A QUANTUM DYNAMICAL, RELATIVISTICALLY INVARIANT RIGID BODY SYSTEM(1)

BY
RICHARD ARENS

1. **Introduction.** The purpose of this work is to propose and investigate a quantum dynamical, relativistically invariant system $\mathfrak{A}$ intended to describe the dynamics of a symmetric rigid body. The Hilbert space associated with $\mathfrak{A}$ is the space $L_2(E(3))$ of square-integrable functions on the ordinary Euclidean group $E(3)$. The (pure) states of $\mathfrak{A}$ are, of course, the rays or one-dimensional subspaces of this Hilbert space.

For any dynamical system one has to specify each dynamorphism or change in state corresponding to a change from one observer to another. For an *invariant* system, this gives rise to a representation (or, more generally, an *action*) of the space-time group. In the relativistic case this is the Poincaré group $\mathscr{P}$. We exhibit this representation of $\mathscr{P}$ in the case of $\mathfrak{A}$. It is, as it should be, an extension to $\mathscr{P}$ of the regular representation of $E(3)$.

In any system (classical or quantum) purporting to describe a rigid body, the orthogonal group $O(3)$ must act in the space of states in a manner induced by the action in the configuration space $E(3)$ which reflects the fact that the rigid body can be orthogonally transformed about its centroid without subjecting the centroid to the motion which the inclusion of $O(3)$ in $E(3)$ defines by left multiplication.

The symmetry of the body means precisely that this action commutes with all dynamorphisms. Accordingly one can construct a new system in which the states are the orbits under this action of $O(3)$. In the case of $\mathfrak{A}$, the new system may be denoted by $\mathfrak{A}/O(3)$. It turns out to be a quantum system. It is indeed the direct sum of the well-known positive energy systems with spin $0, 1/2, 1, 3/2, \ldots$. Thus the positive energy Dirac system is a constituent of $\mathfrak{A}/O(3)$.

The system $\mathfrak{A}$ has also a classical limit $\mathfrak{A}_0$. The system $\mathfrak{A}_0$ has as its space of states the phase space formed in the usual way when the configuration space is $E(3)$. Each of its dynamorphisms preserves the Poisson bracket. It is thus a Hamiltonian dynamical system for a rigid body which is Lorentz invariant. It is of course not the usual one due to Euler which is Galilei invariant.

For the classical system one can also form $\mathfrak{A}_0/O(3)$. This is a system in our general sense, and its state-space has dimension $2 \dim E(3) - \dim O(3) = 12 - 3 = 9$. It has a Poisson bracket and the infinitesimal dynamorphisms have generating

---

functions, but the Poisson bracket is somewhat degenerate. Axioms for systems of this type (which includes the classical systems) are given.

It would go beyond the scope of the present work to explain in what precise sense $\mathfrak{A}_0/O(3)$ is the classical limit of $\mathfrak{A}/O(3)$, but it can be done and this gives rise to a diagram,

$$
\begin{array}{ccc}
\mathfrak{A} & \xrightarrow{\;h\,=\,0\;} & \mathfrak{A}_0 \\
\Big\downarrow {\scriptstyle \div\,O(3)} & & \Big\downarrow {\scriptstyle \div\,O(3)} \\
\mathfrak{A}/O(3) & \xrightarrow{\;h\,=\,0\;} & \mathfrak{A}_0/O(3).
\end{array}
$$

The content of the following sections is as follows. §2 recapitulates the notions of space-time, observes, and dynamorphisms both global and infinitesimal. §3 deals with Hamiltonian systems which among other things have a Poisson bracket which could be degenerate. It defines completely Hamiltonian systems (generalizing a suggestion of D. G. Babbitt) in which the infinitesimal dynamorphisms have generating functions. It is shown that one-particle systems based on the Einstein-Lorentz category of observers which are completely Hamiltonian, describe only free particles. This is not the case with Galilei-Newton observers. In order to permit comparison with the system $\mathfrak{A}_0$, the Galilei invariant Euler system of a rigid body is set up in a group-theoretic way. §4 deals with the classical system $\mathfrak{A}_0$, and shows that the differential equations (i.e., the infinitesimal dynamorphisms) can be integrated to give the necessary action of $\mathscr{P}$ in $E(3)$.

§5 introduces the general concept of completely Hamiltonian systems and shows that the process of dividing out the symmetry group $O(3)$ leads to a completely Hamiltonian system $\mathfrak{A}_0/O(3)$. §6 is purely mathematical and presents a way of deforming Galilean frames into Lorentz frames. This enables us to construct an action of the Lorentz group on $E(3)$. (For system $\mathfrak{A}_0$ we have an action of the Lorentz group on the cotangent bundle of $E(3)$. This action is not induced by the action of §6.) From there we obtain an extension to $\mathscr{P}$ of the regular representation of $E(3)$.

In §7 we construct the quantum system $\mathfrak{A}$. In §8 we divide out the action of $O(3)$ and show that the constituent of $\mathfrak{A}/O(3)$ with spin $1/2$ is the Dirac positive energy system. This is done by using the inverse of the Foldy-Wouthuysen transformation. Since we know no reference for any but the purely temporal dynamorphisms (generated by the Hamiltonian) for the F.-W. system, we calculate the rest of them in §9.

§10 makes precise the notion of the classical limit of quantum systems. This construction is the inverse of the process of quantizing a classical system and seems to avoid the difficulties encountered in the latter process. Then we show that $\mathfrak{A} \to \mathfrak{A}_0$ in the sense defined. The proof is complicated because not all the infinitesimal dynamorphisms of $\mathfrak{A}$ are differential operators.

§11 is a brief disclosure of the local structure of the Poisson bracket axiomatized for the "neo-classical" Hamiltonian systems of §3 in the presence of a condition of constancy of rank.

It is a pleasure to acknowledge helpful discussions with D. G. Babbitt and V. S. Varadarajan on all facets of this work. Also stimulating was a perusal of the former's *Lecture notes on special relativity—a coordinate free approach.*

2. **Coordinators in space-time, and dynamics.**   In connection with manifolds M we use the notation set forth in [2, §IV]. In particular, if ξ is a vector in M

$$\xi = \xi^i(\partial/\partial x^i) \qquad \text{(summation convention)}$$

thus given by components in some coordinate system, then for any (differentiable) function f defined in M, f is defined for vectors in M, by the formula

$$(2.01) \qquad f(\xi) = \xi^i(\partial f/\partial x^i) = \xi(f).$$

Let x, y, z, t be *the* cartesian coordinates in R^4. Then the Lorentz structure in R^4 can be defined by the quadratic (differential) form

$$(2.1) \qquad \dot{t}^2 - \dot{x}^2 - \dot{y}^2 - \dot{z}^2;$$

and the inhomogeneous Lorentz or Poincaré group $\mathscr{P}$ can be defined as those maps T of R^4 onto itself which preserve 2.1 in the sense that for any vector ξ in R^4, 2.1 has the same value for $T\xi$ as for ξ.

Space-time as conceived by Galileo and Newton involves two quadratic forms

$$(2.11) \qquad \dot{t}^2,$$

$$(2.12) \qquad \dot{x}^2 + \dot{y}^2 + \dot{z}^2.$$

We will say that a manifold M with a finite sequence

$$(2.2) \qquad g, g', \dots$$

of one or more quadratic forms has a *space-time structure.*

Now we want to give a definition of space-time isomorphism which, in case of 2.1, singles out the Poincaré (=inhomogeneous Lorentz) group and in case of 2.11–2.12 singles out the Galilean group (see, for example [5, 1346]); and we would like this definition to make sense even when M is not R^4 so that we avoid saying "linear".

A *geodesic* with respect to a set 2.2 is a curve in M such that for each arc thereof the "pseudo-length" of it with respect to each of these quadratic forms is stationary. In the special cases these are just straight lines.

A vector ξ in M is *singular* with respect to g if $g(\xi + \eta) = g(\xi)$ for every vector η with the same base point as ξ.

Now let N be another manifold with space-time structure $h, h', \dots$. Let T be a differentiable mapping of M into N such that

2.21 *geodesics are mapped into geodesics,*

2.22 $h(T\xi)=g(\xi)$ *for every vector* ξ *in* M,

2.23 *if* ξ *is singular relative to* g,

then $h'(T\xi)=g'(\xi)$, $h''(T\xi)=g''(\xi)$, ... *for all the remaining quadratic forms in* 2.2.

Such a map will be called a space-time map. If T and T^{-1} are both space-time maps we call T a space-time isomorphism.

We leave it to the reader to show that this definition does single out the desired groups in the familiar cases mentioned. We call the group of all space-time iso-morphisms the *group* of the space-time structure.

Both structures 2.1 and (2.11, 2.12) have the following property. A space-time structure $g, g', \ldots$ in R^4 is called *Euclidean* if the maps obtained by extending the Euclidean motions in the obvious and trivial way to R^4 are included in the $(g, g', \ldots)$ group.

We now prepare for the definition of the analogue of an "observer", which will play a role similar to that of the space-like hyperplane in the theory of dynamical systems in [7], and to that of observer in [6]. A vector ξ in M is called *time-like* if $g(\xi)$ is positive. A one-dimensional connected smooth submanifold ("curve") W in M is called time-like if all its tangents are time-like vectors. A time-like curve W shall be called a *world line* if it is complete. These definitions are those given in [7, 2.2, 2.21, 2.22] except that the specifically Einstein-Lorentz structure is left out, to allow a simultaneous treatment of Galileo-Newton space-time.

A *space-like section* is a hypersurface S in M which has the property that every world line W intersects S in exactly one point and is not tangent to S.

We now postulate a space-time structure $(g, g', \ldots)$ in R^4 which will not be changed without further notice, and which has the properties

2.3 *the hyperplane* $t=0$ *is a space-like section*,

2.31 *the time translations* $(a, b, c, d) \to (a, b, c, d+\tau)$ *is an isomorphism for each real* τ.

We will call it the structure of R^4 for brevity.

Now let M be a manifold with a space-time structure isomorphic with that of R^4. Then any particular isomorphism $\mathbf{x}: M \leftrightarrow R^4$ shall be called a *coordinator*([2]) for M. A coordinator $\mathbf{x}$ thus consists of four real functions (x^1, x^2, x^3, x^4) defined in M. The coordinator $(x^1, x^2, x^3, x^4-\tau)$, where τ is a constant, is a temporal variant of $\mathbf{x}$. Supposing that τ is positive; for example, the latter coordinator assigns a smaller time coordinate to a given event than does the former.

