Derivation of special relativity from Maxwell and Newton

BY D. J. DUNSTAN*

Department of Physics, Queen Mary, University of London, London E1 4NS, UK

Special relativity derives directly from the principle of relativity and from Newton’s laws of motion with a single undetermined parameter, which is found from Faraday’s and Ampère’s experimental work and from Maxwell’s own introduction of the displacement current to be the $-c^{-2}$ term in the Lorentz transformations. The axiom of the constancy of the speed of light is quite unnecessary. The behaviour and the mechanism of the propagation of light are not at the foundations of special relativity.

Keywords: special relativity; Newton’s laws; Maxwell’s equations; principle of relativity; principle of constancy of speed of light

1. Introduction

The standard derivation of special relativity in most elementary textbooks and lecture courses starts from Einstein’s two postulates, the principle of relativity and the principle of the constancy of the speed of light. The principle of relativity is attributed to Galileo and Newton, and Newton’s first law of motion is a statement of it. The constancy of the speed of light is based on the Michelson–Morley experiment. This approach gives hostages unnecessarily to the cranks who regularly refute relativity by contesting the outcome of the Michelson–Morley experiment or its interpretation, and to the philosophers of science who follow Popper (1959) and Kuhn (1962) to assert that advances in science are little more than a refutation of earlier science and, therefore, that there is no trustworthy or reliable knowledge in science (Miller 1999).

In teaching and in elementary textbooks, it would be preferable to derive special relativity directly from the principle of relativity, Newton’s laws of motion and Maxwell’s equations. We show here how a selection of standard results can be used to do this, letting it be clearly seen how Einstein’s results build upon the earlier work, rather than overthrowing it. It shows how Maxwell’s work in systematizing the results of Ampère and Faraday, and in particular his genius in introducing the displacement current to complete Ampère’s result, is central to special relativity.

*d.dunstan@qmul.ac.uk

Electronic supplementary material is available at http://dx.doi.org/10.1098/rsta.2007.2195 or via http://journals.royalsociety.org.

One contribution of 20 to a Theme Issue ‘James Clerk Maxwell 150 years on’.
2. Principle of relativity

The principle may be stated in the form that experimental results and the equations describing them must be independent of the speed of the inertial frame in which they are measured. This principle may be applied to the co-ordinate transform, from inertial frame $\Sigma$ to frame $\Sigma'$ at velocity $v$ in $\Sigma$, which must take the same form as the back transform from $\Sigma'$ to $\Sigma$ with velocity $-v$ (Lee & Kalotas 1975; Lévy-Leblond 1976). The transform must be linear if linear equations in one frame are to transform into linear equations in the other frame. Neglecting the $y$ and $z$ co-ordinates, we may write the most general linear transform as

$$\begin{pmatrix} x' \\ t' \end{pmatrix} = \begin{pmatrix} a & b \\ \alpha & \beta \end{pmatrix} \begin{pmatrix} x \\ t \end{pmatrix}. \quad (2.1)$$

The back transformation is obtained by inverting equation (2.1),

$$\begin{pmatrix} x \\ t \end{pmatrix} = \frac{1}{a\beta - \alpha b} \begin{pmatrix} \beta & -b \\ -\alpha & a \end{pmatrix} \begin{pmatrix} x' \\ t' \end{pmatrix}. \quad (2.2)$$

The velocity of $\Sigma'$ in $\Sigma$ is the velocity $v$ of its origin with $x' = 0 = ax + bt$, giving its speed $v = -b/a$, or $b = -av$. The principle of relativity requires that the velocity of $\Sigma$ in $\Sigma'$ is $v' = -v = b/\beta$, so that $\beta = -b/v = a$. Making these substitutions, the transforms are

$$\begin{pmatrix} x' \\ t' \end{pmatrix} = \begin{pmatrix} a & -av \\ \alpha & a \end{pmatrix} \begin{pmatrix} x \\ t \end{pmatrix},$$

$$\begin{pmatrix} x \\ t \end{pmatrix} = \frac{1}{a^2 - \alpha av'} \begin{pmatrix} a & av' \\ -\alpha & a \end{pmatrix} \begin{pmatrix} x' \\ t' \end{pmatrix}. \quad (2.3)$$

