where

The Relativistic Dynamics of a System of Particles Interacting at a Distance

L. H. THOMAS

Watson Scientific Computing Laboratory, Columbia University, New York, New York

(Received November 1, 1951)

The dynamics of a system of particles acting on one another at a distance can be relativistically invariant if the assumption of invariant world-lines is given up. This is shown by constructing a particular dynamics in which invariance over the homogeneous Lorentz group is trivial, but space as well as time displacement requires the solution of equations of motion or of a Schroedinger equation. This particular dynamics reduces in the nonrelativistic limit to the most general dynamics of a system of interacting particles admitting the Newtonian group.

I. INTRODUCTION: DIFFICULTIES WITH WORLD LINES

THE Newtonian equations of motion of two interacting particles are, in one dimension of space,

$$m_1 \ddot{x}_1 = -\partial V / \partial x_1, \quad m_2 \ddot{x}_2 = -\partial V / \partial x_2,$$
$$V = V(x_2 - x_1). \tag{1.1}$$

These keep the same form if we change to a frame of reference in uniform relative motion by the equation

$$x' = x - ut, \tag{1.2}$$

and in such an inertial frame of reference the accelerations of the particles depend only on their relative position. This assumes a common time t, but does not identify points of space at one time with points of space at any other time (see Fig. 1).

In relativity theory we replace x' = x - ut by

$$\begin{aligned} x' &= (x - ut) / \sqrt{(1 - u^2/c^2)}, \\ t' &= (t - ux/c^2) / \sqrt{(1 - u^2/c^2)}, \end{aligned}$$
 (1.3)

c being the speed of light. A relation between x and t, transformed from frame to frame by these formulas,



FIG. 1. Nonrelativistic theory.

specifies a world-line for the motion of a particle (see Fig. 2), but when we try to generalize the equations of motion we immediately run into difficulty. In (1.1), $\ddot{x}_1(t)$ and $\ddot{x}_2(t)$ depend on $x_2(t) - x_1(t)$, and this is invariant for the transformation (1.2); but with (1.3), $t_2=t_1$ does not lead to $t_2'=t_1'$ unless $x_2=x_1$. There is, indeed, no reason why any property of the motion of particle 1 at a definite location on its world-line should depend on any one location of the second particle rather than on any other which can be made simultaneous with the location of the first particle by the transformation (1.3).

The current methods of avoiding this difficulty are to deal only with collisions, or with the interactions between particles and fields, for which we can take $x_2=x_1$; or to use retarded interactions, $x_1-x_2=c(t_1-t_2)$ leading to $x_1'-x_2'=c(t_1'-t_2')$. When these methods are applied in detail, consistent finite results are not obtained, and one seems to be led to noncausal theories, or to discontinuous space and time.¹

There is, however, in the above discussion an assumption that does not seem to be logically necessary, the assumption of invariant world lines. We may quite logically give up this assumption and suppose that the state of the system is specified relative to any observer in terms of canonically conjugate dynamical variables, q_1 and p_1 for the first particle, q_2 and p_2 for the second particle, and that these variables transform to the corresponding variables relative to any other observer, displaced in time, position, and velocity, by the transformations of a group given in canonical form. (See Fig. 3.) While $q_1(t)$ may be regarded as giving a world line for the first particle relative to a series of observers displaced only in time, there is no reason why this should be exactly the same line as that given by $q_1'(t')$ for a relatively moving series of observers.

This paper will be devoted to developing a relativistic theory of particles interacting at a distance by giving up the assumption of invariant world lines.

II. KINEMATICS AND DYNAMICS

We shall adopt the general point of view of quantum mechanics, and a notation briefly symmarized as follows.

¹ P. G. Bergmann, An Introduction to the Theory of Relativity (Prentice-Hall, Inc., New York, 1942), p. 85.



FIG. 2. Relativistic theory with invariant world lines.