The physical idea is that the coordinator (x^1, x^2, x^3, x^4) is concerned with recording events on or infinitesimally near the hypersurface in space-time M where $x^4=0$.

Now suppose there is some dynamical system involved. The observations of

---

([2]) We would call it an "observer" if we had not used this term in [6, 3.31] for a closely related, but somewhat different concept. Specifically we are now exploring the case in which the set Q in [7, 3.1] does not intentionally lie in M and moreover is the same for all "observers".

(x^1, x^2, x^3, x^4) should fix and determine which of its dynamically possible per-formances the system is actually doing. Hence the observations of any other coordinator (y^1, y^2, y^3, y^4) relating to that performance are determined by the data supplied by (x^1, x^2, x^3, x^4).

When there is a particular dynamical system involved, there will be an appro-priate space K at hand for each coordinator to mark his observations in, as a point (cf. phase space). This space K is the *kinematics* of the system (cf. [7]).

According to what has been said, there should, for each pair of coordinators $\mathbf{x}, \mathbf{y}$, be a 1:1 mapping $\Delta^{\mathbf{y}}_{\mathbf{x}}$ of K onto itself such that if a given performance provides the data ξ to $\mathbf{x}$ then that performance provides the data $\Delta^{\mathbf{y}}_{\mathbf{x}}(\xi)$ to $\mathbf{y}$. Each such mapping shall be called a *dynamorphism* and the collection of all these, denoted perhaps by Δ, shall be called the *dynamics*. The pair (K, Δ) may be used as a name for the system being discussed.

We impose two conditions on a *dynamics* Δ. These conditions reflect the physical idea. They are

2.4 $\Delta^{\mathbf{z}}_{\mathbf{y}} \circ \Delta^{\mathbf{y}}_{\mathbf{x}} = \Delta^{\mathbf{z}}_{\mathbf{x}}$, *for any three coordinators*, and

2.41 $\Delta^{\mathbf{x}}_{\mathbf{y}}$ *is the inverse of* $\Delta^{\mathbf{y}}_{\mathbf{x}}$, *for any two coordinators.*

We now proceed to a specific example, the case of *second-order interaction between n particles* (cf. [7, 2.6–2.7]). Having chosen a space-time structure we let B denote the class of *velocities* that a world line may have for any coordinator. For Einstein-Lorentz, this is of course those vectors of length less than 1 while for Galileo-Newton there is no restriction and B is linearly isomorphic to R^3. Let $\mathscr{I}$ be a class of n-tuples of world lines (in M) such that for any coordinator $\mathbf{x}$ and any points $\mathbf{a}_1, \ldots, \mathbf{a}_n$ in R^3 and any n velocities $\mathbf{v}_1, \ldots, \mathbf{v}_n$ selected from B there is one and only one n-tuple $(W_1, \ldots, W_n)$ in $\mathscr{I}$ such that their images $\mathbf{x}(W_1), \ldots, \mathbf{x}(W_n)$ in R^4 pass through the given points $\mathbf{a}_1, \ldots, \mathbf{a}_n$ and have the velocities[3] $\mathbf{v}_1, \ldots, \mathbf{v}_n$ respectively.

Thus for any $\mathbf{W} = (W_1, \ldots, W_n)$ in $\mathscr{I}$ on any coordinator $\mathbf{x}$ we obtain a "point" in $K = R^3 \times \cdots \times R^3 \times B \times \cdots \times B$. Let us call this point $\Delta^{\mathbf{x}}(\mathbf{W})$. For any pair of coordinators $\mathbf{x}$ and $\mathbf{y}$, let

(2.42) $\Delta^{\mathbf{x}}_{\mathbf{y}} = \Delta^{\mathbf{x}} \circ (\Delta^{\mathbf{y}})^{-1}$.

This certainly implies 2.4 and 2.41 and is thus the dynamorphism for this system.

REMARK. In this case the dynamorphism arose because we had a family $\{\Delta^{\mathbf{x}}\}$ of 1:1 maps of a fixed set $\mathscr{I}$ onto $\dot{K}$. We prefer not to take this as a basis for a general treatment of dynamics because it is generally hard to find a natural candidate for this space $\mathscr{I}$.

If $\mathbf{x}$ is a coordinator and S is an isomorphism of R^4 (or T is an isomorphism of M) then $S \circ \mathbf{x}$ (as well as $\mathbf{x} \circ T$) is another coordinator and every other coordinator

---

[3] The velocity of a curve C in R^4 is said to be $\mathbf{v}$ if it cuts $t = 0$ in one point $\mathbf{a}$ and the vector with components $(v^1, v^2, v^3, 1)$ is tangent to C at $\mathbf{a}$.

can be obtained in this way. For a coordinator $\mathbf{x}$ and any element S of the group G defined by the space-time structure presumed selected for R^4, we may define

$$(2.5) \qquad\qquad U_{\mathbf{x}}(S) = \Delta_{\mathbf{x}}^{S \circ \mathbf{x}}$$

This is not generally a representation of G. In fact,

$$(2.51) \qquad\qquad U_{\mathbf{x}}(S \circ S_0) = U_{\mathbf{x}}(S) \circ U_{\mathbf{x}}(S_0)$$

for all S_0 in G if and only if S is a *dynamical equivalence* of the system (K, Δ), i.e.

$$(2.52) \qquad\qquad U_{\mathbf{y}}(S) = U_{\mathbf{z}}(S) \quad \textit{for all } \mathbf{y} \textit{ and } \mathbf{z}.$$

This assertion can be established by repeated use of 2.4 and 2.41. In particular

2.53   *The group of dynamical equivalences is represented by each coordinator as a group of permutations in K.*

If every isomorphism of the space-time structure is a dynamical equivalence, then the dynamical system may be called *invariant*([4]).

In the classical and conventional treatments of mechanics [1, 14] those particular dynamorphisms $\Delta_{\mathbf{x}}^{\mathbf{y}}$ are emphasized in which $\mathbf{y}$ is a temporal variant of $\mathbf{x}$. This amounts to studying $U_{\mathbf{x}}(S)$ where S is the one-parameter group of time translations in R^4, and differential equations are introduced to govern the behavior of these one-parameter families (not groups!) of dynamorphisms. In order thus to govern more general dynamorphisms by differential equations we have to make additional assumptions.

2.6   *The group of M has a topology compatible with its group structure.*

2.61   *The set K has a differentiable structure, possibly infinite dimensional.*

2.62   *For each coordinator $\mathbf{x}$ and each continuous one-parameter subgroup T of the group, there is a vector field defined in([5]) K with the property that (in an appropriate sense)*

$$(2.63) \qquad\qquad \left.\frac{d}{ds} U_{\mathbf{x}}(T(s))\right|_{s=0} = \Delta'_{T,\mathbf{x}}.$$

When the dynamics is invariant, then these vector fields have to satisfy the commutation relations of the (Lie algebra of the) group G (see 2.51). It is not known what conditions (beyond smoothness conditions) ought to be imposed on the map $(T, \mathbf{x}) \to \Delta'_{T,\mathbf{x}}$.

These vector fields $\Delta'_{T,\mathbf{x}}$ may be referred to as the *infinitesimal dynamorphisms*. It is of interest to note to what extent they determine the dynamics. The topology 2.6 defines a topology in the space of coordinators. In the Einstein-Lorentz case we get a space homeomorphic to the Poincaré group. In principle, if $\mathbf{x}$ and $\mathbf{y}$ lie in

---

([4]) The principle of relativity would assert that a "natural" dynamical system should be invariant. We do not invoke this principle as part of the definition because that would exclude systems with "external forces".

([5]) That is, on a dense subset of K. The fact that the Hamiltonian operator of a quantum system is usually not defined on all of its Hilbert space justifies not requiring $\Delta'_{T,\mathbf{x}}$ to be defined on all of K.

the same component [15, 13], then the differential equations should serve to determine Δ_y^x. But additional information is needed to find Δ_y^x in the contrary case. Obviously one needs to know, besides the infinitesimal dynamorphisms, also one specimen of the set $\Delta_{I_s \circ \mathbf{x}}^x$ and one specimen of the set $\Delta_{I_t \circ \mathbf{x}}^x$ where I_s, I_t are the *inversions* [15, 10].

In the familiar cases, continuous one-parameter subgroups T can be characterized by the infinitesimal transformations Z which generate them in R^4, and instead of $\Delta_{T,\mathbf{x}}'$ we may write $\Delta_{Z,\mathbf{x}}'$.

2.7 THEOREM. *For a second-order n particle interaction, the infinitesimal dynamorphisms corresponding to the standard dynamorphism 2.42 have the following form. For each coordinator* $\mathbf{x}$ *there is a set of functions defined on K*

(2.71) $$\{A_{\mathbf{x},\lambda}^a : a = 1, 2, 3; \lambda = 1, 2, \ldots, n\}$$

such that for $Z = Z^0(\partial/\partial t) + Z^1(\partial/\partial x^1) + Z^2(\partial/\partial x^2) + Z^3(\partial/\partial x^3)$,

(2.72) $$-\Delta_{Z,\mathbf{x}}' = (Z_\lambda^0 v_\lambda^i - Z_\lambda^i)\frac{\partial}{\partial x_\lambda^i} + [Z_\lambda^0 A_{\mathbf{x},\lambda}^i + v_\lambda^i(Z_{0,\lambda}^0 + v_\lambda^j Z_{j,\lambda}^0) - Z_{0,\lambda}^i - v_\lambda^j Z_{j,\lambda}^i]\frac{\partial}{\partial v_\lambda^i}.$$

Here t, x^1, x^2, x^3 are the cartesian coordinates in R^4, $x_1^1, x_1^2, x_1^3, \ldots, x_\lambda^1, x_\lambda^2, x_\lambda^3, \ldots, x_n^1, x_n^2, x_n^3, v_1^1, v_1^2, v_1^3, \ldots, v_n^1, v_n^2, v_n^3$ are the coordinates in K and the summation convention applies to all repeated indices. Moreover Z_0^i means $\partial Z^i/\partial t$, Z_j^i means $\partial Z^i/\partial x^j$, and the suffix λ in every case means that (t, x^1, x^2, x^3) are replaced by $(0, x_\lambda^1, x_\lambda^2, x_\lambda^3)$, except for the $A_{\mathbf{x},\lambda}^i$ which were defined on K in the first place.

