REVIEWS OF MODERN PHYSICS VOLUME 17, NUMBERS 2 AND 3 APRIL–JULY, 1945

Relativistic Invariance

L. H. THOMAS *
Department of Physics, Ohio State University, Columbus, Ohio

1. INTRODUCTION: "OBSERVERS" AND "STATES"

IT is usually convenient to treat space and time unsymmetrically, not only in the formulation either of classical or of quantum dynamics, but also in the transition from one to the other. The first part (kinematics) of a dynamical theory usually defines a state of the system at a given time by a description in terms of positions in space at that time. Since simultaneity depends on the motion of the observer the description is relative to one or other of a set of observers at relative rest whom we regard as existing at a series of times. The relations between these observers may be visualized as between sets of rectangular Cartesian axes in three dimensions with the help of which they describe the state, and the change of description of a state from one to another is given explicitly. The second part (dynamics) of the theory then tells, by differential, or integro-differential, equations how the state changes with time, the proper time of any one of these observers.

To discuss the invariance of a theory over observers in uniform relative motion, it is convenient to adopt a different point of view, treating space and time as far as possible symmetrically, at any rate at first. We regard each observer of the usual description as a series of observers relatively displaced in time; the relations between these may be visualized as between sets of rectangular axes and time in a three plus one dimensional Minkowski space. We adopt also another conception of state. Each observer may prepare a state according to some specification, and the state may be observed by the same or by any other observer. The change with time of the state of the system in the usual formulation now becomes a change of the specification of the state from an earlier to a later observer; on a par with the change in description in the usual formulation from an observer to a simultaneous observer with a different position or orientation in space, which change is left unaltered; and we shall talk in the same way about a change in the specification of a state from one observer to another in relative motion. Whenever that part of our dynamical laws giving the change of specification to neighboring observers at relative rest at one time is so simple as to be immediately integrable, we can return at will to the usual unsymmetrical formulation.

In the general theory of relativity simultaneity becomes arbitrary. The relations between our observers may be visualized as between sets of orthogonal axes of space and time at the points of a four-dimensional continuum. There seems to be no reason in principle to assume that sets of observers at relative rest at one time exist such that the change of specification from one to another can be given explicitly. A literal interpretation of the principle of relativity would lead us to expect that the change of specification to observers with different orientations and velocities at the same place and time could be given immediately, while the change in specification from an observer to others at different places and times would be given by differential or integro-differential equations. If this were not the case we would have integro-differential equations for the change of specification from one observer to another of a ten parameter family of observers which might or might not fall naturally into a four parameter set of six parameter sub-families.

2. DISPLACEMENTS

It is convenient to describe the observers by a set of values of ten arbitrary parameters $\alpha, \beta, \cdots$ and to use the methods of the general theory of relativity, later specializing for that theory and again for the homoloidal Minkowskian case.

The observers "near" a given observer 0 may be specified as $0+(dx, dy, dz; dt; dl, dm, dn; du, dv, dw)$ with displacement from 0 with

* On leave of absence at Aberdeen Proving Ground since 1943.

components dx, dy, dz, in space, dt in time, dl, dm, dn, in rotation, y to z, z to x, and x to y, and du, dv, dw, in velocity in the x, y, and z directions, relative to 0's axes. These displacements dx, $\cdots$, are to be treated as differentials corresponding to quasi-coordinates, i.e.,

$$\begin{aligned} dx&=\left(\frac{\partial x}{\partial\alpha}\right)d\alpha+\left(\frac{\partial x}{\partial\beta}\right)d\beta+\cdots,\\ dy&=\left(\frac{\partial y}{\partial\alpha}\right)d\alpha+\left(\frac{\partial y}{\partial\beta}\right)d\beta+\cdots,\end{aligned}\tag{2.11}$$

solved by

$$\begin{aligned} d\alpha&=\left(\frac{\partial \alpha}{\partial x}\right)dx+\left(\frac{\partial \alpha}{\partial y}\right)dy+\cdots,\\ d\beta&=\left(\frac{\partial \beta}{\partial x}\right)dx+\left(\frac{\partial \beta}{\partial y}\right)dy+\cdots,\end{aligned}\tag{2.12}$$

where

$$\left(\frac{\partial x}{\partial\alpha}\right),\ \left(\frac{\partial x}{\partial\beta}\right),\ \cdots$$
$$\left(\frac{\partial y}{\partial\alpha}\right),\ \left(\frac{\partial y}{\partial\beta}\right),\ \cdots$$

