General Relativity and Particle Dynamics

L. H. Thomas
Watson Scientific Computing Laboratory, Columbia University, New York, New York
(Received August 18, 1958)

In order to adapt the Hamiltonian dynamics of a system of classical (nonquantum) particles to general relativity it seems necessary to generalize the transformations used from contact transformations of phase space to extended point transformations of a joint observer and phase space. When this is done the curvature of observer space is seen to arise from the physical interaction of particles without the introduction of new gravitational field variables.

In order to extend this theory to quantized particle and field theory it seems necessary to extend the unitary transformations of quantum mechanics in a similar manner, to linear transformations of matrices maintaining the trace of the product of two matrices, the Hermitian nature of a matrix, and the unit matrix. Such transformations do not preserve the phases of wave functions or the product of two matrices, but do preserve the characteristic values of matrices and the traces of the products of any number of matrices. That observers should be designated by q numbers, not c numbers, makes it difficult to interpret the theory except in the classical limit.

INTRODUCTION

The essential starting point of quantum mechanics is that a statistical matrix gives the state of the universe under discussion relative to an observer; and the laws of dynamics provide canonical transformations from this matrix to those giving the same absolute state relative to other observers. In general relativity, however, the observer reached from one observer by a displacement, for instance to a later time, depends on what the state of the universe is; change in the numbering of observers may change this dependence but does not in general remove it. The starting point requires therefore some modification, and what this modification should be may be suggested from classical (that is nonquantum) mechanics by using the correspondence between Poisson brackets and commutators (Dirac's rule), which, however, must be extended to a wider set of transformations.[1]

1. CLASSICAL (NONQUANTUM) DYNAMICS IN CANONICAL FORM RELATIVE TO A GROUP OF OBSERVERS

Although an observer and his equipment are really part of the universe, we shall formulate dynamics first relative to a group of observers specified by parameters.[2]

A specific (nonstatistical) absolute state of the universe under discussion is given by a set of values of a number, say n,[3] of pairs of canonically conjugate variables, $\alpha_1, \beta^1; \alpha_2, \beta^2; \cdots; \alpha_n, \beta^n$. Any given absolute state may be described relative to any one of an s-parameter family of observers numbered by s continuous parameters $\xi^1, \xi^2, \cdots, \xi^s$.[4] This description may be given by a set of values of n pairs of canonically conjugate basic dynamical variables $p_1, q^1; p_2, q^2; \cdots; p_n, q^n$. These variables are functions of $\alpha_1, \beta^1; \alpha_2, \beta^2; \cdots; \alpha_n, \beta^n$ given by a set of contact transformations[5] depending on $\xi^1, \xi^2, \cdots, \xi^s$ as parameters. Any other dynamical variable f, which can be observed by one of the observers, will be a function of the variables $p_1, q^1; p_2, q^2; \cdots; p_n, q^n$ for that observer.

The dynamical variables have physical meaning, so that, while one-to-one transformations from $p_1, q^1; p_2, q^2; \cdots; p_n, q^n$ to another set, $r^1, r^2, \cdots, r^{2n}$, say, may be used for analysis and simplification, the new variables have in general different physical meaning from the old, as when we change from Cartesian components of relative position of two particles to distance apart and position angles. The variables giving the absolute states are subject to arbitrary contact transformations, which may be called "changes of representation," and the parameters giving the observers are subject to arbitrary continuous transformations.

We suppose further than an observer can characterize neighboring observers as having a physical relationship to him, described by an infinitesimal displacement with s components, $dx^1, dx^2, \cdots, dx^s$,[6] suppose, in terms of which changes in $\xi^1, \xi^2, \cdots, \xi^s$ can be expressed linearly, and reversely,

$$\begin{aligned} d\xi^\alpha &= L_a{}^\alpha dx^a, \quad \alpha=1, 2, \cdots, s, \\ dx^a &= L_\alpha{}^a d\xi^\alpha, \quad a=1, 2, \cdots, s, \end{aligned} \tag{1.1}$$

where we have used the usual summation convention over a repeated index, and where $L_a{}^\alpha$ are coefficients that may be functions of $\xi^1, \xi^2, \cdots, \xi^s$, but have nonvanishing determinant, and $L_\alpha{}^a$ are the reciprocal set. $dx^1, dx^2, \cdots, dx^s$ must be regarded as differentials of quasi-parameters,[7] there being no true variables corresponding to them.

