On a Relativity of Causality

Yáng Zhèngyuán

June 30, 2020

***Abstract*—The low velocity and instantaneous velocity limits of the special theory of relativity are discussed. A relativity of causality in terms of matter waves is proposed to provide a conceptual understanding for superluminal propagations, with benefits for reconciling locality and nonlocality.**

***Keywords*—special relativity, causality, superluminal speeds, instantaneous speeds, EPR, nonlocality**

I. Introduction

In its modern formulation, the special theory of relativity is said to be based on two postulates:

1. The Principle of Relativity: the laws of physics stay form invariant in transformations between inertial reference frames.
2. The Principle of an Invariant Speed: there exists a speed c that is the same in all inertial reference frames.

From these postulates, it is then possible to derive the Lorentz transformations. In one spatial dimension, the transformations take the form:

$$x' = \gamma(x - vt) \text{ and } t' = \gamma(t - \frac{vx}{c^2}) \quad (1)$$

where $\gamma = \left(\sqrt{1 - v^2/c^2}\right)^{-1}$ is the Lorentz factor, and v is the speed of the primed reference frame relative to the unprimed reference frame in the common x direction. The more commonly seen Galilean transformations in one dimension:

$$x' = x - vt \text{ and } t' = t \quad (2)$$

are treated as a "classical limit" of the Lorentz transforms [1][2].

Though the Galilean limit $v \ll c$ is often written as $c \to \infty$, this rendering is inappropriate as it leads to an asymmetry: it does not allow consideration of speeds greater than c. The correct form of the limit should instead be the low velocity limit $v \to 0$, as there is in fact another known "classical limit" that receives less attention, that of the superluminal $v \gg c$ regime. This limit is the instantaneous limit $v \to \infty$, which produce the Carroll transformations [3] in one dimension:

$$x' = x \text{ and } t' = t - \frac{vx}{c^2} \quad (3)$$

These transformations are reminiscent of a quote in Lewis Carroll's book *Alice in Wonderland*: "here we must run as fast as we can, just to stay in place," hence the choice of name.

That the Carroll transforms do not receive greater attention is likely due to the common misconception that the invariance of the speed c in all inertial reference frames forbids speeds greater than c as violating the principle of causality. Though not widely appreciated, it is already known that superluminal speeds $v > c$ can exist without violating causality, and there is already experimental observation of superluminal phase velocities in cases of anomalous dispersion of light waves in media [2].

The symmetry between the Galilean and Carroll transforms should not be neglected, as it is connected to particle-wave duality. Historically, it was the treatment of light as a particle that contributed to the formulation of the special theory of relativity, due to the ease of understanding the relationship between the Galilean and Lorentz transformations in particle terms. Meanwhile, the Carroll transformations appeared in the literature decades after the formulation of special relativity, when it was discovered that any attempt to formulate a Galilean limit for electromagnetism must actually return two limits [4], with one corresponding to superluminal speeds $v > c$. This discovery was based on the wave nature of electromagnetism.

For instance, it is by using the Carroll transforms (3) on a wave with angular frequency $\omega = 2\pi/T$ and angular wavenumber $k = 2\pi/\lambda$, where T is the period of the wave and λ is the wavelength, that lead to the classical Doppler shift:

$$k' = k \text{ and } \omega' = \omega - vk \quad (4)$$

and not the Galilean transforms, which would instead lead to:

$$k' = k - \frac{v\omega}{c^2} \text{ and } \omega' = \omega \quad (5)$$

When considering this symmetry between the two classical limits of low velocity $v \to 0$ and instantaneous velocity limit $v \to \infty$, it is possible to use wave terms to extend the relativity of simultaneity to conceptualize a relativity of causality. A relativity of causality in turn implies nonlocal features, which helps reconcile the special theory of relativity with quantum mechanics.

II. Special Relativity

For simplicity, all discussion will be in one spatial and time dimension, with relative motion along the spatial dimension. The results are easily generalizable to three spatial dimensions.

In the Galilean limit, time is invariant under all transforms, meaning that time is held as absolute and simultaneous events remain simultaneous in all inertial reference frames. Under the Lorentz transformations, this is no longer the case, allowing for events that are simultaneous in one reference frame to precede and follow one another in another reference frame, and then reverse order in another reference frame.