are functions of α, β, $\cdots$ forming a matrix of non-zero determinant, and

$$\left(\frac{\partial \alpha}{\partial x}\right),\ \left(\frac{\partial \alpha}{\partial y}\right),\ \cdots$$
$$\left(\frac{\partial \beta}{\partial x}\right),\ \left(\frac{\partial \beta}{\partial y}\right),\ \cdots$$

form the reciprocal matrix; but $(\partial x/\partial\alpha)\cdots$ do *not* in general satisfy the equations

$$\frac{\partial}{\partial\beta}\left(\frac{\partial x}{\partial\alpha}\right)=\frac{\partial}{\partial\alpha}\left(\frac{\partial x}{\partial\beta}\right),\tag{2.2}$$

so that Eqs. (2.11) do not in general determine true parameters x, y, $\cdots$.

We define, however, for x, y, z, t, l, m, n, u, v, w,

$$\frac{\partial f}{\partial x}=\left(\frac{\partial\alpha}{\partial x}\right)\frac{\partial f}{\partial\alpha}+\left(\frac{\partial\beta}{\partial x}\right)\frac{\partial f}{\partial\beta}+\cdots\tag{2.3}$$

with like definitions for other differential coefficients that we shall later introduce, *viz.*,

$$\frac{df}{dx}=\left(\frac{\partial\alpha}{\partial x}\right)\frac{df}{d\alpha}+\left(\frac{\partial\beta}{\partial x}\right)\frac{df}{d\beta}+\cdots\tag{2.31}$$

and

$$D_xf=\left(\frac{\partial\alpha}{\partial x}\right)D_\alpha f+\left(\frac{\partial\beta}{\partial x}\right)D_\beta f+\cdots,\tag{2.32}$$

so that, reversely,

$$\frac{\partial f}{\partial\alpha}=\left(\frac{\partial x}{\partial\alpha}\right)\frac{\partial f}{\partial x}+\left(\frac{\partial y}{\partial\alpha}\right)\frac{\partial f}{\partial y}+\cdots,\tag{2.4}$$

$$\frac{df}{d\alpha}=\left(\frac{\partial x}{\partial\alpha}\right)\frac{df}{dx}+\left(\frac{\partial y}{\partial\alpha}\right)\frac{df}{dy}+\cdots,\tag{2.41}$$

$$D_\alpha f=\left(\frac{\partial x}{\partial\alpha}\right)D_xf+\left(\frac{\partial y}{\partial\alpha}\right)D_yf+\cdots.\tag{2.42}$$

The coefficients $(\partial x/\partial\alpha)$, $\cdots$ may be regarded as describing the topology of the ten-dimensional observer space in a manner somewhat similar to that in which the coefficients $g_{\mu\nu}$ describe the topology of Riemann space, or more nearly similar to that in which functions defining an orthogonal ennuple describe it. We regard $\partial/\partial x$, $\cdots$, as having an intrinsic meaning which $\partial/\partial\alpha$, $\cdots$, relative to arbitrary changes of the arbitrary parameters, do not have.

3. THE DISPLACEMENT OPERATORS

A state may be designated by parameters $(u, v, \cdots)$, say, determining the independent elements $(a, b, \cdots)$ of its specification relative to any observer.

$$\begin{aligned} a&=a(u, v, \cdots; \alpha, \beta, \cdots)\\ b&=b(u, v, \cdots; \alpha, \beta, \cdots).\end{aligned}\tag{3.1}$$

The various observers may set up apparatus according to various specifications, and make observations on the same state (or on states prepared according to the same specifications relative to some particular observer). The results of these observations, one or more numbers, "measures," which may be statistical averages, will be functions of a, b, $\cdots$ the specification of the state relative to the observer, and also of α, β, $\cdots$, because the specification of the apparatus for observation may vary from observer to observer, and because the kinematics may vary from observer to observer,

$$f(a, b, \cdots; \alpha, \beta, \cdots).$$

For a given state $(u, v, \cdots)$, the total rate of change of the measure with the observer, $df/d\alpha$, is the sum of two terms:—the partial rate of change $\partial f/\partial\alpha$ caused by the change in the measure observed for a fixed specification of the state relative to the observer, $a, b, \cdots$; and the rate of change which we shall denote by $D_\alpha f$ caused by the change in specification of the state from observer to observer. Thus

$$D_\alpha f=\frac{\partial f}{\partial a}\frac{da}{d\alpha}+\frac{\partial f}{\partial b}\frac{db}{d\alpha}+\cdots \quad (3.2)$$