[1] L. H. Thomas, Revs. Modern Phys. **17**, 182 (1945), contains the ideas of this paper except for widening the transformations.
[2] L. H. Thomas, Phys. Rev. **85**, 868 (1952), where this formulation is extended to quantum mechanics.
[3] In general n need not be finite.
[4] We shall take $s=10$, for observers with various orientations and velocities at various places in four-dimensional space-time. The simpler case $s=3$, for observers with various velocities at various places in a world with time and one space dimension, may also be considered.
[5] E. T. Whittaker, *Analytical Dynamics* (Cambridge University Press, Cambridge, 1937), fourth edition, p. 292.
[6] In general these will be displacements in time, dt, position, dx, dy, dz, velocity, du, dv, dw, and orientation, dl, dm, dn.
[7] Reference 5, p. 41.

A continuous series of observers is given by making $\xi^1, \xi^2, \cdots, \xi^s$ functions of a variable s, and the rate of change of a dynamical variable observed by this series of observers for a fixed state is given by

$$\frac{df}{ds}=\frac{\partial f}{\partial \xi^\alpha}\frac{d\xi^\alpha}{ds}=\frac{\partial f}{\partial x^a}\frac{dx^a}{ds}, \tag{1.2}$$

where

$$\begin{aligned} \partial f/\partial \xi^\alpha &= L_\alpha{}^a \partial f/\partial x^a, \quad \alpha=1,2,\cdots,s,\\ \partial f/\partial x^a &= L_a{}^\alpha \partial f/\partial \xi^\alpha, \quad a=1,2,\cdots,s, \end{aligned} \tag{1.3}$$

in terms of rates of change $\partial f/\partial x^a$ with respect to the physical displacements $dx^1, dx^2, \cdots, dx^s$. While these rates of change are not in general rates of change with respect to true variables, we regard them as having an invariant physical meaning that the rates of change $\partial f/\partial \xi^\alpha$ with respect to the arbitrary parameters do not have.

The laws of dynamics now give for equations of motion as seen by an observer, the equations, in canonical form,

$$\partial f/\partial x^a = (X_a, f), \quad a=1,2,\cdots,s, \tag{1.4}$$

where $X_1, X_2, \cdots, X_s$ are s functions of the dynamical variables $p_1, q^1; p_2, q^2; \cdots; p_n, q^n$; and the Poisson brackets have the usual meaning[8]

$$(u,v)=\frac{\partial u}{\partial q^k}\frac{\partial v}{\partial p_k}-\frac{\partial u}{\partial p_k}\frac{\partial v}{\partial q^k}. \tag{1.5}$$

If the dynamical system is the whole universe considered, we suppose that these rates of change depend only on the values of the dynamical variables $p_1, q^1; p_2, q^2; \cdots; p_n, q^n$ and, expressed in terms of these, are independent of the observer, that is, of $\xi^1, \xi^2, \cdots, \xi^s$. We suppose indeed that these s functions $X_1, X_2, \cdots, X_s$ specify the nature of the dynamical system in a manner invariant with respect to the transformation of $\alpha_1, \beta^1; \alpha_2, \beta^2; \cdots; \alpha_n, \beta^n$ and of $\xi^1, \xi^2, \cdots, \xi^s$.

The equations giving $q_1, p^1; q_2, p^2; \cdots; q_n, p^n$ in terms of $\alpha_1, \beta^1; \alpha_2, \beta^2; \cdots; \alpha_n, \beta^n$; and $\xi^1, \xi^2, \cdots, \xi^s$ are now the finite equations of a group of transformations of which (1.4) give the infinitesimal transformations. Finding the finite equations is equivalent to integrating the equations of motion (1.4), and the necessary and sufficient conditions that this should be possible are that the Poisson brackets of any pairs of the functions $X_1, X_2, \cdots, X_s$ should be linear combinations of these functions with constant coefficients,

$$(X_a, X_b) = C_{ab}{}^e X_e + h_{ab}. \tag{1.6}$$

The constants $C_{ab}{}^e$ are the structure constants of the group and have definite values determined by the physical meaning of the displacements corresponding to $X_1, X_2, \cdots, X_s$. They necessarily satisfy Jacobi relations.[9] The additive constants h_{ab}, which must also satisfy certain relations, can be adjusted by adding constants to $X_1, X_2, \cdots, X_s$, and in particular, for the inhomogeneous Lorentz group of special relativity, can be reduced to zero simultaneously.

