Alan Bridle

THE FORMULATION OF MECHANICS IN SPECIAL RELATIVITY

A. H. BRIDLE

Trinity College,
Cambridge.

September 1962

THE FORMULATION OF MECHANICS IN SPECIAL RELATIVITY

Summary The matrix formalism for special relativity will be used to develop a Lorentz invariant system of mechanics applicable to mass points. The generalisation to rigid bodies, and the approach to a canonical form will not be considered, and the applications of and experimental verification of the results obtained will only briefly be dealt with. Stress will be laid throughout on the significance of 4-vectors and their connexion with the basic principle of Special Relativity.

In a previous article ("Transformation of Co-ordinates in Special Relativity", referred to hereafter as (I)) it was shown that, with a homogeneous and isotropic definition of inertial frames based on Euclidean geometry, Einstein's Principle of Special Relativity leads to a particular rule for the transformation of space-time co-ordinates between inertial frames. In a mathematical 4-space in which space-time events are represented by vectors with components (ict, x, y, z), the transformation between two frames in standard configuration was shown to be realised in the 4 x 4 orthogonal matrix

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

which defines a rotation in the complex x-t plane through a complex "parametric" angle λ satisfying $\lambda = \arctan(v/ic)$. (See (I), P.3)

It was also shown in (I), as a preliminary to the derivation of the matrix (L), that the corresponding differentials in two inertial frames S and S' satisfy

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

We can therefore define a differential quantity ds^2 through

$$ds^2 = dx^2 + dy^2 + dz^2 - c^2dt^2$$

so that $ds^2 = ds'^2$. The differential ds is then said to be invariant under a Lorentz transformation between reference frames, or more simply, Lorentz invariant. It is called the <u>interval</u> between two events whose space-time co-ordinates differ by the differentials dx,dy,dz,dt. The importance of such Lorentz invariant quantities to the theory is paramount, as reference to the basic principle of Special Relativity will show.

It is a consequence of the Principle of Special Relativity as stated in (I) that all inertial frames are equivalent for the formulation of physical laws, by which we mean that the laws of physics must take the same form in all such frames. But if the transformation of co-ordinates from frame to frame is accomplished by the matrix (L), it follows that the laws of physics must be expressible in terms which are invariant under the transformation (L). The construction of Lorentz invariant quantities as the scalars of the theory is therefore a primary task.

Suppose we define a vector quantity v in an inertial frame S, such as we defined the vector representing an event in (I), i.e. v is a vector with four components in the mathematical 4-space in which we are developing the theory. Suppose further that v' represents the same vector quantity referred to the co-ordinate system of a second inertial frame S'. In general v' will be related to v in an arbitrary manner, but for certain types of vector, the law of transformation between v and v' will be $v' = (L)v$, i.e. the same as for the event-vectors X, X'. Any vector v having this transformation property will be termed a <u>4-vector</u> henceforth, in keeping with conventional usage (perhaps slightly unfortunate as it leaves us no simple name for vector quantities not transforming in this way). It will readily be apparent that we should also seek to formulate physical laws in terms of 4-vectors as well as invariants; suppose a physical idea was represented in a frame S by the equation $v = 0$, where v is a 4-vector. Then the Lorentz transformation of this equation would yield $(L)v = 0$. If $(L)v = v'$, we have $v' = 0$, so that the law would have the same form in any other inertial frame S' in which v' corresponded to v, i.e. the law would have the same form in all inertial frames, as the transformation between any one frame and any other can be expressed in a matrix such as (L).

Clearly, any equation involving only 4-vectors and Lorentz invariants will also preserve its form under a Lorentz transformation. Consider for example a possible physical law $dv/ds = 0$, where v is a 4-vector and ds is the invariant differential of Page 1. Under a Lorentz transformation, this law becomes $d(L)v/ds' = dv'/ds' = 0$. As $ds = ds'$, this law is of the same form in all inertial frames.

Our task is therefore that of discovering Lorentz invariants (scalars) and 4-vectors in terms of which we may write the laws of physics in the same form for all inertial frames. (It may be enquired at this point whether or not we should also concern ourselves with the next highest order in this hierarchical scheme of quantities, viz. some sort of 4-tensors. The definition of quantities with a property of tensor transformation in Lorentz form is indeed possible, but is not required for mechanics. This topic will accordingly be postponed until a later article (in preparation) on special relativistic electrodynamics).