and

$$\frac{df}{dx}=D_\alpha f+\frac{\partial f}{\partial\alpha},\quad \text{etc.,} \quad (3.3)$$

so that also

$$df/dx=D_x f+\partial f/\partial x. \quad (3.31)$$

We may call $D_x f dx$ the change due to displacement dx and D_x the displacement operator. (These operators all obey the law of differentiation of a product in the form

$$D(fg)=(Df)g+g(Df). \quad (3.21)$$

Here $\partial/\partial a$, $\partial/\partial b$, $\cdots$, $\partial/\partial\alpha$, $\cdots$, are for $a, b, \cdots, \alpha, \cdots$ as independent variables; $d/d\alpha, \cdots$ are for $u, v, \cdots, \alpha, \cdots$ as independent variables.

4. EQUIVALENCE

The functions

$$g(a, b, \cdots; \alpha, \beta, \cdots)$$

giving the measures observed with various setups for states with various specifications comprise our kinematics, shrunken in this point of view to the connections between the specifications and measures relative to the same observer. If our observers are equivalent in the sense of the principle of equivalence of general relativity theory,** i.e., if the kinematics is independent of the observer, these functions will be independent of $\alpha, \beta, \cdots$, and it has definite meaning to say that g is the measure of a dynamical variable explicitly independent of the observer (e.g., angular momentum relative to the observer's frame), for which

$$\partial g/\partial\alpha=0, \cdots, \quad (4.1)$$

i.e.,

$$dg/d\alpha=D_\alpha g, \cdots; \quad (4.2)$$

or

$$dg/dx=D_x g, \cdots. \quad (4.3)$$

In this case any sufficient number of the measures g can be used to specify a state relative to an observer, and the states specified in the same way relative to different observers can justly be called displaced states. We can then give an intrinsic definition of the operator D_x; it is just the differential operator giving the change from a state relative to one observer to the state with the same specification relative to an observer displaced in the x direction. If the principle of equivalence does not hold, the operator D_x will depend on what measures were chosen to specify states.

Our dynamical laws which give dg/dx etc., in general or da/dx, db/dx, $\cdots$ in particular in terms of $a, b, \cdots, \alpha, \beta, \cdots$ for changes to the various observers neighboring a given observer are expressed by the form of the operator D_x acting on a function of $a, b, \cdots$. The form of expression of D_x may vary considerably with the nature of the aggregate of $a, b, \cdots$.

This $d/dx\cdots$ for $\partial/\partial\alpha=0, \cdots$, corresponds to the rate of change with time of the dynamical variables in classical mechanics or in the Heisenberg picture in quantum mechanics; and to the rate of change with rotation in the coordinates relative to fixed axes of a point moving with a rotating body, when the fixed axes are those with which the axes moving with the body momentarily coincide.

5. INTEGRALS

Alternatively the dynamical laws may be regarded as determining functions $h(a, b, \cdots; \alpha, \beta, \cdots)$ of variables $a, b, \cdots$, specified relative to the observer $\alpha, \beta, \cdots$, such that

$$dh/d\alpha=0, \cdots, \quad (5.1)$$

and

$$\partial h/\partial\alpha=-D_\alpha h; \quad (5.2)$$

or

$$\partial h/\partial x=-D_x h \quad (5.3)$$

** The principle of equivalence of gravitational fields to those due to an acceleration of the frame of reference, used in Einstein's general theory of relativity, can be regarded as a special case of the equivalence of observers in the sense considered here, that the kinematics is the same for them all.

(e.g., $h=u(a, b, \cdots; \alpha, \beta, \cdots)$ etc., obtained from (3.1)). Such a function may be said to be totally independent of the observer and to determine an integral of the dynamical equations. Equations (5.3) instead of (4.3) now give our dynamical laws when the form of the operator D_x acting on a function of $a, b, \cdots$, is given, and D_x may be defined intrinsically as the negative rate of change of the specification of a fixed state relative to an observer, or to the limiting ratio to dx of the change of specification of some quantity from an observer 0 to an observer 0′ relative to whom 0 is displaced distance dx in the x direction.

This $\partial h/\partial x$ for $dh/d\alpha=0, \cdots$, corresponds to the rates of change with time of the distribution functions of classical statistical mechanics, or of the Schrödinger functions and statistical matrices in the Schrödinger picture in quantum mechanics.

