TRANSFORMATION OF CO-ORDINATES IN SPECIAL RELATIVITY

A. H. BRIDLE

Trinity College,
Cambridge.

August 1962

TRANSFORMATION OF CO-ORDINATES IN SPECIAL RELATIVITY

Summary A rigorous derivation of the Lorentz transformation equations of special relativity is given, using a matrix notation to take full advantage of the assumption of Euclidean geometry and the consequent linearity of the transformation.

An inertial frame is a reference frame in which spatial relations, as would be determined by a system of rigid rods at rest in the frame, are Euclidean, and in which a test particle, once found to be at rest, would remain so for all time. In such a frame the motion of a free particle would be uniform and linear.

Einstein's Principle of (Special) Relativity asserts that a frame in uniform translatory motion relative to an inertial frame cannot be distinguished from that inertial frame by any physical experiment whatsoever. It follows from this that such a frame is itself inertial, and that any two inertial frames which are not trivially coincident must be in uniform translatory motion. The principle implies that all inertial frames are equivalent for the formulation of physical laws, and that in particular the law of light-propagation must be the same for all observers in inertial frames. The work of de Sitter on double star systems has shown that the velocity of light is independent of the velocity of the source, and on the hypothesis of isotropy of an inertial frame, it would be independent of direction.

The stress laid by these ideas on the importance of inertial frames is strong criticism of the special theory, especially as inertial frames can only be regarded as an extrapolation of limited experience. All practical frames of reference contain gravitating matter by their very nature, and the uniform linear motion of test particles cannot be expected to be observed. The Euclidean geometry is also an idealisation of our local observations; these difficulties, together with the unique importance of inertial frames, are removed by the General Relativity theory.

With these definitions and postulates, however, a system of transformations between frames can be set up which provides a satisfactory account of many relationships observed in practice. The extent to which a reference frame can be regarded as inertial depends on the particular application in view, and must ultimately be decided by experiment.

The postulated agreement of all observers on the law of light-propagation is the starting-point for the derivation of these transformations, which will now be undertaken.

Consider an event P in space and time at which a light signal is emitted. An event Q, at which the light signal is absorbed is considered to take place at a nearby point in space and at a neighbouring instant in time. In an inertial frame S, let the event P be assigned co-ordinates (x,y,z,t) and the event Q co-ordinates (x+dx,y+dy,z+dz, t+dt). Similarly, in another inertial frame S', let P be (x',y',z',t') and Q (x'+dx', y'+dy',z'+dz',t'+dt'). From the law of light-propagation and the Euclidean geometry assumed we may put

$$dx^2 + dy^2 + dz^2 = c^2dt^2 \quad \text{and} \quad dx'^2 + dy'^2 + dz'^2 = c^2dt'^2$$

The assumed agreement between inertial observers as regards the value of c justifies the use of the same symbol in the two frames.

Now these relations may be written in the form

$$dx^2 + dy^2 + dz^2 - c^2dt^2 = 0, \quad dx'^2 - dy'^2 - dz'^2 - c^2dt'^2 = 0.$$

It is known from algebra that any two polynomials satisfying the condition that they share all their zeros, if they are of the same degree, can only differ by a constant factor at most. It follows that, for <u>any</u> event Q in the neighbourhood of P, the following relation is true :

$$dx^2 + dy^2 + dz^2 - c^2dt^2 = K(dx'^2 + dy'^2 + dz'^2 - c^2dt'^2), \text{ K being a constant.}$$

Now, because we have taken Euclidean geometry, the space-interval (distance) between P an Q is independent of the position in space of P or Q relative to the origin (i.e. the term $(dx^2 + dy^2 + dz^2)$ is independent of (x,y,z) for given P and Q, and similarly for the primed system, and similarly because of the assumptions regarding time-dependence, the time-intervals dt, dt' are independent of t,t'. Therefore, there is no loss of generality in taking P to be the event (0,0,0,0) in both S and S'. The relationship of S and S' is therefore seen to be completely symmetric, both with respect to P and with respect to each other. It is therefore equally valid to write

$$K(dx^2 + dy^2 + dz^2 - c^2dt^2) = dx'^2 + dy'^2 + dz'^2 - c^2dt'^2$$

It follows therefore that K must equal ∓1. Now in the limit of S and S' becoming coincident for all time (vanishing relative velocity) dx must tend to + dx', etc., so that only K = +1 is acceptable. It follows that transformations of infinitesimally small intervals between S and S' must satisfy

$$dx^2 + dy^2 + dz^2 - c^2dt^2 = dx'^2 + dy'^2 + dz'^2 - c^2dt'^2 \qquad (1)$$

The form of this relation between differentials implies that the law of transformation between S and S' for general events (i.e. not necessarily infinitesimally close events) is linear. The linearity of the transformation also means that the "standard configuration" of two inertial frames commonly used in Special Relativity (viz. one frame in uniform motion along a prescribed space-axis in the other) is a justifiable simplification, in that the spatial co-ordinate systems in any two inertial frames can be oriented to satisfy this condition at all times, whatever the original direction of their relative uniform translatory motion. It is therefore sufficient to derive a transformation law by considering the particular case of "standard configuration". Thus the deduction of linearity of the transformation equations very greatly simplifies the problem.