Statistical states are in this theory defined by a function of the dynamical variables, $\mathrm{P}(q^1, \cdots, p_n)$, the density in phase, such that

$$\int \cdots \int \mathrm{P}(q^1,\cdots,p_n)dq^1\cdots dp_n = 1, \tag{1.7}$$

and a consistent theory can be built up because the volume element of phase space,

$$dq^1\cdots dp_n = d\beta^1\cdots d\alpha_n, \tag{1.8}$$

is invariant for contact transformations.

We can immediately include the observer as part of the dynamical system if there is no interaction between them and if the observer is described by a dynamical system admitting the same group. The variables $p_1, q^1; p_2, q^2; \cdots; p_n, q^n$ are made up of two sets, for the observer and for the rest of the universe, and the functions $X_1, X_2, \cdots, X_s$ are sums of functions of each set separately. If the observer's system by itself provides a faithful representation of the group, the dynamics of the rest of the universe can be described by functions giving its variables in terms of those of the observer, these functions being invariant for the displacement operators (1.4).

When there is interaction between the observer and the rest of the universe, we seem to require, even if the dynamics still admits a group such as the inhomogeneous Lorentz group, at least some of the ideas and notation suitable for more general cases. This seems to be true even in the important special case in which the observer's dynamical system is reduced to one "test particle," no different from others of the universe, and interacting with them.

2. A GENERALIZATION OF CLASSICAL DYNAMICS WHEN THE OBSERVERS DO NOT FORM A GROUP

We desire to generalize the above formulation so that while the laws of nature still appear the same to each observer, their mutual relationship may have a curvature depending on the state of the universe. This will be effected first with the observers specified by parameters. We would like the coefficients $C_{ab}{}^e$ to depend on the dynamical variables, and the coefficients $L_a{}^\alpha$ and $L_\alpha{}^a$ to depend on the dynamical variables or on the variables specifying the state of the universe as well as on the parameters specifying the observer. It is then only in a special class of cases that the transformations of dynamical variables from observer to observer

[8] Reference 5, p. 299. Eisenhart uses the opposite sign; see reference 9, p. 261.

[9] L. P. Eisenhart, *Continuous Groups of Transformation* (Princeton University Press, Princeton, New Jersey, 1933), p. 26.

can be contact transformations of these variables alone (see Sec. 4 below).

We describe a specific state of the universe in terms of m dynamical variables $\gamma^1, \gamma^2, \cdots, \gamma^m$, and any given state may be described relative to any one of an s-parameter family of observers numbered by s continuous parameters $\xi^1, \xi^2, \cdots, \xi^s$ by the values of m dynamical variables $r^1, r^2, \cdots, r^m$, which will be functions of $\gamma^1, \gamma^2, \cdots, \gamma^m$, and of $\xi^1, \xi^2, \cdots, \xi^s$. Any other dynamical variable f is a function of $r^1, r^2, \cdots, r^m$. The dynamical variables have physical meaning, while the variables $\gamma^1, \gamma^2, \cdots, \gamma^m$, describing an absolute state, are subject to arbitrary point-transformations. Thus we are not immediately requiring the variables describing a state to comprise canonically conjugate pairs. Further, we allow arbitrary point transformations of the parameters $\xi^1, \xi^2, \cdots, \xi^s$, which may depend also on the state given by $\gamma^1, \gamma^2, \cdots, \gamma^m$. We still have physical displacements $dx^1, dx^2, \cdots, dx^s$ such that equations (1.1) hold, but the coefficients $L_a{}^\alpha$, $L_\alpha{}^a$ may now be functions of $\gamma^1, \gamma^2, \cdots, \gamma^m$, as well as of $\xi^1, \xi^2, \cdots, \xi^s$. Equations (1.2) and (1.3) also still hold.

The laws of dynamics now give for equations of motion as seen by an observer the sm rates of change of the dynamical variables with physical displacements as functions of the dynamical variables,

$$\partial r^k/\partial x^a = \eta_a{}^k, \quad a=1, 2, \cdots, s; \quad k=1, 2, \cdots, m. \quad (2.1)$$