To prove 2.7 we may identify M with R^4, using $\mathbf{x}$. Let $\mathbf{y}$ be the coordinator (now a map of R^4 onto R^4) which makes a point flow according to Z for ε units of time. We have to compute $\Delta_{\mathbf{x}}^y$. According to 2.42 this means to take our initial conditions, transform them according to $\mathbf{y}$, find the appropriate curves from $\mathscr{I}$ and take their positions and velocities at $t = 0$. A little reflection shows that we need to think only of one particle, and that the dependence of the A's on all the variables takes care of the generalization.

Suppose our particle is at $\mathbf{a}$ with velocity $\mathbf{v}$ when $t = 0$. The tangent to the world line thus determined has components $(1, v^1, v^2, v^3)$. The $\mathbf{y}$ is a transformation which is approximately

(2.73) $$(t + \varepsilon Z^0, x^1 + \varepsilon Z^1, x^2 + \varepsilon Z^2, x^3 + \varepsilon Z^3).$$

This moves our point $(0, a^1, a^2, a^3)$ to

(2.74) $$(+ \varepsilon Z^0, a^1 + \varepsilon Z^1, \ldots, a^3 + \varepsilon Z^3)$$

where the Z's here are evaluated at $(0, \mathbf{a})$ (compare the meaning of the suffix λ). To obtain the transform of the tangent we use the ordinary rules for change of contravariant vector components, as effected by 2.73, and obtain, as direction numbers for the tangent,

$$(1 + \varepsilon X_0^0 + \varepsilon v^i Z_i^0, + \varepsilon Z_0^1 + \varepsilon v^i Z_i^1 + v^1, \ldots, + \varepsilon Z_0^3 + \varepsilon v^i Z_i^3 + v^3).$$

We divide by the first component and ignore ε^2 and obtain another equivalent tangent vector

$$(2.75) \qquad\qquad (1, \bar{v}^1, \bar{v}^2, \bar{v}^3)$$

where $\bar{v}^i = v^i - \varepsilon[v^i(Z_0^0 + v^j Z_j^0) - Z_0^i - v^j Z_j^i]$. Now 2.74, 2.75 provide the position and velocity at $t = \varepsilon Z^0$ (see 2.74) of a particle whose acceleration is, let us say, **A**. Hence at $t = 0$, its *position* will be

$$(a^1 + \varepsilon Z^1 - \varepsilon Z^0 \bar{v}^1, \ldots, a^3 + \varepsilon Z^3 - \varepsilon Z^0 \bar{v}^3)$$

or, ignoring ε^2,

$$(2.76) \qquad\qquad (a^1 + \varepsilon Z^1 - \varepsilon Z^0 v^1, \ldots, a^3 + \varepsilon Z^3 - \varepsilon Z^0 v^3);$$

and its *velocity* will be

$$(2.77) \qquad\qquad (\bar{v}^1 - \varepsilon Z^0 A^1, \ldots, \bar{v}^3 - \varepsilon Z^0 A^3).$$

The infinitesimal dynamorphism desired is simply the rate of change (with ε) of the position (2.76) and velocity (2.77). Taking into account the way ε enters into the $\bar{v}^i$ we see that

$$-\Delta'_{Z,\mathbf{x}} = (Z^0 v^1 - Z^1)(\partial/\partial x^1) + \cdots + (Z^0 v^3 - Z^3)(\partial/\partial x^3)$$
$$+ [Z^0 A^i + v^i(Z_0^0 + v^j Z_j^0) - Z_0^i - v^j Z_j^i](\partial/\partial v^i).$$

This is the assertion of 2.72 for one particle. Now each particle moves according to its own law, except for the way in which the other particles affect the A's. Thus 2.72 is established in general.

As a check, we take $Z = \partial/\partial t$. Then **y** is the transformation that translates upward in the t-direction by an amount ε. Let **a**, **v** be the initial conditions (for a particle) in the sense of **x**. Now **y** would say that this event is at $(+\varepsilon, \mathbf{a})$. With acceleration **A**, the position at $t = 0$ is $\mathbf{a} - \varepsilon\mathbf{v}$ and the velocity $\mathbf{v} - \varepsilon\mathbf{A}$. Thus

$$(2.78) \qquad\qquad -\Delta'_{\partial/\partial t, \mathbf{x}} = v^i(\partial/\partial x^i) + A^i(\partial/\partial v^i)$$

which agrees with 2.72.

3. Hamiltonian systems. Let K be a differentiable manifold, with an anti-symmetric contravariant tensor field A of the type whose components in any coordinate system $x^1, \ldots, x^n$ have two indices A^{ij}. Then we can form a product

$$(3.1) \qquad\qquad \{f, g\} \equiv A^{ij} f_i g_j \qquad \text{(summation convention)}$$

of any two functions f, g defined on open sets in K, where f_i, g_j are $\partial f/\partial x^i$ and $\partial g/\partial x^j$. In the familiar case this "product" satisfies the laws

$$(3.11) \qquad\qquad \{f, g\} = -\{g, f\},$$

$$(3.12) \qquad\qquad \{f, gh\} = \{f, g\}h + \{f, h\}g,$$

and the Jacobi-Lie identity:

$$(3.13) \qquad\qquad \{f, \{g, h\}\} + \{g, \{h, f\}\} + \{h, \{f, g\}\} = 0$$

on the common domain of these functions.

Property 3.13 does not follow from 3.1. A tensor field A satisfying these conditions 3.11–3.13 shall be called an *alternating structure*. These structures are analyzed in our appendix (§11 below).

Any function h defined on K defines a vector field X_h on K, via the formula

$$(3.14) \qquad\qquad\qquad X_h f = \{h, f\}.$$

X_h is a vector field by 3.12 and an infinitesimal contact transformation by 3.13, and h is called the *generating function* of X_h.

It should be noted that if h_1 and h_2 are generating functions for X_1 and X_2 respectively, then

3.15 $\{h_1, h_2\}$ *is a generating function for* $[X_1, X_2]$.

Let (K, Δ) be a dynamical system. Suppose

3.2 *K has an alternating structure* and

3.21 *the infinitesimal dynamorphism* $\Delta'_{\partial/\partial t, \mathbf{x}}$ *has a generating function* $H_{\mathbf{x}}$ (*the* Hamiltonian).

Then the system is Hamiltonian. If the second condition is strengthened to the following extent (3.22–3.24) the system is called *completely Hamiltonian*.

3.22 *For each one-parameter subgroup T the infinitesimal dynamorphism* $\Delta'_{T, \mathbf{x}}$ *has a generating function* $g_{T, \mathbf{x}}$.

Write $g_{Z, \mathbf{x}}$ for $g_{T, \mathbf{x}}$ where Z is the vector field in R^4 corresponding to the subgroup T.

$$(3.23) \qquad\qquad\qquad g_{Z_1 + Z_2, \mathbf{x}} = g_{Z_1, \mathbf{x}} + g_{Z_2, \mathbf{x}},$$

$$(3.24) \qquad\qquad\qquad g_{\lambda Z, \mathbf{x}} = \lambda g_{Z, \mathbf{x}} \quad \textit{for constant } \lambda.$$

If we have a system (K, Δ) in which K is the space $T^1(Q)$ of tangent vectors in a manifold Q (the "configuration" space) and if that system has a Lagrangian L, then using the Legendre transformation [1], we can construct a Hamiltonian system in which the alternating manifold is the space of covectors or cotangent bundle $T_1(Q)$. This system shall be called $(T^1(Q), \Delta)^*$ or $(T_1(Q), \Delta^*)$. It is interesting that it is not necessarily *completely* Hamiltonian. (Cf. also [8].)

3.3 THEOREM. *Suppose we have a second-order one-particle system (i.e., 2.42 with $n = 1$) for which there is a Lagrangian and such that the associated Hamiltonian system is completely Hamiltonian. Then, in the Einstein-Lorentz case, the Lagrangian is equivalent[6] to*

$$(3.31) \qquad\qquad\qquad 1 - (1 - \dot{x}^2 - \dot{y}^2 - \dot{z}^2)^{1/2}$$

and in the Galilei-Newton case, to

$$(3.32) \qquad\qquad\qquad \tfrac{1}{2}(\dot{x}^2 + \dot{y}^2 + \dot{z}^2) - V(x, y, z).$$

---

[6] Let L and M be two Lagrangians. If there are nonzero constants a, b, c such that $aL - bM$ is of the form $\dot{f} + c$ then L and M are here called equivalent. Of course, then they give the same dynamical trajectories.

When there is a Lagrangian L, the *Legendre mapping* establishes a $1:1$ relation between $T^1(R^3)$ and $T_1(R^3)$ wherein a vector ξ is identified with the covector $\mathscr{L}(\xi) = (\partial L/\partial \dot{x}^i)(\xi)\, dx^i$. Equivalently, one can define p_i by $\partial L/\partial \dot{x}^i$, choose $(x^1, x^2, x^3, p_1, p_2, p_3)$ as coordinates and then make $T^1(R^3)$ into an alternating manifold using the familiar formula [14, 141].

Consider 2.72 with λ erased, as there is only one particle. This has reference to the coordinates $(x^1, x^2, x^3, v^1, v^2, v^3)$ where the v^i are the coordinates $\dot{x}^i$ in $T^1(R^3)$ (see 2.01 above or [2, (4.6)]). We have to rewrite it in terms of $(x^1, x^2, x^3, p_1, p_2, p_3)$. This requires care as $\partial/\partial x^i$ relative to one coordinate system is not the same necessarily as relative to the other. Therefore let us denote the $\partial/\partial x^i$ in 2.72 as D_i. Then the following hold (using the summation convention):

$$(3.33) \qquad D_i = \partial/\partial x^i - H_{x^i p_j} K_{jk}(\partial/\partial p_k),$$

$$(3.34) \qquad \partial/\partial \dot{x}^i = K_{ij}(\partial/\partial p_j),$$

where H is the Hamiltonian $p_i \dot{x}^i - L$, subscripted coordinates indicate partial differentiation while K_{ij} is the inverse of the matrix $H_{p_i p_j}$ which, as usual, is assumed never to vanish on $R^3 \times B$ (B was introduced just below 2.41).