A. *Relativity of Simultaneity*

Define two events A and B separated in spacetime. In a certain reference frame, A is located to the left of the origin $(ct, x) = (0,0)$ at the point $(0, -b)$ while B is located to the right at $(0, b)$. In this reference frame, the two events are simultaneous at time $t = 0$.

Transforming to another reference frame moving at velocity v along the x axis, under the Lorentz transformations (1), the spacetime location of event A becomes $(\gamma vb/c^2, -\gamma b)$ while the spacetime location of event B becomes $(-\gamma vb/c^2, \gamma b)$, meaning that event B precedes event A in time. Alternatively, in another reference frame moving at a negative velocity $-v$ along the x axis, the spacetime location of event A becomes $(-\gamma vb/c^2, -\gamma b)$ while event B becomes $(\gamma vb/c^2, \gamma b)$ and instead follows event A in time. If one were to attempt to insert a force carrier particle traveling between the two spacetime events, the carrier particle's velocity would be superluminal. In the frames where the events are not simultaneous, a carrier particle traveling between the events would have velocity $u = c^2/v$, and in the frame where events A and B are simultaneous, the velocity of a particle traveling between them would have to be instantaneous $u \to \infty$. Worse still, a carrier particle would have to leave A to reach B in one frame but leave B to reach A in another. This paradox appears to violate the principle of causality, leading to the argument that that there can be no superluminal speeds $v > c$ in the special theory of relativity.

The luminal speed c therefore allows the definition of two different categories of spacetime intervals between events: time-like and space-like. A time-like interval is an interval more separated by time than space: $s^2 = (ct)^2 - x^2 > 0$, allowing the propagation of particles with subluminal speeds and carry causal effects between two events. A space-like interval is more separated by space than time: $s^2 = (ct)^2 - x^2 < 0$, and a particle connecting two space-like separated events would have to propagate at superluminal speeds. Furthermore, there could exist a reference frame where the two events are simultaneous, leading to an instantaneous speed, along with reference frames where the order of events could reverse. This all would appear to prevent any causal link between two space-like separated events [1][2].

Note however that all this discussion has been using particle terminology, as in imagining photons or other carrier particles traveling between spacetime points. An alternative approach would be to consider relationships between spacetime points in terms of waves.

B. *Relativity of Causality*

Define two space-like separated events, A and B, that are simultaneous in a certain inertial reference frame. However, rather than considering only general events in spacetime, consider a specific case of a lab frame containing two particles of equal mass moving in opposite directions, each with speed u, such that the worldline of particle A intersects spacetime point $(0, -b)$ and the worldline of particle B intersects spacetime point $(0, b)$. In the rest frame of particle A, particle B will be moving with velocity:

$$v = \frac{2u}{1 + u^2/c^2} \tag{6}$$

In the rest frame of particle B, particle A will be moving with velocity $-v$. Now suppose that particle A had an associated scalar plane wave $A(x, t)$. This wave can be conceptualized as a de Broglie matter wave, which does not propagate in space in the rest frame of the particle, but can be treated as appearing to propagate in frames where the particle is in motion [5]. It must be emphasized that the de Broglie matter wave is *not* the wave function of quantum mechanics or the pilot wave of the de Broglie-Bohm formulation of quantum mechanics.

In the lab frame where particle A moves left with velocity $-u$, the associated matter wave has the general form:

$$A(x, t) = y_0 \sin(k_A x + \omega_A t) \tag{7}$$

where y_0 is the wave amplitude, k_A is the angular wavenumber, and ω_A is the angular frequency. Let there be another wave $B(x, t)$ associated with the particle B moving right with velocity u, which would have the general form:

$$B(x, t) = y_0 \sin(k_B x - \omega_B t) \tag{8}$$

In the lab frame, both particles are in motion, such that for an observer in the lab frame located between them, each wave undergoes a relativistic Doppler shift:

$$\frac{\lambda_r}{\lambda_s} = \frac{f_s}{f_r} = \sqrt{\frac{1 + v/c}{1 - v/c}} \tag{9}$$

Where r is the receiver, s is the source, and v is the relative velocity between the receiver and source frames [2]. From the symmetry of the two particles in the lab frame, the relativistic Doppler shift would be of the same magnitude such that $k_A = k_B = k$ and $\omega_A = \omega_B = \omega$, and a superposition of the two waves between the two particles in the lab frame would be:

$$A(x, t) + B(x, t) = 2y_0 \sin(kx) \cos(\omega t) \tag{10}$$

This defines a wave with stationary spatial dependence, that is a standing wave. Note that a standing wave can be loosely interpreted as having zero group velocity $v_g = 0$ and a phase velocity that approaches the instantaneous limit $v_p \to \infty$, though this is not problematic as there is no physical propagation.