It will be convenient in future to denote Lorentz invariants and 4-vectors by capital letters, e.g. dS, V, to distinguish them from other quantities.

Let us now attempt to construct a 4-vector analogous to the velocity vector of non-relativistic theory. By comparison with the form dx/dt of ordinary mechanics, we might seek to define the velocity as dX/dt, where X is the column with terms ict, x, y, z. This vector would therefore have components ic, dx/dt, dy/dt, dz/dt. Inspection shows that it is not a 4-vector however, for although X is a four-vector, dt is not an invariant differential, i.e. dt does not equal dt' when we transform from some S to some S'. Thus we can see the difficulty looming up that our non-relativistic laws of mechanics may need to be extensively re-cast, as this vector which has failed to have the 4-vector property contains the "ordinary" velocity as three of its components. We must look for some differential invariant dT, analogous to dt, so that we may define $V = dX/dT$ as the 4-velocity.

Consider the quantity ds^2. It may be positive, zero, or negative, according to the relations obtaining between dx,dy,dz,dt. If it is positive, dx,dy,dz,dt are such that two events separated by these differentials cannot possibly be illuminated by the same flash of light. Also, as ds is an invariant, dS, there can be no frame in which dx, dy, and dz are all zero. There is therefore an essentially space-like character to the interval dS when it is greater than zero, and such an interval is conventionally called in fact a "space-like" interval. If on the other hand, dS^2 is less than zero, there can be no frame in which $dt = 0$, and the interval dS is called a "time-like" interval. To bring out this idea, let us write

$$dS^2 = - c^2 dT^2$$

thereby defining a differential dT. As dS is an invariant, and so is c (which should strictly be written C in that case, but it will be left as c as the sole exception to the capital rule), dT is an invariant. If dS^2 is negative, so that dS is imaginary, dT^2 is positive, so that dT is real. Thus, for a time-like interval dT is a real quantity, and has a time-like property in this sense. To show the significance of dT more clearly, consider a particle moving with velocity v in an inertial frame S. Then, for that particle,

$$v^2\delta t^2 = \delta x^2 + \delta y^2 + \delta z^2$$

Therefore we can write $\delta x^2 + \delta y^2 + \delta z^2 - c^2\delta t^2 = (v^2 - c^2)\delta t^2$

In the limit $\delta x \to dx$, etc., $(v^2 - c^2)dt^2 = - c^2 dT^2$ therefore

$$dt = \beta . dT$$

Thus dT appears as the time-differential measured on a clock at rest in the rest frame of the particle, corresponding to the time-differential dt in the original frame S. dT is accordingly termed the differential of "pwoper time". The likeliest candidate for the role of 4-velocity is therefore

$$V = \frac{dX}{dT}$$

where X and dT are defined as above.

In exactly the same way, we can see that the acceleration as defined in ordinary mechanics does not transform by the Lorentz matrix, and is therefore not a 4-vector. We can however construct a vector to be called the 4-acceleration as

$$A = \frac{dV}{dT}$$

on exactly the same argument as above, i.e. that V is a 4-vector already, and that dT is an invariant differential.

Before embarking now on the main topic of this discussion, it may be as well to recapitulate on the nature of the problem in hand.

i) Given the transformation matrix (L) for co-ordinates between any two inertial frames xxx S and S', we are seeking to cast the laws of mechanics in a form to be unaltered by this transformation.

ii) In order to do this we shall make use of invariant differentials and 4-vectors, as these are known to have the right transformation properties.

iii) We hope to recover the familiar laws of non-relativistic mechanics in the limit if small velocities.

iv) The ultimate test of the formalism is to be its success in predicting the results of experiment.

As a guide to satisfying iii) above in the final result, we shall attempt to make the laws as formally similar to those of non-relativistic mechanics as possible. The extent to which formal similarity can be achieved while maintaining overall Lorentz invariance cannot be anticipated - it turns out in fact that certain aspects of the non-relativistic theory are closely paralleled in the relativistic form.