Relative to any one observer the world is described in terms of basic dynamical variables a, b, \cdots , that can be represented by Hermitian matrices, various representations being obtained from one another by unitary transformations. A real function of the basic variables is in general a Hermitian matrix depending on the matrices representing the dynamical variables in a manner covariant over the unitary transformations. In particular $\lambda a + \mu b$ where λ and μ are numbers, and $\frac{1}{2}(ab+ba)$, are functions of a and b in this sense, and so is their Poisson bracket

$$(a, b) = (ab - ba)/\iota\hbar,$$
 (2.1)

while the trace $\langle f \rangle$ of any matrix f is invariant. Any property of the system may be described by such a function of the basic variables. A statistical state of the system is also described by a function P(a, b, ...) of the dynamical variables, of unit trace, but it must be analyzable into a sum with positive coefficients over pure states, or a limit of such a sum, the statistical matrix describing a pure state being the open product of a column φ and the complex conjugate row φ^{z}

$$\mathbf{A} = \varphi \varphi^x. \tag{2.2}$$

The kinematics of the system is given by one set of representations of the basic variables, or by relations between them sufficient to determine their representations up to a unitary transformation. Expected values of properties for a state of the system are given by the trace of the product of the matrices describing the property and the state.

A continuous series of observers, referred to by the parameters s, describe the world in terms of basic dynamical variables a(s), b(s), \cdots , defined by each observer relative to himself, for example, in the non-

relativistic case, Cartesian coordinates, momenta, and spins, of the various particles in that observer's frame of reference. We shall suppose that for all the observers the system has the same kinematics. The representations used by the observers can be chosen so that to the same statistical state they give the same matrix. For such a Heisenberg set of representations the matrices representing the dynamical variables, a(s), will in general be different for each observer, and the change from one observer to a neighboring observer of any function of the basic variables not explicitly depending on the observer can be put in canonical form

$$df/ds = (f, S), \qquad (2.3)$$

where S is also a matrix function. These equations go over to the classical equations of motion by replacing the quantum Poisson bracket by a classical Poisson bracket. Again the representations can be chosen so that the basic variables are given by the same matrices whatever the observer. For such a Schrödinger set of representations, the matrices P(s) representing a statistical state must in general be different, and the change from one observer to a neighboring observer can be put in the form

$$\partial \mathbf{P}(s)/\partial s = (\mathbf{S}, \mathbf{P}),$$
 (2.4)

where S must be the previous matrix of (2.3) transformed to the new set of representations. For a column describing a pure state we have, perhaps after adjusting the phases of the representations,

$$\partial \varphi / \partial s = S \varphi / \iota \hbar,$$
 (2.5)

the Schrödinger equation.

The above point of view is familiar enough if our observers' descriptions are the usual descriptions in non-relativistic theory at different times, s being the time



FIG. 3. Relativistic theory with world lines not invariant.

variable t, and S the Hamiltonian function H.² We shall extend this point of view to relativity theory by treating in this displacements of the observer in position, orientation, and velocity, as well as in time. Thus position in space and time, as well as orientation and velocity, will be regarded as properties of observers, who form a ten-parameter family, with the structure of the inhomogeneous Lorentz group. The dynamical variables are not regarded as directly related to space position, and there is no need to introduce many times.

III. THE INHOMOGENEOUS LORENTZ GROUP

We may define the rate of displacement along a series of observers with change of parameter s by the component rates of displacement in position, $\partial x/\partial s$, $\partial y/\partial s$, $\partial z/\partial s$ along the x, y, and z directions, in orientation, $\partial l/\partial s$, $\partial m/\partial s$, $\partial n/\partial s$, about the x, y, and z axes, in velocity $\partial u/\partial s$, $\partial v/\partial s$, $\partial w/\partial s$, in the x, y, and z directions, and in time $\partial t/\partial s$, relative to each observer. We then have rates of change like (2.3) in a Heisenberg set of representations or (2.4) in a Schrödinger set, with

$$S = H \frac{\partial t}{\partial s} + U \frac{\partial u}{\partial s} + V \frac{\partial v}{\partial s} + W \frac{\partial w}{\partial s} - L \frac{\partial l}{\partial s} - M \frac{\partial m}{\partial s} - N \frac{\partial n}{\partial s} - X \frac{\partial x}{\partial s} - Y \frac{\partial y}{\partial s} - Z \frac{\partial z}{\partial s},$$
(3.1)

where the signs have been taken so that H can be identified with the energy, X, Y, and Z, with the components of linear momentum, and L, M, and N, with the components of angular momentum of the whole system.

