnastics, it can also be shown we must have

$$\phi(w) = \phi(u)\,\phi(v) \tag{21}$$

which is understandable, since ϕ is a scale factor, and this expression describes the transform from O to O'' via O'. Since the result for the addition of velocities does not involve the specific values of the ϕ functions, setting them all equal to unity, as Einstein did, nevertheless delivers the correct result.

7 Length contraction

In his paper [1], Einstein states that a rigid body at rest, which has the shape of a sphere, has in a state of motion - viewed from the stationary system - the form of an ellipsoid of revolution. In the direction of motion it appears shortened, whereas the other dimensions are not changed. This effect, called Lorentz contraction, has imprinted itself on our perception of special relativity. However, I shall show in this section that this is incorrect, or at least only correct in Einstein's *SR* abstract thought experiment.

Consider a cube of side length L in motion along the x axis, with y vertical. The front face (xy plane) is indeed contracted along the x direction by the factor $\sqrt{1 - v^2/c^2}$, while in Einstein's *SR* the other faces remain unaffected. This is under Einstein's definition of coordinate time, but if you take into account the fact that light takes longer to reach the observer from the back of the cube, than from the front of the cube, then the motion shows up the previously invisible yz face as rotated towards the x direction by an angle θ given by $\sin\theta = vL/c$. (For further reference, this has been explained neatly by David Appell in [8]). The shortening of the xy face in the x direction by the Lorentz factor is also equivalent to a rotation by the same angle, so that the complete picture for an observer viewing the moving cube orthogonally is that it has rotated about a vertical axis by an angle θ, which increases with velocity v. For an object approaching the speed of light, the angle approaches 90 degrees, which is an intriguing thought in the case of light photons (which I shall not pursue here).

In the case of a sphere in motion, then, the theory does not predict a foreshortening of the sphere into an oblate spheroid, but just a rotation, and so it will appear unchanged in form, i.e. it remains a sphere. Einstein's coordinate time definition in SR is therefore not helpful in trying to understand relativity from a visual point of view, and it is more instructive to take observational time delay into consideration, as I have done in this paper.

8 Further interpretation

I have verified above that Einstein's SR does indeed predict a clock difference - as in the twin paradox - and that even when the function ϕ is included via a different time interval definition, this does not alter that prediction. However, in this section I shall propose an additional postulate that does negate a clock difference.

It is always assumed in such abstract thought experiments that clocks are identical and have been synchronized with each other to tick at the same rate when compared side-by-side. However, this condition is not specifically met in the example here with regard to the clock in the O' frame compared to clocks in the O frame. To be precise we must firstly synchronize clocks O and O' in the same spatial location on the imagined station platform - and then consider what happens. This requires that twin O' boards an initially stationary train, which then accelerates up to the cruising speed v. The train engines convert chemical energy into mechanical work on everything in the train (all the masses) to produce an acceleration, which increases the kinetic energy of the train and its contents, including the clock.

It is well-known from Einstein's other relativity theory, general relativity (GR), that an additional time dilation effect occurs in a curved spacetime, such as that caused by a gravitational field. For sufficiently weak gravitational potential changes $U(\tilde{r})$, one has the (approximate) relation

$$t' = t\sqrt{1 - \frac{v^2}{c^2} + \frac{2U}{m_0 c^2}} \tag{22}$$

where m_0 is the rest mass, and $\tilde{r}$ is a radial coordinate (see, for example, ref. [7]).

There is no gravitational field *per se* in the present thought experiment we are discussing. However, it is suggestive to suppose that an increase in potential energy and an increase in kinetic energy have similar effects on the progression of time. Thus, I shall make this additional postulate: an increase in kinetic energy results in an increase in clock rate identical to that caused by an equivalent increase in potential energy.

We may then write $U = \frac{1}{2} m_0 v^2$ in Equation 22 and the total time dilation becomes zero. In this case, then, an increase in clock rate due to increased energy exactly offsets the decrease in clock rate that occurs as a result of the Lorentzian symmetry of spacetime. The outcome is that no change in clock rate occurs, on the proviso that all clocks have been initially synchronized at rest beforehand.

Kinematical time dilation is, of course, an experimentally measured fact. The reason it is observed in the decay of elementary particles, such as cosmic muons passing through the Earth's atmosphere, is that their subatomic clocks were never synchronized with cosmic muons in the laboratory - they were effectively created at high speed - and then the relativistic kinematical time dilation effect of *SR* becomes specifically apparent.

Many scientists believe that experiments with clocks in aeroplanes, such as the Hafele-Keating experiment, provide unequivocal proof of the age differences predicted by *SR* in the so-called twin paradox. However, I do not think they are a good test, for various reasons:

(a) the circular motion around the Earth is not the standard linear configuration of *SR*, (b) the Sagnac effect plays a vital rôle, i.e., the Earth itself is rotating, and (c) gravitational time dilation itself is the dominant effect.

In conclusion, I have introduced a theoretical modification as a new postulate, because the kinematical approach to time dilation in *SR* does not include all the ingredients necessary for understanding the described thought experiment. The new postulate is possibly just a consequence of the conservation of energy principle, or the equivalence principle.

References

[1] A.Einstein, *Zur Elektrodynamik bewegter Körper*, Annalen der Physik, 17 : 891, (June 1905)

[2] *Criticism of the theory of relativity*, en.wikipedia.org/wiki/

[3] L.I.Verkhovsky, *Memoir on the theory of relativity and unified field theory*, viXra 1802.0136

[4] S.J.Crothers, http://www.sjcrothers.plasmaresources.com/

[5] L.E.Szabó, *Lorentzian theories vs. Einsteinian relativity*, Vienna Circle and Hungary, Berlin and New York, Springer

[6] L.D.Landau, E.M.Lifshitz, *The Classical Theory of Fields,* Pergamon Press Ltd, 1971

[7] U.Krey, A.J.Owen, *Basic Theoretical Physics*, Springer, Heidelberg, 2007

[8] D.Appell, *The invisibility of length contraction*, Physics World, August 2019