In approaching the limit of the rest frame of particle A, the associated wave $A(x, t)$ remains stationary. Though this can be very loosely interpreted as a standing wave which as above implies a phase velocity approaching the instantaneous limit $v_p \to \infty$, there is no actual physical propagation for a standing matter wave. Meanwhile, the other matter wave $B'(x,t)$ undergoes a relativistic Doppler shift due to a relative velocity v as defined in (6). The same is true for wave $A'(x,t)$ in the rest frame of particle B, except in reversed direction.

From the relativistic Doppler shift (9), the received wave $B'(x,t)$ has the general form:

$$B'(x,t) = A_0 \sin\left(\sqrt{\frac{1+v/c}{1-v/c}}\,(k_B x - \omega_B t)\right) \quad (11)$$

Due to the shift in wavenumber and frequency, the superposition $A(x,t) + B'(x,t)$ is no longer a pure standing wave in the rest frame of particle A, but instead forms a beat, an interference pattern with frequency equal to the difference in the component frequencies [6]. Starting from the superposition:

$$A + B' = y_0 \sin(k_A x + \omega_A t) + y_0 \sin(k_B' x - \omega_B{}' t) \quad (12)$$

where the primed values are given by the relativistic Doppler shift as derived from the Lorentz transformations:

$$k_B' = k_B \sqrt{\frac{1+v/c}{1-v/c}} \text{ and } \omega_B' = \omega_B \sqrt{\frac{1+v/c}{1-v/c}} \quad (13)$$

rearranging the terms in (12) makes the beat more explicit:

$$2y_0 \sin\left(\frac{k_A+k_B'}{2}x + \frac{\omega_A-\omega_B'}{2}t\right) \cos\left(\frac{k_A-k_B'}{2}x + \frac{\omega_A+\omega_B'}{2}t\right) \quad (14)$$

In analogy to classical wave beats, the phase velocity of the beat is given by the sine term and the group velocity of the beat is given by the cosine term:

$$v_p = -\frac{\omega_A-\omega_B'}{k_A+k_B'} \quad (15)$$

$$v_g = -\frac{\omega_A+\omega_B'}{k_A-k_B'} \quad (16)$$

As expected, without the Lorentz transformations, the frequency and wavenumber terms in the superposition would be of equal magnitude and cancel, returning a pure standing wave, the phase velocity would be zero $v_p = 0$, and the group velocity would approach the instantaneous limit $v_g \to \infty$. Approaching the limit of the rest frame of particle A where the relative velocity v is positive and the particle B is moving right with respect to particle A, the group velocity takes a net positive value, propagating right from A to B, as if A were sending a signal to B. By symmetry, in approaching the limit of the rest frame of B, there is instead a beat resulting from the superposition $A'(x,t) + B(x,t)$ that has the group velocity in the negative $x$direction as if B were sending a signal to A.

From this perspective, if some signal were to be associated with the group velocity of the beat, a propagation from A to B would readily transform to a propagation from B to A with a change in reference frame. That is, one reference frame could be understood as containing an event A causing an event B, and another reference frame understood as containing an event B causing an event A. From the perspective of the particle frames, A sees that it is sending B a signal, B sees that it is sending A a signal, and in the special theory of relativity both views are apparently correct. Meanwhile, in the symmetrical lab frame of a standing wave due to equal and opposite propagations between A and B, there is a group velocity approaching the instantaneous limit $v_g \to \infty$, so that it appears that the two particles are signaling each other instantaneously, even though there is no net physical propagation of a beat.

III. Quantum Mechanics

A. Analogy to the EPR Argument

The similarity between the above hypothetical experiment of two particles moving in opposite directions and the EPR argument [7][8] is intentional. That is, it is hoped that the wave paradigm helps resolve any conceptual difficulties in appreciating the nonlocality of quantum mechanics as seen in the violation of Bell's Inequality.