Here the ten matrices, H, U, V, W, L, M, N, X, Y, and Z, will be functions of the basic dynamical variables and in general of the observer, but if the system is a complete isolated system, satisfying the same physical laws for each observer, they should be explicitly independent of the observer and should be given by the same matrices for all observers in a Schrödinger set of representations.

Further, in order that we should return to the same description of the system on completing a closed circuit of observers, it is necessary and sufficient that the Poisson brackets of these matrices may take the forms, corresponding to the structure of the inhomogeneous Lorentz group,

The first set of equations, (3.2), gives the Poisson brackets of the six functions, U, V, W, L, M, and N, in terms of themselves only, giving the structure of the homogeneous Lorentz group for changes of orientation and velocity between observers at the same point of space-time. U, V, W, L, M, and N themselves form a six-vector for this group, and the conditions that any other six functions of the basic variables form a sixvector are equations obtained from (3.2) by replacing the right-hand side and one variable on the left-hand side of each with the new set.

The second set of equations, (3.3), gives the Poisson brackets of the six functions U, V, W, L, M, and N, with the four functions H, X, Y, and Z; shows that the latter form a four-vector for the homogeneous Lorentz

group; and gives the changes in the former in spacetime displacements of the observer. Any other set of four functions of the basic variables form a four-vector if they satisfy similar equations.

Lastly, Eqs. (3.4), stating that H, X, Y, and Z commute with each other, complete the conditions that the functions H, U, V, W, L, M, N, X, Y, and Z of the basic variables give a dynamical theory admitting the inhomogeneous Lorentz group.³

IV. THE DYNAMICAL EQUATIONS

In choosing basic dynamical variables and kinematical relations between them to represent a single

² P. A. M. Dirac, *Quantum Mechanics* (Oxford University Press, London, 1930), Chap. VI.

³ Similar theory dealing with a single particle or a system not necessarily composed of particles is to be found in E. Wigner, *Annals of Mathematics* **40**, 145 (1939); C. Møller, Communications Dublin Institute for Advanced Studies 3 (1949), and M. H. L. Pryce, Proc. Roy. Soc. (London) **A195**, 621 (1948).

particle, we may be guided by the consideration that for a single particle by itself we should be able to construct matrices satisfying the relations (3.2), (3.3), and (3.4). This can be done giving them in terms of more special variables, of variables equivalent to this set, or of more general variables. We shall adopt the second possibility and simply take a particle to be described by basic dynamical variables satisfying the relations (3.2), (3.3), and (3.4).

For n particles, then, we take n sets of variables,

$$H_r, U_r, V_r, W_r, L_r, M_r, N_r, X_r, Y_r, Z_r, r = 1, \dots, n, (4.1)$$

each set satisfying the conditions (3.2), (3.3), and (3.4), and with zero Poisson brackets for variables from different particles.

In general, a relativistic theory could be given by any ten functions L, M, N, U, V, W, X, Y, Z, and H of the basic variables, satisfying (3.2), (3.3), and (3.4), but this would be far too general for our purpose.

If we examine customary constructions, we see that they usually introduce invariance over space displacements and rotations implicitly, the construction making this trivial; the effect of time displacement is the main one to be described, by a Hamiltonian or otherwise, and invariance for velocity displacement requires proof. Thus we may take

$$\begin{array}{ll} \mathbf{X} = \boldsymbol{\Sigma} \mathbf{X}_r, & \mathbf{Y} = \boldsymbol{\Sigma} \mathbf{Y}_r, & \mathbf{Z} = \boldsymbol{\Sigma} \mathbf{Z}_r, \\ \mathbf{L} = \boldsymbol{\Sigma} \mathbf{L}_r, & \mathbf{M} = \boldsymbol{\Sigma} \mathbf{M}_r, & \mathbf{N} = \boldsymbol{\Sigma} \mathbf{N}_r, \end{array}$$
 (4.2)

so that invariance over space displacements and rotations is trivial, choose a Hamiltonian function H which commutes with X, Y, Z, L, M, and N, and prove relativistic invariance, which is equivalent to constructing functions U, V, and W to give infinitesimal velocity displacements, such that the whole set X, Y, Z, L, M, N, H, U, V, and W satisfy (3.2), (3.3), and (3.4).

