Einstein and the Special Theory of Relativity

Supurna Sinha

Supurna Sinha got her PhD in condensed matter physics from Syracuse University. She did postdoctoral work in theoretical physics at Indian Institute of Science and Raman Research Institute, Bangalore. Her interests include writing popular physics articles and making picture books for children.

Physics Before 1905

For centuries, physics was dominated by Newton's ideas about space, time and mechanics. These ideas led to a very successful description of the solar system. In the 1800s the theory of electromagnetism came into being with the work of Faraday, Maxwell and others. This theory predicted the existence of electromagnetic waves which move at a speed c, which is a fundamental constant of nature. This was not easy to understand using Newtonian ideas. From the resulting tension between Newtonian physics and electromagnetic theory was born the 'Special Theory of Relativity'.

In the February 2000 issue of *Resonance* we dealt with the role of Poincaré in the evolution of the special theory of relativity. Although Poincaré and other mathematicians and physicists came very close to formulating a theory of relativity, it was Einstein who took the decisive step of giving up the 'ether' and constructing a truly relativistic theory. Here we will explore in some detail how Einstein evolved this theory and get a glimpse of his unique style of thinking.

A New Point of View

Einstein formulated the special theory of relativity during his tenure at the Swiss Patent Office at Berne. In the words of Martin Klein: "In his spare time during those years at Berne, the young patent examiner wrought a series of scientific miracles; no weaker word is adequate. He did nothing less than to lay the main lines along which twentieth-century theoretical physics has developed." Let us try to trace the history of one of these 'scientific miracles' – the special theory of relativity.

In his famous 1905 paper 'On the Electrodynamics of Moving Bodies' Einstein wrote, " ... no properties of observed facts correspond to a concept of absolute rest... for all coordinate systems for which the mechanical equations hold, the equivalent electrodynamical and optical equations hold also... In the following we make these assumptions (which we shall subsequently call the Principle of Relativity) and introduce the further assumption – an assumption which is at the first sight quite irreconcilable with the former one – that light is propagated in vacant space with a velocity c which is independent of the nature of motion of the emitting body. These two assumptions are quite sufficient to give us a simple and consistent theory of electrodynamics of moving bodies on the basis of the Maxwellian theory for bodies at rest."

Even Poincaré who came very close to formulating a principle of relativity was handicapped by the popular bias of retaining the ether.

In 1887 A A Michelson and E W Morley designed an optical interferometric experiment to test the ether hypothesis. An ether hypothesis predicted a shift in the interferometric fringes in such a set up. However, when Michelson and Morley performed the experiment they did not observe *any* fringe shift. This *null result* came as a big surprise to physicists who believed in the existence of the ether. Fitzgerald and independently Lorentz tried to explain the null result of A A Michelson and E W Morley by proposing certain hypotheses within the framework of an ether hypothesis. At that time, Einstein was working in isolation with essentially no contact with other contemporary great thinkers. This isolation played a very important role in shaping his independent style of thinking. Even Poincaré who came very close to formulating a principle of relativity was handicapped by the popular bias of retaining the ether. Einstein, being far removed from the mainstream, approached the problem differently. He made a small number of general postulates and built his theory on them. In contrast, Lorentz made a large number of *ad hoc hypotheses* to arrive at his transformation equations. Einstein arrived at the same transformation equations from a much more elegant point of view. He made the following two postulates:

"Invariance of the velocity of light was, however, in conflict with the rule of addition of velocities we knew of well in mechanics."

1. *The Principle of Relativity*: The laws of physics are the same in all inertial frames. No preferred inertial frame exists.

2. *The Principle of the Constancy of the Speed of Light*: The speed of light in free space has the same value c in all inertial frames.

Then he *derived* the transformation equations from these two basic postulates. The Lorentz transformations are derived, in a non-standard way, in *Box* 1. (Though light played an important role in Einstein's thinking about relativity it is possible to demphasize this role and give up the second postulate. One can argue that logically there are two possibilities: Either there is a limiting speed at which signals can propagate or there isn't. If there isn't such a limiting speed the first postulate leads to Galilean relativity. If there is such a limiting speed we are led to the special theory of relativity. See reference 5 for discussion of this point.)

This desire to search for an elegant and simple approach perhaps helped him see clearer and deeper than any of his contemporaries. In Einstein's words: "In my own development, Michelson's result has not had a considerable influence. I even do not remember if I knew of it at all when I wrote my first paper on the subject (1905). The explanation is that I was, for general reasons, firmly convinced that there does not exist absolute motion and my problem was only how this could be reconciled with our knowledge of electrodynamics. One can therefore understand why in my personal struggle Michelson's experiment played no role, or at least no decisive role."