(The general result, proved in any book on tensor calculus, is that a transformation which takes a metric

$$g_{\lambda\mu}dx^\lambda dx^\mu$$

with constant coefficients into another metric

$$g_{\lambda'\mu'}dx^{\lambda'}dx^{\mu'}$$

also with constant coefficients, must be a linear transformation. The present example is clearly the particular case of this theorem with $g_{11} = g_{22} = g_{33} = 1$, g_{44}(say) $= -c^2$)

Under a linear transformation, the finite co-ordinate differences satisfy the same transformation law as the differentials, so we shall have

$$x^2 + y^2 + z^2 - c^2t^2 = x'^2 + y'^2 + z'^2 - c^2t'^2 \qquad (2)$$

With the linearity of the transformation in mind, the use of a vector representation of events with a matrix transformation law is suggested. In the formalism to be used henceforth, events will be represented by vectors in a mathematical 4-space (a space of four dimensions), such that the event (x,y,z,t) is for example represented by the vector whose components are ict, x, y, and z. The convenience of this particular representation will become apparent shortly. In the matrix notation, the vector will be written as the column:

$$X = \begin{pmatrix} ict \\ x \\ y \\ z \end{pmatrix} \text{, similarly for } X'$$

The transformation law between two inertial frames S and S' moving in standard configuration will take the form $X' = (L)X$, where (L) is a 4 x 4 matrix. Denoting the transpose of a matrix as $\bar{X}$, $(\bar{L})$, etc., the equations of the transformation may be written :

$X' = (L)X, \quad \overline{X}' = \overline{X}(\overline{L}), \quad$ and $\quad \overline{X}'X' = \overline{X}X$ (from equation (2))

Substituting for X' and $\overline{X}'$ in the last of these from the first two,

$$\overline{X}(\overline{L})(L)X = \overline{X}X \text{ , so that } (\overline{L})(L) = 1$$

It is therefore necessary that (L) be an orthogonal 4 x 4 matrix in order that (2) shall be true.

As we are dealing with motion in the standard configuration, $y,z = 0$ imply $y',z' = 0$ for all time. It follows that $y' = Ay$, $z' = Bz$ for all time, where A and B are constants. From the symmetry between inertial frames discussed above, we may deduce that A and B can only take the values ∓ 1 as before, and similarly consideration of the limit as S tends to S' forces the choice of +1. The matrim (L) must therefore take the form:

$$(L) = \begin{pmatrix} L_{11} & L_{12} & 0 & 0 \\ L_{21} & L_{22} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Now the only possible orthogonal 4 x 4 matrix of this form must have $L_{11} = \cos\lambda = L_{22}$, $L_{12} = \sin\lambda$, $L_{21} = -\sin\lambda$, where λ is some angle.

In that case, $\quad x' = ict(-\sin\lambda) + x(\cos\lambda)$

$x' = 0$ must imply $x = vt$, where v is the relative velocity of S' and S, from the definition of the standard configuration, to we must have

$$\tan\lambda = v/ic$$

Therefore $\cos\lambda = \beta$, $\sin\lambda = -iv\beta/c$, where $\beta = (1 - v^2/c^2)^{-\frac{1}{2}}$

The transformation matrix for the special theory of relativity is therefore

$$(L) = \begin{pmatrix} \beta & -iv\beta/c & 0 & 0 \\ iv\beta/c & \beta & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

It may readily be verified that the matrix equation $X' = (L)X$ now contains the two more common equations $x' = \beta(x - vt)$ and $t' = \beta(t - vx/c^2)$.

Discussion.

The assumption of Euclidean geometry ensures the complete symmetry between two inertial frames S and S'. This symmetry is reflected in the fact that the inverse of the final matrix obtained is the same matrix with $-v$ written for v.

It also ensures that the transformation is a linear transformation, thereby permitting the matrix representation which has been used. In the mathematical 4-space in which an event is represented by a vector X, the transformation is seen to be a rotation of the vector in the (complex) x-t plane through a (complex) angle $\lambda = \arctan v/ic$.

It is not usually stressed sufficiently that it is the linearity of the transformation which permits the use of "standard configuration", so that nothing essential to the theory is lost by making the greatly simplifying step of taking S and S' to be in this configuration. Because of this linearity, straight lines in S must become straight lines in S', where they may be moving. By adjusting their direction however, we can always find a set of lines mutually parallel and fixed in S which transform into a similar set in S'. This makes possible the choice of the common x-axis. Further, fixed planes containing the x-axis in S become similar planes in S', and by symmetry two such planes at right angles in S become planes at right angles in S'. This makes possible the choice of common co-ordinate planes.

Because the transformation equations may be written in terms of an orthogonal matrix specifying a rotation (in the mathematical 4-space, it must be emphasised), it follows that the transformations have a group property. In particular, the product of any two transformations will also be a transformation of the same form. Thus, if S' moves in standard configuration at velocity v relative to S, the matrix of transformation between S and S' is L(v). Similarly, if S" moves at u relative to S', the matrix is L(u). Then the transformation between S and S" is, by the group property, given by L(v) x L(u), and will be of the form L(w), where w is the "combined velocity" of v and u. Carrying out the matrix multiplication of L(v) and L(u) shows that

$$w = \frac{u + v}{1 + \frac{uv}{c^2}}$$

The law of combination of velocities therefore follows directly from the group property.

The value of the matrix formulation of Special Relativity cannot be over-stressed. It provides the simplest structure for dealing with events in a 4-dimensional space-time, and suffices for almost all important applications of the theory. In subsequent articles, the matrix method will be applied to relativistic mechanics and to relativistic electrodynamics in order to bring out this point more clearly.

Trinity College,
Cambridge

August 1962