If the system is complete in itself, we still suppose that these rates of change depend only on $r^1, r^2, \cdots, r^m$, and indeed these sm functions of $r^1, r^2, \cdots, r^m$ specify the nature of the dynamical system in a manner invariant with respect to the more general transformations of $\gamma^1, \gamma^2, \cdots, \gamma^m$, and of $\xi^1, \xi^2, \cdots, \xi^s$. The rates of change under displacements of the observer of an arbitrary dynamical variable are now given by the s linear differential operators

$$D_a = \sum_{k=1}^{m} \eta_a{}^k \frac{\partial}{\partial r^k}, \quad (2.2)$$

and the equations of motion (1.4) are replaced by

$$\partial f/\partial x^a = D_a f. \quad (2.3)$$

These coefficients $\eta_a{}^k$ cannot be arbitrary functions of $r^1, r^2, \cdots, r^m$; in order that variables $\xi^1, \xi^2, \cdots, \xi^s$; $\gamma^1, \gamma^2, \cdots, \gamma^m$ should exist such that equations of the form (1.3) hold, it is necessary that the operators $D_1, D_2, \cdots, D_s$ should form a complete set. Conversely, if the operators (2.2) form a complete set, being s operators on m variables, they have in general $m-s$ invariants,[10] and we may transform $r^1, r^2, \cdots, r^m$ to new variables $s^1, s^2, \cdots, s^m$, of which the last $m-s$ are invariants, the first s any other independent combinations, and we shall have

$$D_a = \sum_{k=1}^{s} \zeta_a{}^k \frac{\partial}{\partial s^k}, \quad (2.4)$$

[10] Reference 9, p. 9.

where the coefficients $\zeta_a{}^k$ are functions of $s^1, s^2, \cdots, s^m$. If we now write

$$s^k = \varphi^k(\xi^1, \xi^2, \cdots, \xi^s; \gamma^1, \gamma^2, \cdots, \gamma^m), \quad k=1, 2, \cdots, s, \quad (2.5)$$
$$s^k = \psi^k(\gamma^1, \gamma^2, \cdots, \gamma^m), \quad k=s+1, s+2, \cdots, m,$$

and take

$$L_a{}^\alpha = \sum_{k=1}^{s} \zeta_a{}^k \frac{\partial \xi^\alpha}{\partial s^k}, \quad \alpha=1, 2, \cdots, s; \quad a=1, 2, \cdots, s; \quad (2.6)$$

expressed in terms of $\xi^1, \xi^2, \cdots, \xi^s; \gamma^1, \gamma^2, \cdots, \gamma^m$, by (2.5), then we have the general solution, and, incidentally, the condition is sufficient.

The necessary and sufficient conditions that (2.2) form a complete set are given by the conditions that their commutators

$$(D_a, D_b)f = D_a D_b f - D_b D_a f, \quad (2.7)$$

which are also linear differential operators, should be linear combinations of the operators themselves:

$$(D_a, D_b) = C_{ab}{}^e D_e, \quad (2.8)$$

where the coefficients $C_{ab}{}^e$ which we may call "structure functions" may be functions of $r^1, r^2, \cdots, r^m$.

In this theory statistical mechanics is a little complicated because which observer has a given description depends on the specific state. It is natural to describe a statistical state by a density in $\gamma^1, \gamma^2, \cdots, \gamma^m$ space, as usual:

$$\mathrm{P}(\gamma^1, \gamma^2, \cdots, \gamma^m) d\gamma^1 d\gamma^2 \cdots d\gamma^m \quad (2.9)$$

is to be invariant for transformations of $\gamma^1, \gamma^2, \cdots, \gamma^m$ space. There is, however, no unique meaning for the expected value of a dynamical variable relative to an observer with a given designation, invariant for our general transformations.

We must specify "statistical observers" by densities in observer space, which may be functions of the state also,

$$\varphi(\xi^1, \xi^2, \cdots, \xi^s, \gamma^1, \gamma^2, \cdots, \gamma^m) d\xi^1 d\xi^2 \cdots d\xi^s, \quad (2.10)$$

to be invariant for transformations of $\xi^1, \xi^2, \cdots, \xi^s$ space also perhaps involving $\gamma^1, \gamma^2, \cdots, \gamma^m$. The expectation value of a dynamical variable f is then

$$\int \cdots \int f \varphi \mathrm{P} d\xi^1 \cdots d\xi^s d\gamma^1 \cdots d\gamma^m, \quad (2.11)$$

and is invariant for our general transformations. An ordinary observer is now a special case with φ having as a factor a δ-function in $\xi^1, \xi^2, \cdots, \xi^s$ space, and, if the expectation values of all dynamical variables are given relative to one such observer, P is determined.