Now $\dot{x}^i D_i + A^i(\partial/\partial \dot{x}^i) = H_{p_a}(\partial/\partial x^a) - H_{x^a}(\partial/\partial p^a)$ as Hamilton's canonical equations say. It results from all this that 2.72 takes the form

$$
(3.35) \qquad
\begin{aligned}
&(Z^a - Z^0 H_{p_a})\frac{\partial}{\partial x^a} \\
&\quad - \{-Z^0 H_{x^c} + K_{ac}[Z^b H_{x^b p_a} - Z^a_0 + H_{p_a}(Z^0_0 + H_{p_b}Z^0_b) - H_{p_b}Z^a_b]\}\frac{\partial}{\partial p_c}.
\end{aligned}
$$

Since (by hypothesis) this has a generating function g, it has the form $-g_{p_a}(\partial/\partial x^a) + g_{x^a}(\partial/\partial p_a)$ from which one can see at once that $-g = Z^a p_a - Z^0 H + f$, where f depends only on x^1, x^2, x^3. We calculate g_{x^c} from this and equate it to the coefficient of $\partial/\partial p_c$ in 3.35.

After considerable simplification, we obtain a relation which says that

$$(3.36) \qquad H(Z^0_c H_{p_c} + Z^0_0) = (Z^b_c p_b + f_c)H_{p_c} + Z^c_0 p_c - Z^b H_{x^b} + \varphi$$

where φ also depends only on x^1, x^2, x^3, and suffixes on f (as on the components Z^a) indicate partial derivatives.

Written in terms of L and the original coordinates (using $H_{p_a} = \dot{x}^a$ and $H_{x^a} = -L_{x^a}$) this says that

$$(3.37) \quad (Z^b D_b + \dot{x}^c Z^b_c(\partial/\partial \dot{x}^b))L + Z^b_0(\partial L/\partial \dot{x}^b) + (Z^0_0 + \dot{x}^c Z^0_c)(L - \dot{x}^b(\partial L/\partial \dot{x}^b)) = \varphi - \dot{x}^c f_c.$$

Now we write this down for an infinitesimal Euclidean vector field, so that $Z^0 = 0$, $Z^a_0 = 0$, $Z^a_b + Z^b_a = 0$, and the Z^a_b are constants. 3.37 holds for such vector fields because these belong to the space-time group in either case. Thus

$$(3.38) \qquad \mathscr{E}L = \varphi - \dot{x}^c f_c$$

where $\mathscr{E}$ is the "lifted" form of Z: $\mathscr{E} = Z^b D_b + \dot{x}^c Z^b_c(\partial/\partial \dot{x}^b)$.

Now denote $\partial L/\partial \dot{x}^c$ by L_c. Presently we will also write $L(0)$ and $L_c(0)$ to indicate that we have set all the $\dot{x}$ equal to 0 ("evaluation on zero section"). From 3.38 we see that $\mathscr{E}L_c + Z_c^b L_b = -f_c$. From this it can be readily deduced that $-\mathscr{E}(\dot{x}^c L_c(0)) = \dot{x}^c f_c$. It is also easy to show $-\mathscr{E}(L(0)) = -\varphi$. It results from this that $L = L(0) + \dot{x}^c L_c(0) + M$ where $\mathscr{E}(M) = 0$. (We also have $M(0) = 0$.) This shows that M itself is a *Euclidean-invariant function* on $T^1(R^3)$. Such a function has to have the form $M = M(s)$, $s = \dot{x}^2 + \dot{y}^2 + \dot{z}^2$.

Letting $L(0) = -V(x, y, z)$,

$$(3.39) \qquad\qquad L = -V(x, y, z) + A_i \dot{x}^i + M(s).$$

The next step is to show that $A_i \, dx^i$ is exact. We insert 3.39 again into 3.38 and equate coefficients of $\dot{x}^c$. This gives us an expression for f_c. We differentiate with respect to x^a and express the fact that this is symmetric in a and c:

$$Z_a^b A_c^b + Z^b A_c^{ab} + Z_c^b A_b^a = Z_c^b A_a^b + Z^b A_a^{cb} + Z_a^b A_b^c.$$

Here superscripts on the A's indicate partial derivatives. We evaluate this at the place where $Z^1 = Z^2 = Z^3$, which can be made to be any desired place in R^3. At that place we have $Z_a^b(A_c^b - A_b^c) + Z_c^b(A_b^a - A_a^b) = 0$. We take $a = 1$, $c = 3$ and select the Z's so that $Z_2^1 = -Z_1^2 = 1$ while all the other $Z_b^a = 0$. This yields $A_2^3 = A_3^2$. Similarly $A_2^1 = A_1^2$, $A_3^1 = A_1^3$. Thus $A_i \, dx^i$ is indeed exact and this term in 3.39 has no effect on the solution of Lagrange's equations.

We now take the case of Galilei-Newton space-time which allows $Z^1 = t$ with $Z^0 = Z^2 = Z^3 = 0$. This we insert in 3.37. The result is $L_1 = \varphi - \dot{x}^c f_c$, or $M'(s)2\dot{x}^1 + A_1 = \varphi - \dot{x}^c f_c$ so that $M'(s) = -f_1$ which makes M' constant. Thus $M(s) = ks$, and the assertion 3.32 is proved.

In the situation involved in 3.31, $Z^1 = t$, $Z^2 = Z^3 = 0$ is a permissible vector field. Insertion into 3.37, together with 3.39 as before, leads to the relation

$$2M'\dot{x}^1 + A_1 + \dot{x}^1(M - V - 2\dot{x}^a \dot{x}^a M') = \varphi - \dot{x}^c f_c.$$

Thus $2M' + M - V - 2sM' = -f_1$, $f_2 = f_3 = 0$. Therefore $2M'(1-s) + M - V$ depends only on x^1. Using the vector field $Z^2 = t$, $Z^0 = x^2$, $Z^1 = Z^3 = 0$ we discover that the same expression depends only on x^2 and is thus constant, which we absorb into V, getting $2M'(1-s) + M = V$. Evidently each side here is constant and

$$M = V(1 - (1-s))^{1/2}.$$

This theorem shows that completely Hamiltonian systems of one particle are necessarily invariant in the Einstein-Lorentz case. We take this as a justification to confine the further discussion to invariant systems, although not simply one-particle systems.

3.4 COROLLARY. *Say* $H = \frac{1}{2}(p_1^2 + \cdots + p_3^2)$ *and* $A = A^1(\partial/\partial x^1) + \cdots + A^3(\partial/\partial x^3)$. *Then* $g_{tA} = A^1 x^1 + A^2 x^2 + A^3 x^3$.

This can be deduced from 3.36. We may abbreviate

$$(3.41) \qquad\qquad A^1 x^1 + \cdots + A^3 x^3 = A \cdot \mathbf{x}.$$

We now turn to a brief study of the action of the Galilean group in the space $T^1(E(3))$ of tangent vectors in the Euclidean group. It is fairly obvious that this is the dynamics (in the strong sense of this paper) of a rigid body. However, we need explicit formulas for this dynamorphism so that we can show how, by means of a suitable Hamiltonian, this dynamics can be transferred to the space $T_1(E(3))$ (the phase space for a rigid body), in such a way as to be a completely Hamiltonian system. Corollary 3.4 will provide a check because the tA there is of course a Galilean infinitesimal transformation of the peculiarly Galilean sort.

Let Gal (4) be the Galilean group and let $E(3)$ be the Euclidean group. Each T in Gal (4) acts in Gal (4) by left multiplication, and this action lifts to $T^1(\text{Gal}(4))$, the space of vectors in Gal (4). An element of Gal (4) can be written as a point $(\tau, \mathbf{a})$ of R^4 together with a *Galilean frame*. By this we mean a quartet of vectors where the first is a unit vector in the sense of 2.11 while the other three are singular relative to 2.11, and orthonormal in the sense of (2.12) (this is the generalization of a Lorentz frame). Such a Galilean frame can be characterized by one vector $(u^1, u^2, u^3, \pm 1)$ in R^4 (the leading member of the Galilean frame), plus a Euclidean frame F in R^3. Thus Gal (4) can be identified with the set of all 5-by-5 matrices of the form

$$(3.5) \qquad \begin{bmatrix} & & & u^1 & a^1 \\ & F & & u^2 & a^2 \\ & & & u^3 & a^3 \\ \hline 0 & 0 & 0 & \pm 1 & \tau \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \quad \text{or} \quad \begin{bmatrix} F & \mathbf{u} & \mathbf{a} \\ 0 & \pm 1 & \tau \\ 0 & 0 & 1 \end{bmatrix}$$

where the box marked F has the vectors of F as its columns. This makes the operations in Gal (4) agree with matrix operations.

Given a matrix G of the form 3.5, denote the orthogonal matrix by F_G. Let $\mathbf{v}$ and $\mathbf{b}$ be vectors in R^3 and form the matrix $(+1, \mathbf{v}, \mathbf{b})$. Here $+1$ is the 3-by-3 identity matrix. Form the new matrix $F_G(+1, \mathbf{v}, \mathbf{b}) G^{-1}$. It can be seen to have the form $(+1, \mathbf{v}', \mathbf{b}')$ and we can therefore define

$$(3.51) \qquad\qquad T_G(\mathbf{v}, \mathbf{b}) = (\mathbf{v}', \mathbf{b}').$$

It is easily seen that

$$(3.52) \qquad\qquad T_{G_1 G_2} = T_{G_1} \circ T_{G_2}.$$

It may be verified that the set of points in R^4 for which $\mathbf{x} = \mathbf{b} + t\mathbf{v}$ holds is transformed by G, then the image is the set of points for which $\mathbf{x} = \mathbf{b}' + t\mathbf{v}'$. Thus T_G gives the dynamorphism of Theorem 2.7 for a single free particle. We want to use

it to construct the dynamorphism for a rigid body. The configuration space for such a body is $E(3)$. An element of $E(3)$ may be written as a matrix

$$(3.53) \qquad \begin{bmatrix} F_0 & \mathbf{b} \\ 0 & 1 \end{bmatrix}$$

where F_0 belongs to the orthogonal group $O(3)$. Usually one requires here $SO(3)$, but we must allow sense-reversing matrices because they represent what a left-handed observer, so to speak, would see.