Requiring that the coefficients in the forward and back transforms are the same requires that

$$a^2 - \alpha av' = 1, \quad \alpha = \frac{1 - a^2}{-av'} = \frac{1 - a^2}{av}. \quad (2.4)$$

yielding the general transform

$$\begin{pmatrix} x' \\ t' \end{pmatrix} = \begin{pmatrix} x - avt \\ \frac{1 - a^2}{av} x + at \end{pmatrix}. \quad (2.5)$$

as the only linear transform consistent with the principle of relativity, with the undetermined coefficient $a$. The Galilean transform has $a=1$.

Note that $y' = \kappa y$, $z' = \kappa z$, with $\kappa = 1$. Other solutions are possible (Rindler 1977), but these may all be ruled out by requiring the limiting case of the transform for $v=0$ to be $x' = x$, $y' = y$, $z' = z$, $t' = t$. 

Phil. Trans. R. Soc. A (2008)
3. Newton’s laws

Newton’s second and third laws of motion permit a tighter constraint on the coefficient \(a\) and the functional form of its dependence on \(v\). The second law defines momentum and the third law states the conservation of momentum. We need to analyse a problem such as a collision in which particles have different speeds \(u\) in a convenient frame \(\Sigma\). Most elementary textbooks give the analysis of a suitable problem (Rosser 1967). Two identical particles \(P_1\) and \(P_2\) undergo an elastic collision at the origin of two frames. In \(\Sigma\), \(P_2\) travels down the \(y\)-axis \((x_2=z_2=0)\) at speed \(u_2\) before the collision and up it at \(u_2\) afterwards. \(P_1\) travels in the \(x-y\) plane \((z_1=0)\) at speed \(u_1\). In \(\Sigma'\) the centre of mass is stationary. By symmetry, in \(\Sigma'\) masses \(m'_1 = m'_2\) and speeds \(u'_1\) and \(u'_2\) are equal and unchanged in the collision. Conserving the momenta \(mu\) resolved in the \(x\) and \(y\) directions in both frames gives after some algebra the noteworthy equation (Rosser 1967; see also electronic supplementary material),

\[
m_1 \sqrt{u_1^2 + \frac{a^2 v^2}{(1-a^2)}} = m_2 \sqrt{u_2^2 + \frac{a^2 v^2}{(1-a^2)}}. \tag{3.1}
\]

This requires that the masses \(m_1\) and \(m_2\) in \(\Sigma\) vary with speeds \(u\) in \(\Sigma\). It may be worth noting that we need not—indeed should not—introduce any philosophical assumptions about what the \(m(u)\) is, other than that it is the quantity that an observer in any frame will find appropriate to use in his/her version of Newton’s second law. Since the mass of each particle in \(\Sigma\) can be a function of its speed in \(\Sigma\) only, and cannot be a function of \(v\) or of the speed of the other particle, equation (3.1) is separable, i.e. each side has to be separately equal to a constant. Then the term in \(a\) has to be a constant, so we may write it as

\[
\frac{1}{k} = \frac{a^2 v^2}{1-a^2} \quad \text{or} \quad a = \frac{1}{\sqrt{1 + kv^2}} \tag{3.2}
\]

and then

\[
\begin{pmatrix} x' \\ t' \end{pmatrix} = \frac{1}{\sqrt{1 + k v^2}} \begin{pmatrix} 1 & -v \\ k v & 1 \end{pmatrix} \begin{pmatrix} x \\ t \end{pmatrix}. \tag{3.3}
\]

This may be recognized as the Galilean transform for \(k=0\) and the Lorentz transform for \(k=-c^{-2}\).