We shall depart from this scheme and instead make trivial the homogeneous Lorentz transformations, space rotations, and changes of velocity for a fixed space-time position. This seems to be more elegant mathematically and to lead to simpler results, and has the further advantage that it will extend directly to the usual form of the general theory of relativity, where just the homogeneous Lorentz group of transformations at each space-time position is taken trivial.⁴

Thus we assume to start with

$$L = \Sigma L_r, \quad M = \Sigma M_r, \quad N = \Sigma N_r, \\ U = \Sigma U_r, \quad V = \Sigma V_r, \quad W = \Sigma W_r,$$
(4.3)

which, it is trivial, satisfy (3.2). We need in addition functions H, X, Y, and Z, satisfying (3.3) and (3.4). The conditions (3.4) require that they shall mutually commute. The conditions (3.3) require that they are four functions of the basic variables L_r , M_r , N_r , U_r , V_r , W_r , H_r , X_r , Y_r , and Z_r , $r=1, \dots, n$, forming a fourvector for the transformations of the homogeneous Lorentz group. Any four-vector function of the basic variables with mutually commuting components will give us a relativistically invariant dynamics.

V. CONSTRUCTION OF A SPECIAL DYNAMICS

Assume

$$X = \frac{1}{2} (\mu \Sigma X_r + \Sigma X_r \mu),$$

$$Y = \frac{1}{2} (\mu \Sigma Y_r + \Sigma Y_r \mu),$$

$$Z = \frac{1}{2} (\mu \Sigma Z_r + \Sigma Z_r \mu),$$

$$H = \frac{1}{2} (\mu \Sigma H_r + \Sigma H_r \mu),$$

(5.1)

where μ is a single function of the basic variables, a scalar for the homogeneous Lorentz group; then X, Y, Z, and H will be Hermitian and will form a four-vector satisfying Eqs. (3.3). In order that the expressions (5.1) should commute, it is sufficient that μ should commute with the ratios of ΣX_r , ΣY_r , ΣZ_r , and ΣH_r .

$$\Sigma X_r \mu \Sigma Y_r = \Sigma Y_r \mu \Sigma X_r, \quad \text{etc.} \tag{5.2}$$

Thus μ may be any scalar function of expressions that have this property, in particular of X_r , Y_r , Z_r , and H_r , $r=1, \dots, n$, and of expressions constructed by the following method. Take any determinant D with five rows and columns, with rows chosen out of the 6n+1rows

where Q is one of L_r , M_r , N_r , U_r , V_r , or W_r , $r=1, \dots, n$. For all these Poisson brackets when evaluated contain only X_r , Y_r , Z_r , and H_r , $r=1, \dots, n$, which commute with each other. Thus, for example,

$$(\Sigma X_r, D) \Sigma Y_r - (\Sigma Y_r, D) \Sigma X_r,$$

where D is one of these determinants, is a determinant with corresponding rows and columns from

$$\begin{array}{cccc} \Sigma X_r, & \Sigma Y_r, & \Sigma Z_r, & \Sigma H_r, \\ (\Sigma X_r, Q), & (\Sigma Y_r, Q), & (\Sigma Z_r, Q), & (\Sigma H_r, Q), \\ & & 0, \\ & & (\Sigma X_r, Q) \Sigma Y_r - (\Sigma Y_r, Q) \Sigma X_r, \end{array}$$

and, since all these terms commute, is seen by combining the columns to vanish, and this is just the condition required. In this way we obtain, in all, 10n-4allowed combinations of the original 10n variables, which is as many as we should expect.

We may arrange these more symmetrically as the components of four-vectors.