A vector in $E(3)$ can be described by its base point, say, and another matrix

$$(3.54) \qquad \Gamma = \begin{bmatrix} \Omega & \mathbf{w} \\ 0 & 0 \end{bmatrix}$$

where Ω is a 3-by-3 skew symmetric matrix. The vector this is intended to describe is the one which is tangent *at* M to the curve given parametrically, with parameter ε, as $M \exp \varepsilon\Gamma$. This is clearly Euclidean. If M is as in 3.53 then the moving point in $E(3)$ is given approximately by

$$(3.55) \qquad \begin{bmatrix} F_0 + \varepsilon F_0 \Omega & \mathbf{b} + \varepsilon F_0 \mathbf{w} \\ 0 & 1 \end{bmatrix}.$$

We have to recognize the kinematic significance of describing vectors in $E(3)$ in this way. If 3.53 represents the position ($\mathbf{b}$) of the centroid of a body while F_0 gives the components of an orthogonal triple of vectors embedded in the body, then $\mathbf{w}$ is not the velocity of the centroid but rather, as 3.55 shows,

3.56 *the velocity of the centroid is $F_0\mathbf{w}$.*
Similarly, the *angular velocity* is not given by Ω alone, but is rather the vector associated with the skew symmetric matrix $F_0 \Omega F_0^{-1}$ [4, Appendix].

The advantage of describing a vector in $E(3)$ by a pair M, Γ is that the effect of left-multiplication in $E(3)$, when extended to $T^1(E(3))$ is merely to left-multiply the M and leave the Γ alone.

In order to use 3.51 and 3.52 and also because the velocity and the angular velocity (matrix) are more tangible than the Ω and $\mathbf{w}$ in 3.54, we will denote the vector defined by 3.53, 3.54 (and described by 3.55) by the quadruple

$$(3.57) \qquad [F_0, \, F_0 \Omega F_0^{-1}, \, F_0 \mathbf{w}, \, \mathbf{b}].$$

If a vector is denoted by $(F_0, \omega, \mathbf{v}, \mathbf{b})$ then the scheme of 3.57 can easily be reversed to give the Ω and $\mathbf{w}$.

For a Galilean transformation G we have already defined the orthogonal part F_G. We note another homomorphism into the real numbers. If G is 3.5 let τ_G be the product of nonzero elements of the fourth row. Then

$$(3.58) \qquad \tau_{G_1 G_2} = \tau_{G_1} + \tau_{G_2}.$$

For each element 3.57 of $T^1(E(3))$ and each G as in 3.5 let

$$\mathbf{T}_G[F_0, \omega, \mathbf{v}, \mathbf{b}] = [F_G \exp{(-\tau_G\omega)}F_0, \pm F_G\omega F_G^{-1}, T_G(\mathbf{v}, \mathbf{b})].$$

The sign $\pm$ is to be the same as in the G (3.5). From 3.52 and 3.58 we obtain

(3.59) $$\mathbf{T}_{G_1G_2} = \mathbf{T}_{G_1} \circ \mathbf{T}_{G_2}$$

and thus we have an action of Gal (4) in $T^1(E(3))$.

3.6 THEOREM. *For each G in* Gal (4) *and each Galilei-Newton coordination* **y** *let* $\Delta_\mathbf{y}^{G \circ \mathbf{y}} = \mathbf{T}_G$.

Then Δ defines a Gal (4)-*invariant dynamics with state space* $T^1(E(3))$.

The proof consists in observing that the $U_\mathbf{x}$ defined in 2.5 is nothing but $\mathbf{T}_G$. Hence 2.51 holds by virtue of 3.59. Thus 2.4 and 2.41 hold.

We now proceed to clarify the essentially classical nature of this system, namely that it is the ordinary dynamics of a rigid body subject to no forces. In the first place, we obviously have the right configuration space, viz. $E(3)$.

The element $[F_0, \omega, \mathbf{v}, \mathbf{b}]$ gives the attitude, angular velocity (matrix), velocity and position of the centroid. Anyone familiar with the motion of a rigid body subject to no forces knows how to transform from one observer to another. We have to convince ourselves that $\mathbf{T}_G$ reproduces these facts.

First consider the effect of passage of time. Then G is an orthochronous time translation (i.e., everything in 3.5 is trivial except the τ). Then

$$\mathbf{T}_\tau[F_0, \omega, \mathbf{v}, \mathbf{b}] = [\exp{(-\tau\omega)}F_0, \omega, \mathbf{v}, \mathbf{b} - \tau\mathbf{v}].$$

If the columns of F_0 are the three-unit vector fixed in the body and the relation of matrix to vector is that specified in [4, Appendix] then this triod is indeed[7] spinning with angular velocity ω.

Next we consider the case where G is *Euclidean*, i.e., when the ± 1 is $+1$, the $\mathbf{u}$ is 0, and the τ is 0. Then

(3.61) $$\mathbf{T}[F_0, \omega, \mathbf{v}, \mathbf{b}] = [FF_0, F\omega F^{-1}, F\mathbf{v}, F\mathbf{b} + \mathbf{a}].$$

According to [4, (A.2)] this says that the *new angular velocity equals the old one multiplied* by F^{-1}. Taking the transpose, the new angular velocity *column* vector is F times the old one, as one would expect.

It is useful to note the following.

3.62 *The action* $\mathbf{T}$ *when restricted to G from $E(3)$, is the action induced in $T^1(E(3))$ by left multiplication in $E(3)$.*

This is seen as follows. Take the vector $[F_0, \omega, \mathbf{v}, \mathbf{b}]$. The Γ corresponding to it (3.54) is

$$\Gamma = \begin{bmatrix} F_0^{-1}\omega F_0 & F_0^{-1}\mathbf{v} \\ 0 & 0 \end{bmatrix}$$

---

[7] We have to keep in mind that the effect of T_τ is supposed to give the conditions τ seconds *earlier*. Compare 2.78.

and the 3.53 is just as written there and may be called M. The one-parameter curve is $M \exp \varepsilon \Gamma$. Let us left multiply by

$$E = \begin{bmatrix} F & \mathbf{a} \\ 0 & 1 \end{bmatrix}.$$

Then the new one-parameter curve is $EM \exp \varepsilon \Gamma$, and this is the key to the induced action. We compute the symbol (3.54) corresponding to its tangent and get the right side of 3.61, as asserted.

For later use we study briefly the peculiarly Galilean transformation of passing to a moving frame of reference. In such a case, F is 1, $\mathbf{a}$ is 0, τ is 0 and the ± 1 is $+1$. Thus there is left only the $\mathbf{u}$ and we may write $\mathbf{T_u}$ for $\mathbf{T}_G$. The result is

(3.64) $\mathbf{T_u}(F_0, \omega, \mathbf{v}, \mathbf{b}) = (F_0, \omega, \mathbf{u}+\mathbf{v}, \mathbf{b})$.

If $\mathbf{u} = \varepsilon \mathbf{c}$ then the infinitesimal transformation is $\mathbf{T}'_c = (0, 0, \mathbf{c}, 0)$ for which the differential operator or vector field in $T^1(E(3))$ is

(3.65) $\Delta'_c = c^k(\partial/\partial \dot{x}^k)$

in terms of coordinates $(x^1, x^2, x^3, \ldots)$ for $E(3)$ where x^i is the coordinate which assigns to 3.53 the value b^i. For the same infinitesimal Galilean transformation acting, however, in R^4 the vector field is (cf. 3.4) $t(c^1(\partial/\partial x^1)+c^2(\partial/\partial x^2)+c^3(\partial/\partial x^3))$ where now (x^1, x^2, x^3, t) are the cartesian coordinates in R^4.

Thus 3.64 is not unexpected. The purpose of studying this system is to enable us to prove that its Legendre transform provides us with the first nontrivial example of a *completely* Hamiltonian system.

A purely geometric theorem is now presented which settles the "contact" nature of most of the transformations which we shall meet.

3.7 PROPOSITION. *Let Q be a manifold and let $\mathfrak{X}$ be a vector field in Q. Then $\mathfrak{X}$ induces an infinitesimal contact transformation in $T_1(Q)$, the space of covectors in Q. If $x^1, \ldots, x^n$ is a coordinate system in Q and $X^1, \ldots, X^n$ are the components of $\mathfrak{X}$, then over the domain of that coordinate system this infinitesimal contact transformation has a generating function $-p_1 X^1 - \cdots - p_n X^n$.*

It is well known of course that a *point transformation S* of Q onto Q can be extended to a *contact transformation* of $T_1(Q)$ to $T_1(Q)$, although in several ways. The way intended in 3.7 is to form the map called S^* in [1, 14.16], but we prefer the notation S_1. (Functorial notation would suggest $T_1(S)^{-1}$.)

As usual, $p_1, \ldots, p_n$ are defined in $T_1(Q)$ whenever a coordinate system $(x^1, \ldots, x^n)$ is given for Q by the specification that if α is a covector[8] at q in Q

---

(8) *Old terminology*: covariant vector. *Additional notation*: δ^i_j is the Kronecker delta. $\varepsilon_{jkm} = \pm 1$ or 0 according as jkm is an even, odd, or not a, permutation of 1, 2, 3.

with $(x^1, \ldots, x^n)$-components $b_1, \ldots, b_n$ then $b_i = p_i(\alpha)$. Extending the use of x^i, we set $x^i(\alpha) = x^i(q)$. Thereupon $(x^1, \ldots, x^n, p_1, \ldots, p_n)$ is a coordinate system in $T_1(Q)$ relative to which the alternating structure is (cf. for example [1, *loc. cit.*]).

$$(3.7) \qquad \{f, g\} = \frac{\partial f}{\partial x^k} \frac{\partial g}{\partial p_k} - \frac{\partial f}{\partial p_k} \frac{\partial g}{\partial x^k}.$$

Now let S be any $1:1$ map of Q onto Q. If α is a covector at q we define $S_1(\alpha)$ to be that *covector* at $S(q)$ whose $(x^1, \ldots, x^n)$-components are the $(x^1 \circ S, \ldots, x^n \circ S)$-components of α (just as $S(q)$ is the *point* whose $(x^1, \ldots, x^n)$-coordinates are the $(x^1 \circ S, \ldots, x^n \circ S)$-coordinates of q).