To be able to build up thermodynamics conveniently, we may assume further that all our transformations preserve the volume element

$$dr^1dr^2\cdots dr^m = d\gamma^1 d\gamma^2\cdots d\gamma^m \tag{2.12}$$

(or if we make transformations of $\gamma^1, \gamma^2, \cdots, \gamma^m$, or of $\xi^1, \xi^2, \cdots, \xi^s$, that do not, we introduce corresponding multipliers explicitly). This implies that the physical transformations given by D_a preserve the volume element, and so have unity for a multiplier.[11] Their commutators then also possess this property, as does a subclass of the complete set they determine. (Members of the subclass are of the form $\psi^a D_a$, where $D_a\psi^a = 0$.)

In this theory the distinction between dynamical variables and parameters describing observes becomes almost nonexistent. The operators

$$\eta_a{}^k \frac{\partial}{\partial r^k} - L_a{}^\alpha \frac{\partial}{\partial \xi^\alpha}, \tag{2.13}$$

with $L_a{}^\alpha$ functions of $r^1, r^2, \cdots, r^m, \xi^1, \xi^2, \cdots, \xi^s$, are such that a point transformation from $r^1, r^2, \cdots, r^m, \xi^1, \xi^2, \cdots, \xi^s$ to $\gamma^1, \gamma^2, \cdots, \gamma^m, \xi^1, \xi^2, \cdots, \xi^s$ takes them to

$$-L_a{}^\alpha \frac{\partial}{\partial \xi^\alpha}, \tag{2.14}$$

where $L_a{}^\alpha$ are the transformed functions of $\gamma^1, \gamma^2, \cdots, \gamma^m, \xi^1, \xi^2, \cdots, \xi^s$. The operators (2.14) certainly form a complete set, since they comprise s operators in s differentiations, so (2.13) must be a complete set, and if $\eta_a{}^k$ are functions of $r^1, r^2, \cdots, r^m$ only, D_a form a complete set. The requirement of preserving volume may easily be introduced.

Reversely, if we can introduce s variables into the operators D_a, as in (2.13), so that the new operators are still a complete set, and if we express the variables $r^1, r^2, \cdots, r^m$ in terms of the invariants of this set and of the new s variables, we have a solution of the equations of motion corresponding to D_a. If we take these s variables to be some among those used to define D_a, for example those describing one particle, the situation is similar except that the variables now corresponding to ξ^α enter $\eta_a{}^k$, representing reaction of this particle on the rest of the system. This will not prevent a point transformation from $r^1, r^2, \cdots, r^m, \xi^1, \xi^2, \cdots, \xi^s$ to $\gamma^1, \gamma^2, \cdots, \gamma^m, \xi^1, \xi^2, \cdots, \xi^s$ being found, with the same degree of arbitrariness as before.

3. THE RIEMANNIAN SPACE OF GENERAL RELATIVITY

We now assume Einstein's principle of relativity, perhaps slightly strengthened. The ten-parameter family of observers breaks up into a four-parameter set of six-parameter groups of observers with various orientations and velocities at the point-events of four-dimensional space-time.[12] Each of these six-parameter groups is to be a realization of the homogeneous Lorentz group in transformations of the dynamical variables, so that the corresponding physical displacements form a six-vector. We suppose further that the four remaining physical displacements to neighboring subsets form a four-vector.

Thus we divide ξ^α into two sets; the first we call ξ^α, $\alpha = 1, 2, 3, 4$; they specify point-events comprising six-parameter groups of observers, and are subject to arbitrary transformations involving the state of the universe; the second we call θ^λ, $\lambda = 1, 2, \cdots, 6$, specifying which observer of the six-parameter group at that point-event is involved; these are subject to arbitrary transformations involving the ξ^α and the state of the universe. Likewise we divide dx^a into two sets; the first we call dx^a, $a = 1, 2, 3, 4$, the four-vector of physical displacements to neighboring point-events, which neighboring point-event depending on the state of the universe; the second we call dx^{ab}, forming a six-vector, so that

$$dx^{ba} \equiv -dx^{ab}, \tag{3.1}$$

comprising infinitesimal rotations and velocity changes of the homogeneous Lorentz group.