(a)
$$X_r, Y_r, Z_r$$
, and $H_r, r=1, \dots, n.$ (5.4)

The length of this four-vector may be regarded as the rest-mass of the particle.

(b) $L_rH_r/c^2+W_rY_r-Y_rZ_r$, $M_rH_r/c^2+U_rZ_r-W_rX_r$, $N_rH_r/c^2+V_rX_r-U_rY_r$, and $L_rX_r+M_rY_r+N_rZ_r$, $r=1, \cdots, n.$ (5.5)

⁴L. H. Thomas, Revs. Modern Phys. 17, 182 (1945); P. M. Dirac, Revs. Modern Phys. 21, 392 (1949).

The factors of the terms of the four-vector (b) commute. The components of the four-vector (b) commute with X_r , Y_r , Z_r , and H_r . The four-vector (b) has zero scalar product with the four-vector (a), and it may be regarded as the product of the rest-mass of the particle and its spin angular momentum.

(c) Differences for different values of r of the ratios of

$$\begin{array}{l} U_r \Sigma H_s + M_r \Sigma Z_s - N_r \Sigma Y_s, \quad V_r \Sigma H_s + N_r \Sigma X_s - L_r \Sigma Z_s, \\ W_r \Sigma H_s + L_r \Sigma Y_s - M_r \Sigma Z_s, \end{array}$$

and

$$U_{r}\Sigma X_{s} + V_{r}\Sigma Y_{s} + W_{r}\Sigma Z_{s},$$

to $H_{r}\Sigma H_{s}/c^{2} - X_{r}\Sigma X_{s} - Y_{r}\Sigma Y_{s} - Z_{r}\Sigma Z_{s}.$ (5.6)

The factors of terms of these four-vectors do not commute, but the changes made in them by taking the factors in the opposite order are proportional to X_r , Y_r , Z_r , and H_r . The four-vectors have zero scalar product with ΣX_s , ΣY_s , ΣZ_s , and ΣH_s .

Thus to give our dynamics we take for μ any scalar function of the four-vectors (5.4), (5.5), and (5.6), which means any function of their lengths and scalar products, and the expressions (5.1) will commute so that (3.2), (3.3), and (3.4) will all be satisfied.

VI. THE NONRELATIVISTIC APPROXIMATION⁶

If
$$m_r^2 c^4 = H_r^2 - c^2 (X_r^2 + Y_r^2 + Z_r^2),$$
 (6.1)

 m_r is a scalar and commutes with all the other variables. The nonrelativistic limit for the basic variables is obtained by keeping all except H_r , $r=1, \dots, n$, finite as c tends to infinity, but putting

$$H_r - m_r c^2 \rightarrow \frac{1}{2m_r} (X_r^2 + Y_r^2 + Z_r^2),$$
 (6.2)

following (6.1).

In the limit, U_r , V_r , and W_r commute, and $(U_r, X_r) = (V_r, Y_r) = (W_r, Z_r) = m_r$, so we can write

$$\mathbf{U}_r = m_r \mathbf{x}_r, \quad \mathbf{V}_r = m_r \mathbf{y}_r, \quad \mathbf{W}_r = m_r \mathbf{z}_r, \quad (6.3)$$

where x_r , y_r , and z_r are canonically conjugate to X_r , Y_r , and Z_r .

$$(\mathbf{x}_r, \mathbf{X}_r) = 1, \quad (\mathbf{y}_r, \mathbf{Y}_r) = 1, \quad (\mathbf{z}_r, \mathbf{Z}_r) = 1.$$
 (6.4)

We then find that if we write

$$L_r = \lambda_r + m_r (\mathbf{y}_r \mathbf{Z}_r - \mathbf{z}_r \mathbf{Y}_r),$$

$$M_r = \mu_r + m_r (\mathbf{z}_r \mathbf{X}_r - \mathbf{x}_r \mathbf{Z}_r),$$

$$N_r = \nu_r + m_r (\mathbf{x}_r \mathbf{Y}_r - \mathbf{y}_r \mathbf{X}_r),$$

(6.5)

 λ_r , μ_r , and ν_r have Poisson brackets as for components of angular momentum,

$$(\lambda_r, \mu_r) = \nu_r, \quad (\mu_r, \nu_r) = \lambda_r, \quad (\nu_r, \lambda_r) = \mu_r.$$
 (6.6)

The remaining Poisson brackets of x_r , y_r , z_r , X_r , Y_r , Z_r , λ_r , μ_r , ν_r , and m_r vanish.