Non-relativistic mechanics defines the "force" acting on a particle in terms of the acceleration evidenced by the particle's motion. The definition is accomplished through the equation

$$f = ma$$

where a is the acceleration and m is a determinable constant proper to the particle. The same approach will be adopted here, so that we define the 4-force through

$$F = MA$$

where A is as above, and M is a scalar. For M to be characteristic of the particle and in no way characteristic of the frame of reference we must now assert that M is to be an invariant. The precise nature of M will become apparent later, and its relation to the non-relativistic m also. For the moment it is to be regarded as an "abstract" invariant somehow characteristic of the particle. Similarly, by analogy with non-relativistic mechanics, we shall label the 4-vector $P = MV$ as the "4-momentum" of the particle.

The relationship between these definitions and the familiar quantities of non-relativistic mechanics will now be considered. Consider first the quantity dX/dt, which, as was remarked earlier, bears a simple relation to the "3-velocity".

$$\frac{dX}{dt} = \frac{dX}{dT}\cdot\frac{dT}{dt} = \frac{V}{\beta} \quad \text{from Page 2}$$

Thus $V = \beta v$ in its second, third and fourth components.

It follows from this that $P = MV = \beta Mv$, in its second, third, and fourth components.

Now the non-relativistic "3-momentum", p, is mv, and as v --) 0, β --) 1, so that P --) Mv, in its second, third, and fourth components. We can therefore say that the invariant M represents the mass of the particle as determined in the frame in which it is at rest. M is therefore called the "proper mass" or "rest-mass" of the particle. The proper mass M is clearly to be identified with the "non-relativistic" mass m appropriate to ordinary mechanics, and we can then say that the second, third, and fourth components of the 4-momentum may be obtained from the ordinary 3-velocity by multiplication, not by m, as in ordinary mechanics, but by $\beta M = \beta m$. This is equivalent to saying that the mass of a particle in relativistic mechanics is a velocity-dependent quantity $m(v) = \beta m(0)$, if one insists on using the non-relativistic velocity 3-vector in one's definition of momentum. This has the result that, for a particle with finite $m(0)$, the mass will apparently diverge to an infinite value as the particle is accelerated to velocities near that of light, for then β becomes very great. This is to say that it becomes progressively harder to accelerate a particle uniformly as its velocity approaches that of light, which is a fact recognised in the behaviour of high-energy particle accelerating machines. There is therefore a correspondence already between the somewhat formal definition of the 4-momentum and our physical experience. Note however that in relativity the "mass" is M, an invariant, and the β is taken up with a different definition of the velocity. It is because we insist on using the velocity as measured in the laboratory frame that the apparent mass of a particle is velocity-dependent.

Let us now examine the definition $F = MA$. This definition can be regarded as the "law of motion" of relativistic mechanics. It will be of interest once again to compare its nature with that of the corresponding law in non-relativistic mechanics.

We have $F = \frac{d}{dT}\left(M\frac{dX}{dT}\right) = \frac{d}{dT}\left(\beta M\frac{dX}{dt}\right) = \beta.\frac{d}{dt}\left(\beta M\frac{dX}{dt}\right)$

Now write $F/\beta = f$, which is not a 4-vector because β is not an invariant. For the second, third, and fourth components of the above equation we then have

$$f = \frac{d}{dt}(\beta Mv)$$

which bears a striking resemblance to the non-relativistic law, where the mass is again taken as a function of velocity $m(v) = \beta M$. With this modification, then, f corresponds to the non-relativistic "force". We note then that non-relativistic "3-forces", as calculated in one inertial frame, will not be transformed into their counterparts in another frame by the Lorentz transformation analogously to position co-ordinates. Instead, the quantities which are given by βf will transform in this way, as the components of the 4-vector F. Thus the idea of force intrinsic to Newtonian mechanics is modified in relativity. The measure of a force within the relativistic theory is different for observers in different inertial frames; in particular, if two forces acting at different points are considered equal in one inertial frame, they will not necessarily be considered so in another frame. Newton's Third Law is preserved in fact only for two bodies in actual contact.