Equations (1.1), (1.2), and (1.3) now take the form

$$\begin{aligned} d\xi^\alpha &= l_a{}^\alpha dx^a, & \alpha &= 1, \cdots, 4 \\ d\theta^\lambda &= L_a{}^\lambda dx^a + \tfrac{1}{2} L_{ab}{}^\lambda dx^{ab}, & \lambda &= 1, \cdots, 6 \\ dx^a &= l_\alpha{}^a d\xi^\alpha, & & \\ dx^{ab} &= l_\alpha{}^{ab} d\xi^\alpha + L_\lambda{}^{ab} d\theta^\lambda, & a, b &= 1, \cdots, 4 \end{aligned} \tag{3.2}$$

$$\frac{df}{ds} = \frac{\partial f}{\partial \xi^\alpha}\frac{d\xi^\alpha}{ds} + \frac{\partial f}{\partial \theta^\lambda}\frac{d\theta^\lambda}{ds} = \frac{\partial f}{\partial x^a}\frac{dx^a}{ds} + \frac{1}{2}\frac{\partial f}{\partial x^{ab}}\frac{dx^{ab}}{ds}, \tag{3.3}$$

and

$$\begin{aligned} \frac{\partial f}{\partial \theta^\lambda} &= \tfrac{1}{2} L_\lambda{}^{ab} \frac{\partial f}{\partial x^{ab}}, & \lambda &= 1, \cdots, 6 \\ \frac{\partial f}{\partial \xi^\alpha} &= l_\alpha{}^a \frac{\partial f}{\partial x^a} + \tfrac{1}{2} l_\alpha{}^{ab} \frac{\partial f}{\partial x^{ab}}, & \alpha &= 1, \cdots, 4 \\ \frac{\partial f}{\partial x^{ab}} &= L_{ab}{}^\lambda \frac{\partial f}{\partial \theta^\lambda}, & & \\ \frac{\partial f}{\partial x^a} &= l_a{}^\alpha \frac{\partial f}{\partial \xi^\alpha} + L_a{}^\lambda \frac{\partial f}{\partial \theta^\lambda}, & a, b &= 1, \cdots, 4. \end{aligned} \tag{3.4}$$

The coefficients omitted must vanish, and the factors $\frac{1}{2}$ occur because each is included twice in the summation.

Equations (2.1), the laws of dynamics, become

$$\begin{aligned} &\partial r^k/\partial x^a = \eta_a{}^k, \quad a, b = 1, \cdots, 4; \quad k = 1, 2, \cdots, m; \\ &\partial r^k/\partial x^{ab} = \eta_{ab}{}^k; \end{aligned} \tag{3.5}$$

[11] E. Goursat, *Mathematical Analysis* (Ginn and Company, Boston, 1917), Vol. 2, part 2, p. 82.

[12] Reference 1, p. 186.

and (2.2) and (2.3) become

$$D_a=\sum_{k=1}^{n}\eta_a{}^k\frac{\partial}{\partial r^k},$$
$$D_{ab}=\sum_{k=1}^{n}\eta_{ab}{}^k\frac{\partial}{\partial r^k},\tag{3.6}$$

and

$$\partial f/\partial x^a=D_af,$$
$$\partial f/\partial x^{ab}=D_{ab}f.\tag{3.7}$$

The assumption that the operators D_{ab} are physical displacement operators corresponding to the homogeneous Lorentz group allows us to choose them and the corresponding four-vector D_a, so that dx^1, dx^2, and dx^3 are orthogonal displacements in space, dx^4 a time displacement, dx^{12}, dx^{23}, dx^{31} rotations from dx^1 to dx^2, dx^2 to dx^3, and dx^3 to dx^1, while c^2dx^{14}, c^2dx^{24}, c^2dx^{34} are velocity displacements in directions dx^1, dx^2, and dx^3, c being the velocity of light. With this choice there goes consistently a fundamental tensor g_{ab} of the special form

$$g_{11}=g_{22}=g_{33}=1,\quad g_{44}=-c^2,\quad g_{ab}=0,\quad a\neq b,\tag{3.8}$$

and we use this and its inverse to raise and lower the indices a, b, $\cdots$, after the manner of general relativity theory.

The commutators of D_{ab} must be given by

$$D_{ab}D_{cd}-D_{cd}D_{ab}=g_{ac}D_{bd}-g_{bc}D_{ad}+g_{ad}D_{cb}-g_{bd}D_{ca},\tag{3.9}$$

and those of D_{ab} and D_c by

$$D_{ab}D_c-D_cD_{ab}=g_{ac}D_b-g_{bc}D_a.\tag{3.10}$$

More specially, operating on a vector function of dynamical variables

$$D_{ab}A_c=g_{ac}A_b-g_{bc}A_a,\tag{3.11}$$

with corresponding results for tensors.