Einstein's line of reasoning can be understood from his Kyoto address of December 1922: "At any rate, at that time I felt certain of the truth of the Maxwell–Lorentz equations in electrodynamics. All the more, it showed to us the relations of the so-called invariance of the velocity of light that those equations should hold also in the moving frame of reference. This invariance of the velocity of light was, however, in conflict with the rule of addition of velocities we knew of well in mechanics."

Box 1. The Lorentz Transformations using the Doppler Shift

The Lorentz Transformations mentioned in this article can be derived in a number of different ways. College students will be familiar with the usual textbook derivations. Here we outline an approach due to Hermann Bondi that starts with the idea of the Doppler shift.

Imagine an inertial frame F with two observers Radhakrishnaswami and Swapnasundari (hereafter R and S) stationary with respect to each other. R is at the origin of the coordinate system and S is at $(x, 0, 0)$. At a time $t = 0$, Tarundeep (hereafter T) crosses R and moves with velocity v along the x axis in the direction of S. T's frame is denoted F', and in this frame S's position is $(x', 0, 0)$.

To understand the Doppler relation between F and F', imagine a monochromatic light wave being emitted in F. Suppose that the interval between two successive crests of the wave is τ in F. In F' the interval between the same two crests is τ'. A space-time diagram will show that τ' is longer that τ. Let us set $D = \tau'/\tau$. Because of the symmetry between F and F', it is clear that if the light wave were emitted in F' instead, then D would be τ/τ'.

Let us now go back to R, S, and T. At $t = t_1$, shortly after $t = 0$, when T crosses R, R sends a light beam to S. This beam crosses T at a time that he records as t_1'. The beam goes through to S, and she immediately reflects it towards R. As the beam is on its way back, T intercepts it, at t_2' (in F'), but instantaneously re-emits it in its original trajectory towards R, who receives it at t_2 (in F).

Because of the Doppler relations between intervals τ and τ' measured in F and F', we know that $t_1' = D\, t_1$, but $t_2 = D\, t_2'$. According to R, S is at $x = c\,(t_2 - t_1)/2$, and the time at which she reflected the beam is $t = (t_1 + t_2)/2$. According to T, however, S's position is $x' = c\,(t_2' - t_1')/2$ and the time at which she reflected the beam is $t' = (t_1' + t_2')/2$. (Once again, a space-time diagram will make things clear.) Substituting the Doppler relations in the space-time coordinate equations, we get

$$x' = \frac{D^2+1}{2D}\left[x - \frac{D^2-1}{D^2+1}ct\right]$$

and

$$t' = \frac{D^2+1}{2D}\left[t - \frac{D^2-1}{D^2+1}\frac{x}{c}\right]$$

To determine D in terms of v we notice that for $x' = 0$, $x = vt$. Thus $D = \sqrt{(1+v/c)/(1-v/c)}$. Inserting this expression for D in the equations for x' and t' we get the one-dimensional Lorentz transformations.

Question: How do the two postulates of Einstein come into this derivation of the Lorentz transformations?

Einstein abandoned many fruitless attempts, 'until at last it came to me that time was suspect!'

The resolution of this conflict came from questioning the nature of time. This fundamental issue of time had bothered Einstein as a 16 year old boy. A clearer understanding of the nature of time led him to formulate the special theory of relativity. In the words of R S Shankland: "I asked Professor Einstein how long he had worked on the special theory of relativity before 1905. He told me that he had started at age 16 and worked for ten years; first as a student when, of course, he could spend only part-time on it, but the problem was always with him. He abandoned many fruitless attempts, 'until at last it came to me that time was suspect!' ". He questioned the existence of an universal time which is the same for all observers. We will all agree that if an observer on a moving train sees two events occur at the same railway compartment at different times, another observer standing on the railway platform will see them occur at different places. However, from everyday experience we believe and assume that if an observer sees two events occurring at the same time at different places, all observers would agree that these events are simultaneous. Einstein questioned the validity of the second statement. He believed that, just as events occurring at the same spatial location for one observer may appear to be taking place at different locations for another observer, events which are simultaneous according to an observer may appear to be occurring at distinct times from another observer's point of view. This enabled him to put space and time on an equal footing and eventually arrive at the special theory of relativity. In an earlier article (February 2000, *Resonance*) we had discussed how Poincaré had questioned the issue of absolutism of simultaneity. He, however, did not quite arrive at a perfect understanding of the issue. He had used words like 'fixed' and 'moving' which are inconsistent with the true spirit of relativity. He did not give up the 'ether'. Einstein, in contrast, arrived at a complete understanding of the issue of relativity of simultaneity by approaching the problem from a fresh perspective. Both Poincaré and Einstein independently noticed that two successive Lorentz transformations lead to another Lorentz transformation and this observation led them to realise that the Lorentz transformations (here we use the term to include spatial rotations) form a group (see *Box* 2 for an interesting observation).