We apply this concept to the one-parameter group of transformations S_ε generated by $\mathfrak{X}$ with the intent of thus defining a vector field $\mathfrak{X}_1$ in $T_1(Q)$. Let $y^i = x^i \circ S_\varepsilon$. Then $y^i = x^i \circ S_\varepsilon$ is (to a suitable degree of approximation) $x^i + \varepsilon X^i$. Let α be a covector at q, where $x^i(q) = a^i$. Then $x^i(S_1(\alpha)) = a^i + \varepsilon X^i$. Let the x^i components of α be $b_1, \ldots, b_n$. Then the y^i components of α are $\bar{b}_\alpha = b_i(\partial x^i/\partial y^\alpha)$. Ignoring ε^2, $\bar{b}_\alpha = b_i(\delta^i_\alpha - \varepsilon X^i_\alpha)$. Hence the change in p_α for α is $-\varepsilon X^i_\alpha p_i$ while the change in x^i is (of course) εX^i. Hence the one-parameter family S_ε determines in $T_1(Q)$ a one-parameter family generated by the vector field

$$(3.71) \qquad \mathfrak{X}_1 = X^i(\partial/\partial x^i) - p_i X^i_\alpha(\partial/\partial p_\alpha).$$

This clearly has the generating function proposed in 3.7 which is thus proved. Incidentally, a coordinate-free form for it is given by $g(\alpha) = \langle \alpha, \mathfrak{X} \rangle$.

To transform the dynamical system of 3.6 onto the phase space $T_1(E(3))$ for a rigid body we need a Lagrangian to set up a Legendre mapping

$$(3.72) \qquad \mathscr{L}: T^1(E(3)) \to T_1(E(3)),$$

via the formula: $\alpha = \mathscr{L}(\xi)$ if $p_i(\alpha) = (\partial L/\partial \dot{x}^i)(\xi)$ and $x^i(\alpha) = x^i(\xi)$. We take

$$(3.73) \qquad L = \tfrac{1}{2}\dot{\mathbf{x}} \cdot \dot{\mathbf{x}} + L'$$

where L' is the kinetic energy of rotation in terms of some coordinates for $E(3)$. When these are selected one can form the Hamiltonian (cf. [4, (4.7)]):

$$(3.74) \qquad H = \tfrac{1}{2}(p_1^2 + p_2^2 + p_3^2) + H'$$

where H' does not depend on the x's and p's. Because of 3.62 and 3.7 we can be sure that the action of $E(3)$ in $T_1(E(3))$ will be given by contact transformations, as will of course the time translations. There remains only to discover a generating function for the peculiarly Galilean transformation 3.65. We can use 3.33, 3.34 here, provided we understand them in a general way for any Lagrangian system. Because of the separation of variables in H, using 3.34 yields

$$(3.75) \qquad \Delta^{*\prime}_c = c^k(\partial/\partial p_k) = c^1(\partial/\partial p_1) + \cdots + c^3(\partial/\partial p_3).$$

It has a generating function (cf. 3.4)

(3.76) $\mathbf{c} \cdot \mathbf{x}.$

We used Δ^* instead of Δ because we are now speaking of the dynamorphism on $T_1(E(3))$. We sum up our findings.

3.8 THEOREM. *The dynamics* Δ^* *defined on* $T_1(E(3))$ *which*

3.81 *for Euclidean mappings, is induced by left multiplication,*

3.82 *for time translations, is generated by* 3.74, *and*

3.83 *for the remaining Galilean transformations, is governed by* 3.76, *is a Galilean invariant dynamics which is completely Hamiltonian.*

The theorem would remain just as significant and valid if "left" in 3.81 were replaced by "right".

In the next section we will present the analogue of this, in which the Galilean group is replaced by the Poincaré group.

4. Poincaré-invariant rigid body dynamics.   We will say that a dynamical system (K, Δ) is a *rigid body dynamical system* if

(4.1) K *is* $T^1(E(3))$

and

4.11 *whenever S is Euclidean, then* $U_x(S)$ *is the mapping in* $T^1(E(3))$ *induced by left-multiplication by S in* $E(3)$. [This does make each S in $E(3)$ a dynamical equivalence in the sense of 2.52. (It does not mean that the temporal evolution of the system commutes with rotations and translations. This latter condition is expressed by $U_x(T) \circ U_x(S) = U_x(S) \circ U_x(T)$ for S in $E(3)$ and T a translation in time. This latter condition would follow, via 2.51, if we assumed that each translation in time were also a dynamical equivalence. This unexpected twist is traceable to the fact that these transformations S, T, etc. work in the *coordinate space* R^4 and not in the "physical" space-time, as they do in [5] and [7]. On the other hand, in the context of [5] and [7], it is not possible to speak of time-translation or euclidean transformation except with reference to a particular observer. However, there is no reason for anxiety when dealing, as we will, with completely invariant systems.)]

U_x was defined in 2.5. Compare also 3.62.

We will show that such systems exist even for the Einstein-Minkowski space time structure. To construct such a system we have to begin with $\mathbf{T}$ as defined in 3.61 and extend the definition of $\mathbf{T}_G$ to all G in the Poincaré group $\mathscr{P}$. Since we are going to make an invariant system we must insure 3.59 also.

The $\mathbf{T}_G$ we have in mind is rather hard to present directly. We will define first the dual dynamical system $(\Delta^*, T_1(E(3)))$ and then pass back to $T^1(E(3))$ using a suitable Hamiltonian and the Legendre transformation.

We will make use of the fact that $E(3)$ can be written as $O(3) \times R^3$, with multiplication $(F, \mathbf{a})(F_0, \mathbf{a}) = (FF_0, \mathbf{a} + F_0\mathbf{a})$. In R^3 we use the coordinates x^1, x^2, x^3. In $O(3)$, we choose any coordinate system, say y^1, y^2, y^3. This defines a coordinate system $(x^1, \ldots, y^3)$ in $E(3)$, and induces, in the natural way, coordinates

$$(x^1, \ldots, y^3, p_1, p_2, p_3, r_1, r_2, r_3)$$

in $T_1(E(3))$. Here the r_i are the "momenta" associated with the y^i.

Let f be a 3-by-3 skew symmetric matrix, and $\mathbf{b}$ an element of R^3. Then $(\exp \varepsilon f, \varepsilon \mathbf{b}) = S_\varepsilon$ is an element of $E(3)$. Let $(F, \mathbf{a})$ be a generic element of $E(3)$. Then $S_\varepsilon(F, \mathbf{a})$ defines a curve and its tangent for $\varepsilon = 0$ is a vector at $(F, \mathbf{a})$ in $E(3)$. Varying $(F, \mathbf{a})$, we obtain a vector field in $E(3) = O(3) \times R^3$ which may be written $X + Z$. If f has the entries f_j^i (column index high) and $\mathbf{b}$ has the components b_i (as is appropriate since $\mathbf{b}$ is a one-column matrix), Z has the form $x^j f_k^i(\partial/\partial x^k)$ $+ b_k(\partial/\partial x^k)$ whereas X has some form $e^1(\partial/\partial y^1) + \cdots + e^3(\partial/\partial y^3)$. It must be observed that

4.2 $X + Z$ *is the infinitesimal left-multiplication corresponding to the infinitesimal transformation* Z *in* R^3.

Accordingly $X + Z$ and also X is *right*-invariant. X is in fact a vector field on $O(3)$. It is possible to choose f so that (with $b = 0$) we get $Z = Z_i \equiv -\varepsilon_{ijk}x^j(\partial/\partial x^k)$. The corresponding X we will now call X_i, and introduce its components (using, as always, the summation convention),

$$(4.21) \qquad\qquad X_i = e_i^j(\partial/\partial y^j).$$

It is easily seen that $[Z_i, Z_j] = \varepsilon_{ijk}Z_k$. Hence

$$(4.22) \qquad\qquad [X_i, X_j] = \varepsilon_{ijk}X_k,$$

whence

$$(4.23) \qquad\qquad e_i^k e_{jk}^m - e_j^k e_{ik}^m = \varepsilon_{ijk}e_k^m$$

where the second suffix indicates differentiation with respect to the y with that index. We define functions ω_i on $T_1(E(3))$ by the formula $\omega_i = e_i^j r_j$. It follows from 4.23 that we have the Poisson bracket relations

$$(4.24) \qquad\qquad \{\omega_i, \omega_j\} = -\varepsilon_{ijk}\omega_k.$$

We use these to define seven functions J_i, K_i and H, as follows: $-J_i = \omega_i + \varepsilon_{ijk}p_j x^k$, H is the positive square root of $1 + p_1 p_1 + \cdots + p_3 p_3 + \theta(\omega_1\omega_1 + \cdots + \omega_3\omega_3)$ where θ is any positive constant (related to the moment of inertia of the body), and $K_i = -x^i H + \varepsilon^{ijk}p_j\omega_k/(H + H_0)$, where H_0 is the square root of

$$1 + \theta(\omega_1\omega_1 + \cdots + \omega_3\omega_3).$$

We let $P_i = -p_i$.