6. THE CONSISTENCY CONDITIONS

In order that such a set of dynamical laws should be consistent it is necessary and sufficient that we should come back to the same values after traversing a circuit of observers.†

The differential conditions of consistency may be found either from the form (4) or from the form (5). Using the latter, we must have

$$\frac{\partial}{\partial\beta}\frac{\partial h}{\partial\alpha}-\frac{\partial}{\partial\alpha}\frac{\partial h}{\partial\beta}=0; \tag{6.1}$$

and since

$$\begin{aligned}\frac{\partial}{\partial\beta}\frac{\partial h}{\partial\alpha}&=-\frac{\partial}{\partial\beta}(D_\alpha h)\\&=-\left(\frac{\partial}{\partial\beta}D_\alpha\right)h-D_\alpha\frac{\partial h}{\partial\beta}\\&=-\left(\frac{\partial}{\partial\beta}D_\alpha\right)h+D_\alpha D_\beta h,\end{aligned}$$

we obtain

$$(D_\alpha D_\beta-D_\beta D_\alpha)h=\frac{\partial}{\partial\beta}(D_\alpha)h-\frac{\partial}{\partial\alpha}(D_\beta)h.$$

Thus

$$\begin{aligned}&(D_xD_y-D_yD_x)h\\&=\sum_\alpha\sum_\beta\left(\frac{\partial\alpha}{\partial x}D_\alpha\frac{\partial\beta}{\partial y}D_\beta-\frac{\partial\beta}{\partial y}D_\beta\frac{\partial\alpha}{\partial x}D_\alpha\right)h\\&=\sum_\alpha\sum_\beta\frac{\partial\alpha}{\partial x}\frac{\partial\beta}{\partial y}(D_\alpha D_\beta-D_\beta D_\alpha)h\\&=\sum_\alpha\left(\frac{\partial\alpha}{\partial x}\frac{\partial D_\alpha}{\partial y}-\frac{\partial\beta}{\partial y}\frac{\partial D_\beta}{\partial x}\right)h\\&=\left[\frac{\partial}{\partial y}(D_x)-\frac{\partial}{\partial x}(D_y)\right]h\\&\qquad+\sum_\alpha\sum_z\left[\frac{\partial\alpha}{\partial x}\frac{\partial}{\partial y}\left(\frac{\partial z}{\partial\alpha}\right)-\frac{\partial\beta}{\partial y}\frac{\partial}{\partial x}\left(\frac{\partial z}{\partial\beta}\right)\right]D_zh;\end{aligned}$$

i.e.,

$$D_xD_y-D_yD_x=\frac{\partial D_x}{\partial y}-\frac{\partial D_y}{\partial x}+\sum_z C_{xy}{}^zD_z. \tag{6.2}$$

If the dynamical laws are independent of the observer, $\partial D_x/\partial y, \cdots$, vanish and $C_{xy}{}^z$ are constants, so that

$$D_xD_y-D_yD_x=\sum_z C_{xy}{}^zD_z, \tag{6.3}$$

and $D_x, \cdots$, are the infinitesimal operators of a group with $C_{xy}{}^z$ for structure constants.††

The terms $\partial D_x/\partial y\cdots$ enter in practice if we wish to consider invariance in the presence of given external fields the specification of which varies in a given way from observer to observer, without actually including in D_x the operators to carry out this variation.

For ordinary general relativity fifteen of the forty-five relations (6.2) reduce to

$$\begin{aligned}&D_mD_n-D_nD_m=D_l, && D_mD_w-D_wD_m=D_u,\\&D_vD_n-D_nD_v=D_u, && D_lD_u-D_uD_l=0,\\&D_nD_l-D_lD_n=D_m, && D_nD_u-D_uD_n=D_v,\\&D_wD_l-D_lD_w=D_v, && D_mD_v-D_vD_m=0,\\&D_lD_m-D_mD_l=D_n, && D_lD_v-D_vD_l=D_w,\\&D_uD_m-D_mD_u=D_w, && D_nD_w-D_wD_n=0;\end{aligned}$$

$$\begin{aligned}D_vD_w-D_wD_v&=-\frac{1}{c^2}D_l,\\D_wD_u-D_uD_w&=-\frac{1}{c^2}D_m,\\D_uD_v-D_vD_u&=-\frac{1}{c^2}D_n,\end{aligned} \tag{6.4}$$

† A more exact statement would be "the same physical values after traversing a circuit of observers;" the phase of a Schrödinger function need not be the same, but we do not regard it as physical.