Consequences of the Theory of Relativity

There are several consequences that follow from Einstein's two basic postulates. Some of these are :

a. The Lorentz transformations which we have already discussed.

b. From the Lorentz transformations he was led to Fitzgerald–Lorentz contraction of lengths and the dilation of time: $\gamma l = l_0$ and $t = \gamma t_0$ where l_0 and t_0 are, respectively, a length and a duration of time in the rest frame. $\gamma = 1/\sqrt{1 - v^2/c^2}$, where v is the speed of the moving frame relative to the rest frame and c is the speed of light in vacuum.

c. The composition rule for velocities: Two successive Lorentz transformations with velocities v_1 and v_2 in the same direction result in a new Lorentz transformation with a velocity given by $v = (v_1 + v_2)/(1 + v_1 v_2/c^2)$.

d. The relativistic expression for the angle α of aberration of starlight coming from the zenith: $\tan\alpha = \gamma v/c$ where v is the speed of earth relative to the star.

e. An immediate consequence of the composition rule for velocities formula is an explanation of the 'Fizeau effect'. In 1817 Fresnel had predicted that if a nondispersive liquid moves through a tube with a velocity v relative to the ether and if a light beam traverses the tube in the same direction, then the light velocity c' in the laboratory is given by $c' = c/n + v(1 - 1/n^2)$ where n is the refractive index of the liquid. This effect was experimentally confirmed by Fizeau in 1851. One can easily check that one gets the Fresnel formula for the speed of light using the special relativistic law of addition of velocities. This result was however not mentioned in Einstein's June 1905 paper.

f. The relativistic equation for the Doppler effect: $\nu' = \gamma\nu(1 - v\cos\phi/c)$ where ϕ is the angle between a monochromatic light

To this day not a single experimental result has been found to be inconsistent with Einstein's special theory of relativity.

ray with frequency ν and the x direction. ν' is the modified frequency due to Doppler effect. Notice that for $\phi = \pi/2$ classically there is no Doppler effect. However, in the relativistic regime there is a modification of the frequency even for $\phi = \pi/2$. This inherently relativistic feature of the Doppler effect, the transverse Doppler effect, was Einstein's discovery.

g. He mentioned that if two synchronous clocks C_1 and C_2 are at the same initial position A and if C_2 leaves it and moves along a closed orbit, then upon return to A, C_2 will run slow relative to C_1. This effect is sometimes called the 'clock paradox' or the 'twin paradox'. There is no 'paradox' here however: the effect is a logical consequence of special relativity and is seen in the lifetimes of particles received on earth from cosmic rays.

h. He checked the invariance of the Maxwell–Lorentz equations under a Lorentz transformation and in the process obtained the expression for the Lorentz force which Lorentz had simply introduced as a new assumption.

i. He gave an expression for the kinetic energy W of a particle moving with a speed v: $W = mc^2(\gamma - 1)$, which led him to conclude that one cannot have a velocity greater than that of light since W becomes infinite for $v = c$.

j. The measured mass of a body depends on its velocity relative to the observer. This follows from the relativistic invariance of the classical form of the momentum conservation law. Thus, m, the measured mass and m_0, the rest mass of a body, are related via $m = m_0 \gamma$.

k. The June 1905 paper also contained the law of transformation for the energy E of a light beam: $E' = \gamma E\,(1 - v \cos\phi/c)$. It is a bit surprising that Einstein who was aware of the connection $E = h\nu$ between the energy and the frequency of a light quantum did not see the energy transformation law as following from the Doppler shift formula. He makes the following comment: 'It is remarkable that the energy and the frequency of a light complex vary with the state of motion of the observer in accordance with the same law'.