These functions satisfy the relations:

$$
\begin{array}{lll}
\{H, P_i\} = 0 & \{H, J_i\} = 0 & \{H, K_i\} = -P_i \\
(4.25) \quad \{P_i, P_j\} = 0 & \{P_i, J_j\} = \varepsilon_{ijk}P_k & \{P_i, K_j\} = -\delta_{ij}H \\
\{J_i, J_j\} = \varepsilon_{ijk}J_k & \{J_i, K_j\} = \varepsilon_{ijk}K_k & \{K_i, K_j\} = -\varepsilon_{ijk}J_k.
\end{array}
$$

The proof of this is tedious and the reader may prefer to trust us. The table 4.25 is patterned after [10, (1.1)]. It is patterned after the infinitesimal generators of $\mathscr{P}$, which we now list in an order corresponding to H, P_i, J_i, K_i:

$$
(4.26) \qquad \frac{\partial}{\partial t}, \frac{\partial}{\partial x^i}, -\varepsilon_{ijk}x^j \frac{\partial}{\partial x^k}, -x^i \frac{\partial}{\partial t} - t \frac{\partial}{\partial x^i}.
$$

By virtue of 3.15 we see that

(4.27) *the vector fields generated by H, P_i, J_i, K_i satisfy the commutator relations of the generators 4.26 of the Poincaré group $\mathscr{P}$.*

4.3 THEOREM. *There is a Poincaré invariant, completely Hamiltonian, dynamical system (Δ, K) with $K = T_1(E(3))$ and such that for each Lorentz coordinator z, the dynamorphism $\Delta_z^{S \cdot z}$.*

4.31 *for S in $E(3)$ is defined by lifting to $T_1(E(3))$ the left-multiplication by S, and more generally,*

4.32 *is defined infinitesimally by assigning to the generators (4.26) of $\mathscr{P}$ the vector fields generated in $T_1(E(3))$ by H, P_i, J_i, K_i.*

Proof. We consider first 4.31. The vector field called Z_i above appears in the list 4.26. We have already noted (4.2) that the corresponding vector field is $X_i + Z_i$. According to 3.7, it has a generating function $-e_i^j r_j + \varepsilon_{ijk}x^j p_k$ and this is indeed J_i.

Now consider the generator $\partial/\partial x^i$. Taking it as the Z in 4.2, we obtain $X = 0$, so $X + Z$ is $\partial/\partial x^i$ whose generating function is in fact $-p_i$.

We now turn to the K_i. One must show that there is a one-parameter group of transformations in $T_1(E(3))$ whose generator (a differential operator) has this K_i for a generating function. This is indeed true. We will only sketch an argument which can be filled out to give a rigorous proof.

Let f be any function defined on $T_1(E(3))$. Let $V(\varepsilon)$ be the one-parameter group generated by K_i. Let $f(\varepsilon) = f \circ V(\varepsilon)$. Then

$$
(4.33) \qquad (d/d\varepsilon)f(\varepsilon) = \{f(\varepsilon), K_i\}.
$$

If one can solve the *system* given by 4.33 where f runs over a sufficiently large class, and the solutions are valid for all ε, then one has demonstrated the existence of the desired $V(\cdot)$. The main obstacle to a routine application of this idea is that the coordinates y^1, y^2, y^3 are not defined everywhere and hence cannot be inserted in 4.33. We begin, however, by inserting p_1, p_2, p_3, ω_1, ω_2, ω_3. This gives a system in which the x's and y's do not appear, and has an analytic solution. We adjoin the equations for x^1, x^2, x^3. This system is greatly simplified by the integral relation $K^i = \text{const}$.

It has analytic solutions. These solutions show how the **a** in the *base point* $(F, \mathbf{a})$ under a generic element of $T_1(E(3))$ changes under the influence of $V(\varepsilon)$. This brings us to the final question, how does the F change? We must introduce a matrix valued function M on $E(3)$ by defining $M(F, \mathbf{a}) = F$. The differential of this M is dM, also a matrix. Let $\sim$ denote transposition. Then the matrix $M\, dM^{\sim}$ is evidently *right*-invariant on $E(3)$. It is also antisymmetric. Denoting the elements of M by M_{ik} (k here the row index), we obtain from the antisymmetry that $M_{ik}\, dM_{jk} = \varepsilon_{ijk} A_m^k\, dy^m$. From the right-invariance we conclude that the contraction with 4.21, namely $\varepsilon_{ijk} A_m^k e_n^m$, are right-invariant functions, hence constants. Hence $A_m^k e_n^m$ is a constant c_n^k.

Now $\{y^n, K_i\} = e_j^n L_j$ (a sum) where L_j is $\partial K_i / \partial \omega_j$. In other words $dy^n = e_j^n L_j\, d\varepsilon$. Hence $M_{ik}\, dM_{kj} = \varepsilon_{ijk} A_m^k e_n^m L_n\, d\varepsilon = \varepsilon_{ijk} c_n^k L_n\, d\varepsilon$. The important thing is that the right side here is defined over all of $T_1(E(3))$ and in fact depends only on **p**, **x**, and **ω**. Recalling that the transpose here is also the inverse, we see that

$$(4.34) \qquad M^{-1}\, dM = N(\mathbf{x}, \mathbf{p}, \boldsymbol{\omega})\, d\varepsilon,$$

where N is analytic. Into the N here we insert the solutions for $\mathbf{x}, \ldots, \boldsymbol{\omega}$ already obtained, valid for all ε. The resulting equation has a solution $M(\varepsilon)$ for all ε, with $M(0)(F, \mathbf{a}) = F$. This completes the construction of the one-parameter group generated by K_i in $T_1(E(3))$.

It is actually more useful to work with the generating function $\mathbf{c} \cdot K = c_1 K_1 + \cdots + c_3 K_3$ where the c_i are constants. In this case let us denote the one-parameter group by $V_\mathbf{c}(\varepsilon)$. It is easy to verify that $V_\mathbf{c}(\varepsilon) = V_{\varepsilon \mathbf{c}}(1)$, so we may abbreviate it by $V(\mathbf{c})$. This $\mathbf{c} \cdot K$ corresponds to an element of the Lie algebra of the Lorentz group. We denote the exponential of this element by $\exp(\mathbf{c} \cdot K)$.

We have to carry out this type of argument also for H. This turns out to be rather simple and indeed $M_{ik}\, dM_{jk}$ is some constant times $d\varepsilon$. Denote the one-parameter group in this case by $V(\varepsilon)$. εH itself corresponds to an element of the Lie algebra, and we denote its exponential in $\mathscr{P}$ by (ε).

For the element T of time reversal we define U_T to preserve the point of $E(3)$ but reverse the direction of each tangent vector there. Each element P of $\mathscr{P}$ can be written uniquely as a product

$$(4.35) \qquad P = T^m(\varepsilon) \exp(\mathbf{c} \cdot K) E \qquad (m = 0 \text{ or } 1)$$

where E belongs to $E(3)$, and each of the factors depends differentiably on the given P. We define

$$(4.36) \qquad U(P) = (U_T)^m V(\varepsilon) V(\mathbf{c}) V(E)$$

where $V(E)$ is the action induced by left-multiplications as already discussed in connection with 4.31. It must now be observed that

$$(4.37) \qquad U(P_1 \circ P_2) = U(P_1) \circ U(P_2).$$

First we select a coordinate neighborhood Γ in $T_1(E(3))$ and consider the Lie algebra of vector fields generated by the ten functions in 4.22. According to [12, Theorem 88] there is a unique *local* Lie group G acting in Γ for which this is the corresponding set of vector fields. Thus G must be a neighborhood of the identity of $\mathscr{P}$ because it has the same Lie algebra. Now 4.37 certainly holds for the local action assured us by [12, Theorem 88]. One can easily verify that the actions $V_c(\varepsilon)$ and $V(\varepsilon)$ constructed above for the one-parameter groups there discussed are the (unique!) ones assigned by [12, Theorem 88] to the vector fields in question. The analogous problems are trivial for the $V(E)$ in 4.36. Thus 4.37 holds in a neighborhood of the identity in the Poincaré group. Let U restricted to this neighborhood be called U_0. By [12, Theorem 63] it extends to a homomorphism of the simply connected covering group of the group $(\mathscr{P}\uparrow)_+$ which might be two-valued on $(\mathscr{P}\uparrow)_+$. If so, this double-valuedness would persist on restriction to the orthogonal group. But we know (4.31) that there is no double-valuedness at that stage.

To extend the validity of 4.37 we must show that

$$(4.38) \qquad\qquad U(TPT) = U_T U(P) U_T$$

for each P in $(\mathscr{P}\uparrow)_+$. Again, the hardest case is for $P = V(\mathbf{c})$. In the first place, $T \exp(\mathbf{c}\cdot K)T = \exp(-\mathbf{c}\cdot K)$. Thus the question 4.38 in this case is the following: if one reverses the signs of $\mathbf{p}$ and $\boldsymbol{\omega}$ and applies $V(-\mathbf{c})$ and then reverses the signs again, is that the same as applying $V(\mathbf{c})$? The answer is "yes" for the following reasons:

(1) $\{x^i, \mathbf{c}\cdot K\}$ retains its value when the signs of $\mathbf{c}, \mathbf{p}, \boldsymbol{\omega}$ are reversed; (2) same for $\{y^i, \mathbf{c}\cdot K\}$; (3) $\{p_i, \mathbf{c}\cdot K\}$ changes sign when $\mathbf{c}, \mathbf{p}, \boldsymbol{\omega}$ are reversed; (4) so does $\{\omega_i, \mathbf{c}\cdot K\}$.

Now we prove also that

$$(4.39) \qquad\qquad U(SPS) = V(S) U(P) V(S)$$

where S is space reversal, in which $\mathbf{x} \to -\mathbf{x}$, $\mathbf{p} \to -\mathbf{p}$, and $\boldsymbol{\omega} \to \boldsymbol{\omega}$. Considering $P = V(\mathbf{c})$, we note that $S \exp(\mathbf{c}\cdot K)S = \exp(-\mathbf{c}\cdot K)$. Again the question is whether reversing $\mathbf{x}, \mathbf{p}$, applying $V(-\mathbf{c})$ and then reversing $\mathbf{x}, \mathbf{p}$ again, amounts to $V(\mathbf{c})$. Applying $V(S)$ takes us from the component of $E(3)$ to the other and we can use the coordinates y^i in each component, more specifically, $y^i(SE) = y^i(E)$. Thus y^i is preserved. We observe that as the signs of $\mathbf{c}, \mathbf{x}, \mathbf{p}$ are reversed, (1) $\{x^i, \mathbf{c}\cdot K\}$ changes (as does $\mathbf{x}$); (2) $\{p_i, \mathbf{c}\cdot K\}$ changes; (3) $\{y^i, \mathbf{c}\cdot K\}$ stays the same; and (4) $\{\omega_i, \mathbf{c}\cdot K\}$ also stays the same. Hence 4.39 holds for $V(\mathbf{c})$.

The verification of 4.37 for $P = (\varepsilon)$ (time translation) is easy. With that shown, we may consider 4.3 established.

As mentioned at the beginning of this section, we may use the Hamiltonian H to transfer this dynamics to $T^1(E(3))$. It is clear that this can be done and the details are of little interest. One might note, however, that the speed of the centroid cannot approach 1 unless the angular velocity approaches zero.