4. Maxwell’s equations

Two of Maxwell’s equations in free space (with no charges or currents) are

\[
\partial_x E_y = \mu_0 \partial_t H_z, \tag{4.1}
\]

\[
\partial_x H_z = \varepsilon_0 \partial_t E_y. \tag{4.2}
\]

These come from Faraday’s law of magnetic induction (an experimental result) and Ampère’s law (experimental) as modified by Maxwell’s introduction of the displacement current on theoretical grounds and are obtained from macroscopic
experiments carried out at rest in $\Sigma$. We need not consider the modern account in terms of the transforms of the Lorentz force on moving charges. More simply, if the same macroscopic experiments are carried out at rest in $\Sigma'$, the principle of relativity requires that the same expressions are obtained, with all variables primed and with the same numerical values for the constants,

$$\partial_x' E_y' = \mu_0 \partial_t' H_z', \quad (4.3)$$

$$\partial_x' H_z' = \varepsilon_0 \partial_t' E_y'. \quad (4.4)$$

Using the chain rule to transform the partial differential operators into the frame $\Sigma$, we obtain

$$\partial_x' = a\partial_x - \frac{1 - a^2}{av} \partial_t \quad \partial_t' = av\partial_x + a\partial_t. \quad (4.5)$$

Making these substitutions in equation (4.3) and rearranging, we get

$$\partial_x (aE_y' - av\mu_0 H_z') = \mu_0 \partial_t \left( aH_z' + \frac{1 - a^2}{av\mu_0} E_y' \right), \quad (4.6)$$

and similarly from equation (4.4)

$$\partial_x (aH_z' - av\varepsilon_0 E_y') = \varepsilon_0 \partial_t \left( aE_y' + \frac{1 - a^2}{av\varepsilon_0} H_z' \right). \quad (4.7)$$

These are the predictions from the primed frame of what equations will be found to hold in the unprimed frame, and so the principle of relativity permits us to identify the terms in brackets with $E_y$ and $H_z$. That is, comparing equation (4.6) with equation (4.1), and equation (4.7) with (4.2),

$$E_y = aE_y' - av\mu_0 H_z' = aE_y' + \frac{1 - a^2}{av\varepsilon_0} H_z', \quad (4.8)$$

$$H_z = aH_z' + \frac{1 - a^2}{av\mu_0} E_y' = aH_z' - av\varepsilon_0 E_y'. \quad (4.8)$$

Just as we do not require to interpret $m(u)$ in equation (3.1), we do not require to interpret the expressions of equation (4.8) in terms of the transforms of electric and magnetic fields. All we require is the equalities of the second and third terms in each line, so that from the first line of equation (4.8) we have

$$-av\mu_0 = \frac{1 - a^2}{av\varepsilon_0}, \quad (4.9)$$

while from the second line of equation (4.8) we have

$$\frac{1 - a^2}{av\mu_0} = -av\varepsilon_0. \quad (4.10)$$

Solving the expressions of either equation (4.9) or (4.10) for $a$ gives

$$a = \frac{1}{\sqrt{1 - \varepsilon_0\mu_0 v^2}} \quad \text{or} \quad k = -\varepsilon_0\mu_0. \quad (4.11)$$

Having derived the wave equation, Maxwell identified $\varepsilon_0\mu_0$ as $c^{-2}$, so that $k = -c^{-2}$, which yields the standard Lorentz transform when put into the transform of §2. But invoking the speed of light is unnecessary. Equation (4.11) is
already sufficient to quantify $k$ and $a$, and therefore by substitution into equation (2.5) and equation (3.1) to yield the entire theory of special relativity, quantitatively, without any mention of the speed of light.

5. Conclusions

Special relativity derives directly from the principle of relativity and from Newton’s laws of motion. The parameter values of $a=1$ or $k=0$ were compatible with all experimental information available in Newton’s day. However, Maxwell’s equations permit a more accurate determination, from Faraday’s and Ampère’s experimental work and Maxwell’s own introduction of the displacement current. Discussions of the Michelson–Morley experiment and of theories of the ether are quite unnecessary. The behaviour and the mechanism of the propagation of light are not at the foundations of special relativity.

This approach, if used in undergraduate teaching and in elementary textbooks, would significantly help to confound both the cranks who seek to reinstate the ether and the philosophers who seek a paradigm shift in which Newton’s edifice is falsified and overthrown. On the contrary, Einstein introduced no paradigm shift or falsification into physics. His genius, rather, lay in the recognition that the $x^0$ and $t^0$ of equation (2.1) are the space and time coordinates in the frame $S^0$, a paradigm expansion, a realization that had escaped Lorentz, Fitzgerald and even Poincaré.

References