†† The relations between the $C_{xy}{}^z$ necessary for this will follow from their expression in terms of $\partial\alpha/\partial x$ and $\partial z/\partial\alpha$.

and the observers fall into a four parameter set of six parameter sub-families, each sub-family being a realization of the homogeneous Lorentz group, for which D_l, D_m, D_n, D_u, D_v, D_w, form a set of infinitesimal operators.

If one observer is picked from each sub-family in a continuous manner, an orthogonal ennuple in four parameter space-time is provided, and the remaining operators D_x, D_y, D_z, and D_t, provide the rates of change of the specification of a state with changes of the observer along the curves of the ennuple. It is this situation that may be regarded as guaranteed by the physical principles of relativity and equivalence.

For flat Minkowski space the remaining thirty relations are

$$\begin{aligned}
&D_mD_z-D_zD_m=D_x, \quad D_yD_n-D_nD_y=D_x,\\
&D_lD_x-D_xD_l=0, \quad D_lD_t-D_tD_l=0,\\
&D_nD_x-D_xD_n=D_y, \quad D_zD_l-D_lD_z=D_y,\\
&D_mD_y-D_yD_m=0, \quad D_mD_t-D_tD_m=0,\\
&D_lD_y-D_yD_l=D_z, \quad D_xD_n-D_nD_x=D_z,\\
&D_nD_z-D_zD_n=0, \quad D_nD_t-D_tD_n=0;\\
&D_vD_z-D_zD_v=0, \quad D_yD_w-D_wD_y=0,\\
&D_uD_x-D_xD_u=\frac{1}{c^2}D_t, \quad D_uD_t-D_tD_u=D_x,\\
&D_wD_x-D_xD_w=0, \quad D_zD_u-D_uD_z=0,\\
&D_vD_y-D_yD_v=\frac{1}{c^2}D_t, \quad D_vD_t-D_tD_v=D_y,\\
&D_uD_y-D_yD_u=0, \quad D_xD_v-D_vD_x=0,\\
&D_wD_z-D_zD_w=\frac{1}{c^2}D_t, \quad D_wD_t-D_tD_w=D_z;\\
&D_yD_z-D_zD_y=0, \quad D_xD_t-D_tD_x=0,\\
&D_zD_x-D_xD_z=0, \quad D_yD_t-D_tD_y=0,\\
&D_xD_y-D_yD_x=0, \quad D_zD_t-D_tD_z=0.
\end{aligned} \tag{6.5}$$

For changes to observers at rest at the same time and place, differing only in orientation, the invariance of our equations will follow at once from their forms when expressed in ordinary three-dimensional vector notation, which we, therefore often use. It is then necessary to verify only a selection of the relations (6.4), (6.5), for example

$$D_xD_t-D_tD_x=0, \quad D_xD_y-D_yD_x=0,$$

$$D_uD_x-D_xD_u=\frac{1}{c^2}D_t, \quad D_xD_v-D_vD_x=0,$$

$$D_uD_t-D_tD_u=D_x,$$

and

$$D_uD_v-D_vD_u=-\frac{1}{c^2}D_n. \tag{6.6}$$

Verification of invariance would be still simpler if we made invariance over the homogeneous Lorentz group obvious from the form; it would be then only necessary to verify

$$D_xD_t-D_tD_x=0; \tag{6.7}$$

but the connection with the usual unsymmetrical formulation would be much less direct, involving solution of the equations of motion. On the other hand, this formulation would be much better for introducing the general theory of relativity.

SUMMARY

It is convenient and usual in the formulation both of classical and of quantum dynamics and especially in the transition from one to the other to treat the time variable in a special manner so that invariance for the Lorentz transformations of the restricted theory of relativity is not at all obvious. Further generalization to include general relativity and a gravitational field then becomes difficult.

An analysis of the situation sufficiently wide to cover quantum dynamics as well as classical dynamics from the point of view of a ten-parameter family of observers "equivalent" in the sense of general relativity unifies the whole treatment and serves to show that "general relativity" in which the ten parameter family falls into a four parameter set of six parameter sub-families is by no means the most general case.

REFERENCES

1. Dirac, *Quantum Mechanics*, first edition (quantum mechanics in relativistic form).
2. Dirac, *Quantum Mechanics*, second edition (quantum mechanics in non-relativistic form).
3. Eisenhart, *Theory of Continuous Groups* (tensor analysis applied to groups of transformations).
4. Routh, *Advanced Rigid Dynamics* (moving axes formulas).