Thus \mathbf{x}_r , \mathbf{y}_r , \mathbf{z}_r , \mathbf{X}_r , \mathbf{Y}_r , \mathbf{Z}_r , λ_r , μ_r , ν_r , and m_r have just the right Poisson brackets to be interpreted as the Cartesian coordinates, components of linear momentum, spin angular momentum, and mass of a nonrelativistic particle, with kinetic energy $\mathbf{H}_r - m_r c^2$, and angular momentum components \mathbf{L}_r , \mathbf{M}_r , and \mathbf{N}_r , while Eqs. (2.4) and (3.1) will give its displacement as an isolated particle in space, time, orientation, or velocity, correctly.

If we further write

$$\mu = 1 + V / \Sigma m_r c^2, \tag{6.7}$$

we find in the nonrelativistic limit,

$$L = \Sigma L_r = \Sigma \{ m_r (y_r Z_r - z_r Y_r) + \lambda_r \},$$

$$M = \Sigma M_r = \Sigma \{ m_r (z_r X_r - x_r Z_r) + \mu_r \},$$

$$N = \Sigma N_r = \Sigma \{ m_r (x_r Y_r - y_r X_r) + \nu_r \},$$
(6.8)

$$U = \Sigma U_r = \Sigma m_r x_r,$$

$$V = \Sigma V_r = \Sigma m_r y_r,$$

$$W = \Sigma W_r = \Sigma m_r z_r.$$

(6.9)

and

$$\mathbf{X} = \boldsymbol{\Sigma} \mathbf{X}_r, \quad \mathbf{Y} = \boldsymbol{\Sigma} \mathbf{Y}_r, \quad \mathbf{Z} = \boldsymbol{\Sigma} \mathbf{Z}_r, \quad (6.10)$$

with finally

$$\mathbf{H} - \Sigma m_r c^2 = \Sigma \frac{1}{2m_r} (\mathbf{X}_r^2 + \mathbf{Y}_r^2 + \mathbf{Z}_r^2) + V, \quad (6.11)$$

where V has to be constructed out of the limiting principal parts of scalar products of the four-vectors (5.4), (5.5), and (5.6).

These four-vectors take the forms

(a)
$$X_r, Y_r, Z_r, m_r c^2 + (X_r^2 + Y_r^2 + Z_r^2)/2m_r$$

(b) $\lambda_r, \mu_r, \nu_r, X_r\lambda_r + Y_r\mu_r + Z_r\nu_r$

and (c) differences for two different values of r of

$$\mathbf{x}_r, \mathbf{y}_r, \mathbf{z}_r, \mathbf{x}_r(\Sigma \mathbf{X}_s/\Sigma m_s) + \mathbf{y}_r(\Sigma \mathbf{Y}_s/\Sigma m_s) + \mathbf{z}_r(\Sigma \mathbf{Z}_s/\Sigma m_s).$$

The limiting principal parts reduce to scalar products of (a) differences for different values of r of the three-vectors

$$X_r/m_r, Y_r/m_r, Z_r/m_r,$$
 (6.12)

(b) the three-vectors

$$\lambda_r, \ \mu_r, \ \nu_r,$$
 (6.13)

and (c) differences of the three-vectors

$$x_r, y_r, z_r.$$
 (6.14)

Thus in nonrelativistic limit we have the usual theory with a potential energy any scalar function of the relative positions, relative velocities, spin angular momenta, and masses of the particles.

⁵ The equations given by C. G. Darwin, Phil. Mag. **39**, 537 (1920), and G. Breit, Phys. Rev. **34**, 553 (1929); **51**, 248 (1937); **53**, 153 (1938), may be regarded as approximations to order v^2/c^2 .