To review the EPR argument, consider an EPR-Bohm [9] spin configuration, where a particle of zero spin at rest in the lab frame decays into a particle and antiparticle traveling in opposite directions from one another. By the conservation of angular momentum, when the spin of one particle is measured along an axis, it is immediately known that the spin of the other particle along that axis must be in the opposite direction. In the orthodox interpretation of quantum mechanics, the value of an observable is not fixed until it is measured, and there is no value prior to measurement. In the EPR argument, this interpretation is unacceptable because it would require an instantaneous "spooky action at a distance" propagation from one particle to the other. The EPR argument therefore assumes locality to infer a realist interpretation of quantum mechanics, that the particles already had spin values prior to measurement. The information on these prior values would be an unknown "hidden variable" that does not appear in standard quantum mechanics. However, any local hidden variable theory would have to return probabilities that obey Bell's Inequality [10], and there is experimental evidence that quantum mechanics violates Bell's Inequality [11]. Quantum mechanics, with or without hidden variables, must therefore be nonlocal. As Bell [10] concluded:

> *In a theory in which parameters are added to quantum mechanics to determine the results of individual measurements, without changing the statistical predictions, there must be a mechanism whereby the setting of one measuring device can influence the reading of another instrument, however remote. Moreover, the signal involved must propagate instantaneously, so that such a theory could not be Lorentz invariant.*

Bell may have been too hasty in concluding that a theory of instantaneous propagations could not be made Lorentz invariant, for as seen above, the special theory of relativity can be shown to naturally include such instantaneous propagations.

B. *Speed of Information*

By interpreting the instantaneous propagations between the EPR entangled particles as the result of a superposition of their corresponding matter waves, it should be possible to predict a "speed of information" or "speed of correlation" between the entangled particles in terms of their relative velocities by associating such a speed with the group velocity of the beat.

In quantum mechanics, the angular wavelength and angular frequency are related to the energy and momentum of a particle by the de Broglie relations:

$$E = \hbar\omega \text{ and } p = \hbar k \tag{17}$$

Note that in formulating his original matter wave argument, de Broglie used relativistic energy and momentum, which transform as $E' = \gamma E$ and $p' = \gamma p$ [5]. Therefore, using the mass-energy relation in the special theory of relativity:

$$E = m_0 c^2 \tag{18}$$

where m_0 is the rest mass of a particle, it then follows that the associated velocities of a beat (16) in the rest frame of particle A would be related to the velocity of particle B as:

$$v_p = -\frac{\omega-\omega'}{k+k'} = \frac{-m_0c^2+\gamma m_0 c^2}{\gamma m_0 v} = \frac{(\gamma-1)}{\gamma}\frac{c^2}{v} \tag{19}$$

$$v_g = -\frac{\omega+\omega'}{k-k'} = \frac{m_0c^2+\gamma m_0 c^2}{\gamma m_0 v} = \frac{(\gamma+1)}{\gamma}\frac{c^2}{v} \tag{20}$$

Restricting to positive relative velocity v, then:

$$v_p = \frac{c}{v}\left(c - \sqrt{c^2 - v^2}\right) \tag{21}$$

$$v_g = \frac{c}{v}\left(c + \sqrt{c^2 - v^2}\right) \tag{22}$$

Where the group velocity is superluminal for $v < c$ and can grow arbitrarily large toward the instantaneous limit as the relative velocity v approaches the low velocity limit.

Therefore, it would be expected that if the group velocity is associated with the speed of nonlocal correlations in EPR type experiments, measurements of the speed [11] should find varying speeds dependent on the relative motion of the entangled particles, and there is a lower bound of $v_g \to c$ corresponding to the luminal limit $v \to c$ of the relative velocity of the other particle.

IV. Discussion

In a wave paradigm, the instantaneous velocity limit $v \to \infty$ should be considered as valid as the low velocity limit $v \to 0$, and indeed the two limits are closely connected in wave phenomena: as one velocity approaches one limit there is often another velocity approaching the other. A wave paradigm also resolves the conceptual difficulties of accepting the possibility of a relativity of causality that could be implied by the relativity of simultaneity. Whereas it might be difficult to conceive how it would be valid for a particle to propagate from A to B in one frame but B to A in another, such a situation could readily be conceptualized as a superposition of waves between A and B.