This energy transformation law of his June 1905 paper played a crucial role in his derivation of the celebrated mass-energy equivalence relation $E = mc^2$. This relation appeared in his September 1905 paper. The equivalence of mass and energy had been known for special cases before his September 1905 paper. For instance, Fritz Hasenohrl, the Austrian physicist had discovered that the kinetic energy of a cavity increases when it is filled with radiation, in such a way that the mass of the system appears to increase. The novelty of Einstein's 1905 paper lay in the generality of this relation. In Einstein's words: "If a body gives off the energy E in the form of radiation, its mass diminishes by E/c^2. The fact that the energy withdrawn from the body becomes energy of radiation evidently makes no difference, so that we are led to the more general conclusion that the mass of a body is a measure of its energy content .... It is not impossible that with bodies whose energy-content is variable to a high degree (e.g., with radium salts) the theory may be successfully put to the test. If the theory corresponds to the facts, radiation conveys inertia between the emitting and absorbing bodies." Experiments have indeed confirmed Einstein's predictions.

Einstein's special theory of relativity with its modified notion of space and time, led directly to Minkowski's mathematical framework for relativistic kinematics in terms of four-dimensional space-time.

Einstein's Approach – Simple Yet Profound

Einstein's approach was bold, simple and general. The success of any theory lies in its predictive power and in its ability to explain the results of existing experiments. In fact, Einstein's special theory of relativity was not only able to explain existing experiments, it led to many predictions which were later confirmed by experiments. To this day not a single experimental result has been found to be inconsistent with Einstein's special theory of relativity. Just to get a feel for how successful this theory has been one can compare the results of thirteen crucial experiments (see page 29 of [1]) *all* of which agree with Einstein's special theory of relativity. A considerable fraction of these experiments actually disagree with other competing theories (ether and emission theories). The overwhelming experimental success of the special theory of relativity sets it apart from all other competing theories. On the theoretical end, Einstein's special

Box 2. Thomas Precession

In 1927, after special relativity had been subjected to over two decades of scrutiny, Llewellyn Thomas discovered a consequence of the theory that astounded everyone, including Einstein. George Uhlenbeck and Samuel Goudsmit had proposed a theory of the spin of the electron in 1925. The interaction between the magnetic moment of the electron, which is inseparably linked with its spin, and the electric field of the nucleus produces an interaction energy that was able to account for the splitting of doublets in the spectra of certain atoms – only the answer was off by a factor of two. Thomas provided this factor as follows:

An electron moving around a nucleus changes Lorentz frames as it circles the nucleus. This sequence of frame changes results in an overall rotation of the electron spin when the electron returns to its original Lorentz frame. This rotation is termed Thomas precession and it needs to be taken into account to get the observed rate of precession of the electron, or equivalently, the observed splitting of doublets in atomic spectra. (Please note that Thomas precession is a relativistic kinematic effect and it has nothing to do with the quantum mechanical scale of the atom. A classical top behaves exactly the same way.)

(See [6] for a detailed discussion of Thomas precession and its observational consequences.)

theory of relativity with its modified notion of space and time, led directly to Minkowski's mathematical framework for relativistic kinematics in terms of four-dimensional space-time. This paved the way for Einstein's general theory of relativity. Einstein initially dismissed Minkowski's mathematical work as 'superfluous learnedness'. However, he later realised how Minkowski's approach greatly facilitated the transition from special to general relativity.

In retrospect, Einstein's approach appears to be the most obvious one to follow as a way of resolving the incompatibility between electrodynamics and Newtonian relativity. Let's end this article with Herman Bondi's statement: "The special theory of relativity is a necessary consequence of any assertion that the unity of physics is essential, for it would be intolerable for all inertial systems to be equivalent from a dynamical point of view yet distinguishable by optical measurements. It now seems almost incredible that the possibility of such a discrimination was taken for granted in the nineteenth century, but at the same time it was not easy to see what was more important – the universal validity of the Newtonian principle of relativity or the absolute nature of time."

Suggested Reading

[1] Robert Resnick, *Relativity and Early Quantum Theory*, John Wiley and Sons, New York, 1972.

[2] Abraham Pais, 'Subtle is The Lord...' *The Science and the Life of Albert Einstein*, Oxford University Press, 1982.

[3] Joseph Samuel, 'Torches, Clocks and Lorentz Transformation', Physics Education, Vol.11, p.402, 1995.

[4] J D Jackson, *Classical Electrodynamics*, Wiley Eastern Limited, New Delhi, 2nd Edn., 1978.

[5] Wolfgang Rindler, *Introduction to Special Relativity*, Clarendon Press, Oxford, 1982.

[6] Yuan Zhong Zhang, *Special Relativity and its Experimental Foundations*, World Scientific, Singapore, 1997.

Address for Correspondence
Supurna Sinha
307, Sampige Road
Malleswaram
Bangalore 560003, India.