We find that for these dynamical laws to fit into a Riemann space it is necessary and sufficient that the commutators of D_a have the form

$$D_aD_b-D_bD_a=\tfrac{1}{2}R^{cd}{}_{ab}D_{cd},\tag{3.12}$$

without any terms in D_a on the right-hand side. $R^{cd}{}_{ab}$ are physical components of the Riemann curvature tensor as functions of the dynamical variables.

To show that this is necessary we follow the text books.[13] We introduce the fundamental tensor of general relativity by

$$\gamma_{\alpha\beta}d\xi^\alpha d\xi^\beta\equiv g_{ab}dx^adx^b,\tag{3.13}$$

so that

$$\gamma_{\alpha\beta}=l_\alpha{}^al_\beta{}^bg_{ab}.\tag{3.14}$$

The physical displacements define an orthogonal ennuple, specified by $l_\alpha{}^a$ or their inverses $l_a{}^\alpha$. If we consider a vector (over the homogeneous Lorentz group) function of position and dynamical variables, given by A_b or A_α where $A_b=l_b{}^\alpha A_\alpha$, we have

$$D_aA_b=l_a{}^\alpha\frac{\partial}{\partial\xi^\alpha}A_b+\tfrac{1}{2}L_a{}^\lambda L_\lambda{}^{cd}D_{cd}A_b$$
$$=l_a{}^\alpha\frac{\partial}{\partial\xi^\alpha}(l_b{}^\beta A_\beta)+\tfrac{1}{2}L_a{}^\lambda L_\lambda{}^{cd}(g_{cb}A_d-g_{db}A_c),$$

which should agree with the covariant derivative in terms of Christoffel symbols,

$$A_{\beta,\,\alpha}=\frac{\partial A_\beta}{\partial\xi^\alpha}-A_\gamma\begin{Bmatrix}\gamma\\ \beta\ \alpha\end{Bmatrix},$$

giving us

$$\begin{Bmatrix}\gamma\\ \beta\ \alpha\end{Bmatrix}=-l_\beta{}^b\frac{\partial}{\partial\xi^\alpha}(l_b{}^\gamma)+l_c{}^\gamma L_\alpha{}^\lambda L_\lambda{}^{ac}g_{ab}l_\beta{}^b.\tag{3.15}$$

That $\begin{Bmatrix}\gamma\\ \beta\ \alpha\end{Bmatrix}$ is symmetrical in β and α turns out to be just the condition that when we form the commutator of D_a and D_b the terms in D_a cancel; in addition the coefficients of D_{cd} take the form (3.11).

The sufficiency follows from the general argument of Sec. 2. If we solve the dynamical equations, which form a complete set, the conditions (3.9) allow us to separate out the local Lorentz groups, and (3.10) and (3.12) make the geometry at which we arrive Riemannian.

If we regard the observers as specified by some of the dynamical variables of the system, we must arrive at just the above results, provided that we can neglect the reaction of the observer on the system. When this cannot be neglected, the curvature tensor will depend also on the observer.

Einstein's principle of the equivalence of gravitational and accelerational fields is now seen simply to state that if the physical displacement operators are given satisfying (3.9), (3.10), and (3.12), when expressed in terms of the dynamical variables, gravitation is already included, and requires no new field variables. We see in addition that the state relative to one observer gives, by solving the equations of motion, the state relative to any other observer. To set up statistical mechanics it is convenient to make the additional assumption that unity is a multiplier for each physical displacement operator.

4. CANONICAL FORM AND QUANTIZATION

Any set of linear differential operators can be put in classical canonical form by introducing a new variable for $\partial/\partial r^u$ to be canonically conjugate to r^u; the important question is whether they can be put in canonical form with fewer than $2m$ variables, since the new

[13] A. S. Eddington, *Mathematical Theory of Relativity* (Cambridge University Press, Cambridge, 1923), p. 217; L. P. Eisenhart, *Riemannian Geometry* (Princeton University Press, Princeton, New Jersey, 1926), p. 97.

variables are not physically observable. It is well known that a single operator can be put into canonical form without introducing extra variables,[14] and this theorem can be extended to any set of operators giving the infinitesimal transformations of a group.

For a general complete set of linear differential operators we must use the $2m$ variables, but we do not wish to consider all contact transformations of these $2m$ variables but only those that are extended point transformations[15] in the m variables $r^1, r^2, \cdots, r^m$. Our complete set of s linear operators together with the m variables $r^1, r^2, \cdots, r^m$, themselves now form a function group,[16] reducible in general to s pairs of canonically conjugate variables and $m-s$ singular variables. Indeed the singular variables are just the invariants and the reduction is equivalent to a solution of the equations of motion.