If one were to attempt to return to a carrier particle picture, one could attempt to associate a particle with the propagating beat. Then, in the other frame where the direction is reversed, there could be an associated antiparticle propagating in the opposite direction, in accord with the observation that the mathematical description of an antiparticle resembles that of a particle traveling backwards in time due to CPT symmetry [12]. A relativity of causality in this form would resolve any grandfather paradoxes: in one reference frame, an event A would appear to emit a superluminal particle that is absorbed at event B, but in the perspective of the other reference frame, it must instead appear that B emitted an antiparticle that was absorbed at event A. The required symmetry would enforce the restriction of all backwards causality to closed causal loops and prevent a paradox. Furthermore, since the relativity of causality view requires each physical particle A and B to observe itself as the emitter in its own reference frame, it resolves any problem of freedom of choice in measurement of an EPR type experiment. In any frame where A is measured first, the measurer sees particle A propagating a signal to particle B and not vice versa. There is even freedom of choice for particle B to not even be measured.

However, the main hope in this discussion of the conceptual features of reintroducing a wave terminology is to allow a reinterpretation of physical phenomenon that potentially removes any unease over apparent contradictions between locality and nonlocality. Specifically, it demonstrates that although the special theory of relativity is commonly believed to enforce locality by interpreting the invariant speed c as a limiting speed, a wave picture allows the special theory to be interpreted as a nonlocal theory as well, which addresses any concerns on how to reconcile the nonlocality of quantum mechanics with the special theory of relativity.

From a more metaphysical standpoint, though classical wave theories were intended to restore locality by providing a medium of local propagations between two points, the concept of waves actually reintroduces nonlocality, since the values of a point on a wave necessarily depends on all other points in the wave. Furthermore, though nonlocality is a feature of the action at a distance formulation of Newtonian gravity between two particles, the particle concept leads to locality by allowing the separability of two points into distinct particles.

Much as the development of quantum theories led the conceptual understanding of physics to move away from "wave *or* particle" and toward "wave *and* particle," future physical theories will likely have to move toward a "nonlocal *and* local" paradigm.

V. Conclusion

The low velocity and instantaneous velocity limits of the special theory of relativity relate to the concept of particle-wave duality in quantum mechanics. Applying a matter wave interpretation to the special theory of relativity allows a conceptual understanding of a relativity of causality, which in turn addresses some of the conceptual difficulties of nonlocal and superluminal interactions.

Acknowledgment

This work was privately sponsored.

References

[1] Goldstein, H., Poole, C. P., Safko, J. L. *Classical Mechanics*, 3rd Edition. San Francisco, CA: Addisson-Wesley, 2002.

[2] Jackson, John David, *Classical Electrodyanmics*, 3rd Edition. New York, NY: Wiley, 1999.

[3] Lévy-Leblond, J.-M. "Une nouvelle limite non-relativiste du groupe de Poincaré." *Ann. Inst. H. Poincaré* A **3**: 1-12, 1965.

[4] Le Belac, M., Lévy-Leblond, J.-M. "Galilean Electromagnetism" *Il Nuovo Cimento B*. **14**: 217-234, 1973.

[5] De Broglie, L. "The wave nature of the electron." Nobel Lecture, 1929.

[6] Marion, J. B., Thornton, S. T., *Classical Dynamics of Particles and Systems*, 4th Edition. Orlando, FL: Harcourt, 1995.

[7] Griffiths, David J. *Introduction to Quantum Mechanics*, 2nd Edition. Upper Saddle River, NJ: Pearson Prentice Hall, 2005.

[8] Einstein, A., Podolsky, B, and Rosen, N., "Can quantum-mechanical description of physical reality be considered complete?" *Phys Rev*, **47**: 777-780, 1935.

[9] Bohm, David. *Quantum Theory*, 1st Edition. Englewood Cliffs, NJ: Prentice-Hall, 1951.

[10] Bell, J. S. "On the Einstein Podolsky Rosen Paradox," *Physics*, **1** (3): 195-200, 1964.

[11] Yin, J. Cao, Y. Yong, H.-L., Ren, J.-G., Liang, H. Liao, S.-K., Zhou, F. Liu, C. Wu, Y.-P., Pan, G.S., Zhang, Q. Peng, C.-Z., Pan, J.-W. "Bounding the speed of 'spooky action at a distance.'" *Phys. Rev. Lett.* **110**, 260407, 2013.

[12] Peskin, M. E., Schroeder, D. V., *An Introduction to Quantum Field Theory*, Boulder, CO: Westview, 1995.