Thus far we have been able to find strong similarities between non-relativistic mechanics and the Lorentz-transforming equations written in terms of 4-vectors. The only modification of the non-relativistic ideas to emerge so far is the apparent dependence of mass on velocity, and the non-absolute nature of the magnitude of a force as defined by non-relativistic mechanics. Suppose now that we try to form an expression for the work in terms of 4-vectors. Analogously to $w = f.\delta x$ we try $W = F.\delta X$

i.e. $W = \bar{F}\delta X = M\bar{A}\delta X = M.\frac{d\bar{V}}{dT}V.\delta T$ (where $\bar{N}$ denotes the transpose of N)

$$= \tfrac{1}{2}M.\frac{d}{dT}(\bar{V}V).\delta T$$

Now $\bar{V}V = \delta\bar{X}\delta X/\delta T^2 = -c^2\delta T^2/\delta T^2$, as $\delta\bar{X}\delta X = \delta s^2 = -c^2\delta T^2$ from the definition of δT, dT

$$W = \tfrac{1}{2}M.\frac{d}{dT}(-c^2).\delta T = 0$$

Therefore the work-scalar vanishes identically in this representation, that is to say, the 4-vectors F and δX are orthogonal (if we "4-push" a particle, it "moves" in a "4-direction" orthogonal to the "push"). This strange-looking result has a very interesting consequence.

As F is orthogonal to δX, F/β must also be orthogonal to δX, for division by β cannot destroy the orthogonality property in any inertial frame, β merely being a number. Thus the four-dimensional vector f is orthogonal to δX. Let us denote the four components of the vector f corresponding to the four components of δX (icδt, δx, δy, δz) by f_t, f_x, f_y, f_z. then the orthogonality of f and δX can be written

$$f_t.ic\delta t + f_x.\delta x + f_y.\delta y + f_z.\delta z = 0$$

Now the quantity $f_x.\delta x + f_y.\delta y + f_z.\delta z$ is what is meant in non-relativistic mechanics by the work δw.

We therefore have that

$$\delta w = - f_t.ic\delta t$$

To find f_t, we take the t-component of $f = \frac{d}{dt}(\beta M\frac{dX}{dt})$, i.e. $f_t = \frac{d}{dt}(\beta M\frac{d(ict)}{dt}) = ic\frac{d(\beta M)}{dt}$

Thus $\delta w = c^2\delta(\beta M)$ on substitution for f_t

But δw must equal δE, the change in energy of the particle, so that we must now equate the total energy E of a particle to

$$E = \beta Mc^2$$

This is Einstein's famous mass-energy equivalence, and its appearance in the present formalism must be regarded as strong evidence for the applicability of the equations. In terms of M, we have

$$E = Mc^2(1 - v^2/c^2)^{-\frac{1}{2}} = Mc^2 + \tfrac{1}{2}Mv^2 + \ldots.$$

The second term is the classical expression for the kinetic energy, and we see that relativity has added to this a "rest-energy" Mc^2, which a particle is to possess by virtue of its mass alone, and higher-order terms in v^4/c^2. The form of this expression for the total energy of a particle is well borne out by microscopic physics.

The final topic to be raised in this article is the extension of these ideas to more than one particle. This requires a further postulate, namely the conservation of total 4-momentum in any interactions. For then, if the sum of all individual P's is a constant, in another frame, the sum of all the P' vectors is obtained by Lorentz-transforming the sum of all the P vectors, each P transforming to a P' by the Lorentz transformation. Thus if the sum of all the P vectors is constant in one frame, the sum of all the P' vectors must be a constant in the other frame, for the Lorentz transformation of a constant can only give another (though possibly different) constant. The components of P are, by inspection, (icβM, βMdx/dt, βMdy/dt, βMdz/dt), so that the postulate of conservation of 4-momentum implies both the conservation of 3-momentum and the conservation of mass, (where the velocity-dependent mass is to be used) in non-relativistic mechanics. This postulate then enables a complete system of dynamics to be set up, including the effects of collisions. This will not be undertaken here.

Our final conclusion must therefore be that if we set out to write the laws of mechanics in terms of 4-vectors and Lorentz invariants, in the attempt to satisfy the requirement that they have the same form in all inertial frames, and furthermore make all our definitions and "laws" as similar in 4-dimensional form as possible to the Newtonian counterparts, we do indeed achieve an acceptable system of mechanics capable of dealing with high-velocity particles as observed experimentally. The fact that such ideas as the apparent variation of mass with velocity, and the Einstein mass-energy relationship could be obtained from such simple formal requirements must be regarded as convincing proof of the suitability of the basic approach. For full discussion of the experimental aspects, and a possibly less formal method of attack, any standard text on relativity should be consulted (for example, MOLLER, "The Theory of Relativity", O.U.P. 1952)