If the complete set were so special that it could be expressed in canonical form in terms of the m variables $r^1, r^2, \cdots, r^m$, only, the s functions representing the operators and at most half the remaining $m-s$ variables would have to form a function group of which the latter would be singular variables, so that the structure functions could involve only the first s functions and $(m-s)/2$ variables with zero Poisson brackets with them and with each other. This justifies the statement made at the beginning of Sec. 2.

In classical mechanics we have generalized from canonical transformations of dynamical variables to more general point transformations of these, a special class of canonical transformations in a "larger" space of twice as many dimensions. In quantum mechanics our dynamical variables and statistical states correspond to infinitesimal unitary transformations of the Hilbert space of wave functions; in fact to Hermitian matrices over this space. Now all the matrices over this space may be regarded as the vectors of a "larger" Hilbert space, the trace of the matrix product of two of them giving the scalar product of vectors of the "larger" space.[17] We shall regard unitary transformations of this "larger" space as corresponding to canonical transformations of the "larger" classical space. The extended point transformations will be those that maintain Hermitian matrices of the smaller space Hermitian, which are characterized in the larger space as commuting with a certain conjugation.[18] Finally, it is convenient to restrict our physical displacements to those infinitesimal transformations which maintain the unit matrix of the smaller space: this corresponds to having unity for a multiplier in the classical case.

The most general linear infinitesimal transformation of the matrices of the smaller space has the form

$$F'=F+\iota\epsilon\sum_r A_rFB_r, \tag{4.1}$$

where A_r and B_r are matrices. It is an infinitesimal unitary transformation of the larger space if it has the form

$$F'=F+\iota\epsilon\sum_r (A_rFB_r-B_rFA_r), \tag{4.2}$$

and it further maintains the Hermitian nature of F if A_r and B_r are Hermitian, while any such transformation that does this can be reduced to this form. Finally it maintains the unit matrix I if

$$\sum_r (A_rB_r-B_rA_r)=0. \tag{4.3}$$

[It can then be written in such forms as

$$F'=F+\tfrac{1}{2}\iota\epsilon\sum_r \{A_r(FB_r-B_rF)+(FB_r-B_rF)A_r - B_r(FA_r-A_rF)-(FA_r-A_rF)B_r\}.]$$

While transformations of this form do not apply to the original wave functions, and do not in general preserve the product of two matrices, they do preserve the trace of the product of any number of matrices, and therefore also the characteristic values of Hermitian matrices. Thus the statistical matrix corresponding to a wave function is transformed into a matrix that breaks up in the same way, though the phase of the resulting wave function is not determined.

By taking the point of view of the larger space, we see that what corresponds to multiplying an infinitesimal point transformation in classical theory by a function of the variables is here replacing a term A_rFB_r by a sum of terms

$$\sum_s \{(C_sA_r+A_rC_s)F(D_sB_r+B_rD_s) + (D_sB_r+B_rD_s)F(C_sA_r+A_rC_s)\},$$

C_s and D_s being Hermitian matrices. This leads immediately to a definition of a linear combination of infinitesimal transformations with functions of dynamical variables as coefficients. This preserves the form of (4.2) but not (4.3), just as in classical mechanics the corresponding operation does not leave the infinitesimal transformation in general with a unit multiplier. We can then define a complete set of infinitesimal transformations as a set for which the alternants of any two members of the set are linear combinations of the members of the set, with dynamical variables as coefficients. I then conjecture that all such transformations that leave unaltered the matrices that the set leaves unaltered are linear combinations of the members of the set.

These sets are expected to have enough of the properties of the complete sets of the classical theory to allow us to extend that theory. Even if the difficulty that our observers and test particles are defined by q numbers rather than c numbers makes it impossible to set up a Riemann space except in the classical limit, that would perhaps be sufficient.[19] We see that classical theories extended in this way can be quantized, and that quantum theory can be extended in general to allow of the wider range of transformations.

[14] The theorem of Lie and Konigs. See reference 5, p. 275.
[15] Reference 9, p. 85.
[16] Reference 9, p. 283.
[17] M. H. Stone, *Linear Transformations of Hilbert Space* (American Mathematical Society, New York, 1932), p. 67.
[18] Reference 17, p. 357.
[19] Perhaps Everett's " 'Relative State' Formulation of Quantum Mechanics" may help. H. Everett, III, Revs. Modern Phys. **29**, 454 (1957).