5. **Hamiltonian systems of a novel type.**   The definition (3.2–3.24) does not imply explicitly or implicitly that K be the cotangent space $T_1(Q)$ of some manifold Q. We proceed at once to exhibit a great class of alternating structures not necessarily of the classical type (wherein, for example, the dimension of K has to be *even*).

5.1 THEOREM. *Let $\mathfrak{G}$ be a Lie algebra, with real coefficients. Let K be the dual of the linear space $\mathfrak{G}$. Then there is exactly one alternating structure on K such that if p, q, r are the linear forms on K defined by $X, Y,$ and $Z = [X, Y]$ in $\mathfrak{G}$, then $\{p, q\} = r$.*

Proof. Let $X_1, \ldots, X_n$ be a basis for $\mathfrak{G}$. Let $p_1, \ldots, p_n$ be the linear forms they define on K. Then we must have $\{p_i, p_j\} = c_{ij}^k p_k$ where these c's are the structure constants of $\mathfrak{G}$ relative to this basis. Hence we must have $A_{ij} = c_{ij}^k p_k$ in this co-ordinate system, so that A, if it exists, must be unique. Conversely, the components A_{ij} define an alternating contravariant tensor (although the indices for natural reasons appear as suffixes). The rule $\{f, g\} = (\partial f/\partial p_i)(\partial g/\partial p_j)c_{ij}^k p_k$ defines a product for which 3.13 also holds. (A little reflection shows that in view of 3.12 it is enough to check it for f, g, h being coordinates, and for these it amounts to precisely the Jacobi-Lie identity for $[\ ,\ ]$ in $\mathfrak{G}$.)

When applied to the group $O(3)$ and K is identified with R^3, this operation comes down to $\{f, g\} = (\nabla f \times \nabla g) \cdot \mathbf{r}$ where $\mathbf{r}$ is the familiar position vector (field) with components x, y, z.

Since dim $K = 3$ in this example, it cannot be a classical Hamiltonian structure. Not unrelated is the existence of a function, viz. r, for which $\{r, f\} = 0$ for all f.

A system of this type can be constructed out of one of the more familiar Hamiltonian type, such as the examples of §3 and §4, when such a system has what we will call *geometric symmetry*.

We will say that a system (K, Δ) has a geometric symmetry group G, where G is some group, if G acts in K in such a way that

$$(5.2) \qquad\qquad \Delta_x^y \circ g = g \circ \Delta_x^y$$

for each pair of coordinators, and each g in G. By G acting in K we mean that for ξ in K and g in G there is defined $g(\xi)$ in K, subject to the rules $g_2(g_1(\xi)) = (g_2 g_1)(\xi)$ and $g(\xi) = \xi$ for all ξ when g is the identity of G.

In the Hamiltonian case we will call it a *Hamiltonian* geometric symmetry if $\{f_1 \circ g, f_2 \circ g\} = \{f_1, f_2\} \circ g$ for any pair of functions on K and each g in G.

Suppose (K, Δ) has a group G of geometric symmetry. For ξ in K let $G(\xi)$ denote the set of all $g(\xi)$, g in G. These sets ("orbits") $G(\xi)$ form the elements of a space called K/G. We can define a dynamics Θ here by setting $\Theta_x^y(G(\xi)) = G(\Delta_x^y(\xi))$. It is easy to verify that $(K/G, \Theta)$ is a dynamical system.

We illustrate this first for the system 3.6. As our group G we choose $O(3)$. For g in $O(3)$ and an element (3.53)

$$(5.3) \qquad\qquad S = \begin{bmatrix} F_0 & \mathbf{b} \\ 0 & 1 \end{bmatrix}$$

in $E(3)$ we define $g(S)$ to be

(5.31)
$$\begin{bmatrix} F_0 g^{-1} & \mathbf{b} \\ 0 & 1 \end{bmatrix}.$$

This means leaving the centroid of the body fixed, but rotating the body. This action of $O(3)$ in $E(3)$ lifts to an action in $T^1(E(\varepsilon))$ because the vector 3.53, 3.54 tangent to the curve 3.55, when that curve is transformed in the manner of 5.31, becomes a vector *at* 5.3 characterized by a matrix in the manner described by 3.54, 3.55 of the form

$$\begin{bmatrix} g\Omega g^{-1} & g\boldsymbol{\omega} \\ 0 & 1 \end{bmatrix}.$$

Hence the velocity of the centroid is *not* changed and neither is the angular velocity. Thus each state of the "quotient" system $(T^1(E(3))/O(3), \Theta)$ has a well defined position and velocity for the centroid, and angular velocity of spin about the centroid, but is equivocal about the "attitude" of body. (Since the group $O(3)$ is compact, this is in fact a random = probability = statistical state for $(T^1(E(3)), \Delta)$.)

If the vector is described as in 3.57, by $[F_0, \omega, \mathbf{v}, \mathbf{b}]$ then

$$g[F_0, \omega, \mathbf{v}, \mathbf{b}] = [F_0 g^{-1}, \omega, \mathbf{v}, \mathbf{b}].$$

Obviously $\mathbf{T}_G g = g \mathbf{T}_G$ as 5.2 requires. The fact that $g[F_0, \omega, \mathbf{v}, \mathbf{b}] = [F_0 g^{-1}, \omega, \mathbf{v}, \mathbf{b}]$ shows that in this case K is of the form $N \times G$ where the action of G is confined to the G. Then K/G will be N, presumably a manifold.

Now we consider the example of 4.3. The space $T_1(E(3))$ has a factor $O(3)$ just as $T^1(E(3))$ does, and there is no difficulty identifying $T_1(E(3))/O(3)$. The new question that arises is this: Has $T_1(E(3))/O(3)$ an alternating structure which makes the dynamics Θ completely Hamiltonian?

5.4 THEOREM. *Suppose N is a $\mathscr{C}^\infty$ manifold and let G be a compact Lie group. Suppose $K = N \times G$ has an alternating structure $\{ \, , \, \}$ which is invariant under the action of G in which $g(n, g_0) = (n, gg_0)$. Suppose the function h on K generates an infinitesimal transformation which commutes with the action of G. Then this infinitesimal transformation can be generated by a function $\bar{h}$ invariant under the action of G.*

Proof. For *any* function h on K let $(L_g h)(n, g_0) = h(n, gg_0)$. Taking the given h, define $\bar{h}$ by $\bar{h} = \int L_g h \, dg$ where we are integrating over the normalized Haar measure of G. Evidently $\bar{h}$ is an invariant function. The question remains, is $\{f, h\} = \{f, \bar{h}\}$? We note that $L_{g^{-1}}\{f, h\} = \{L_{g^{-1}}f, h\}$ because h commutes with the group action. Hence $\{f, h\} = L_g\{L_{g^{-1}}f, h\} = \{f, L_g h\}$ because L_g preserves the alternating structure. Thus

$$\{f, h\} = \int \{f, L_g h\} \, dg \stackrel{?}{=} \left\{ f, \int L_g h \, dg \right\}.$$

Let X be the vector field generated by f. Then the question is

$$\int X(L_g h)\, dg \stackrel{?}{=} X\left[\int L_g h\, dg\right].$$

This is analogous to the question of equality for

$$\int \left[A(x, h)\frac{\partial}{\partial x} + B(x, y)\frac{\partial}{\partial y}\right]h(z+y)\varphi(z)\, dz$$

and

$$\left[A(x, y)\frac{\partial}{\partial x} + B(x, y)\frac{\partial}{\partial y}\right]\int h(z+y)\varphi(z)\, dz$$

(integrations over the whole line, φ and h of compact support) and is a standard technical exercise in Lie groups. The theorem is proved.

We note that on the hypotheses of 5.4

(5.41) *there is an induced alternating structure* $\{\ ,\ \}_N$ *on N such that*

$$(5.42)\qquad\qquad\qquad \{f_1, f_2\}_N = \{f_1^*, f_2^*\}$$

where for f_i defined on N, $f_i^*(n, g) = f_i(n)$.

One takes 5.42 as the definition and verifies its requisite properties.

5.5 THEOREM. *Suppose N is a $\mathscr{C}^\infty$ manifold and G, a Lie group. Let $K = N \times G$. Suppose we have a completely Hamiltonian dynamical system (K, Δ) in which the action of G (as in 5.4) defines a Hamiltonian geometric symmetry. Then the quotient system (N, Θ) is completely Hamiltonian.*

Proof. Let a certain infinitesimal dynamorphism of (K, Δ) be generated by a function h. Then there is an invariant function $\bar{h}$ generating the same dynamorphism. This $\bar{h}$ is of the form f^* for some f defined on N. This f generates the corresponding infinitesimal dynamorphism for (N, Θ).

When Theorem 5.5 is applied to the example of §4, we obtain a system (N, Θ) in which N is $\mathfrak{G}' \times T_1(R^3)$ where $\mathfrak{G}'$ is the dual of the Lie algebra $\mathfrak{G}$ of $O(3)$. Therefore *a state of this system is a pair of things*:

5.51 *a state for a one-particle system in R^3*;

5.52 *an angular momentum.*

It is our contention that this system contains the "classical" (at any rate, nonquantum) analogue of the Dirac electron system. To support this contention we must explore its dynamorphisms.

It can be seen in several ways that the functions on $T_1(E(3))$ which are $O(3)$-invariant are those which depend only on the x^i, p_i, and ω_i introduced earlier. Consequently these can be regarded as coordinates on $T_1(R^3) \times \mathfrak{G}'$. The induced alternating structure here has then the characterization

$$\{x^i, x^j\} = 0, \qquad \{x^i, p_j\} = \delta^i_j, \qquad \{x^i, \omega_j\} = 0,$$
$$\{\omega_i, \omega_j\} = -\varepsilon_{ijk}\omega_k, \qquad \{p_i, p_j\} = 0, \qquad \{p_i, \omega_j\} = 0.$$

The functions involved in 4.25, which generate the infinitesimal dynamorphism of the system (4.3) are expressible in terms of the x's, p's, and ω's and thus those very expressions generate the infinitesimal dynamorphisms in the quotient system.

$$(5.6) \qquad \boldsymbol{\omega} \cdot \boldsymbol{\omega} = \omega_1 \omega_1 + \omega_2 \omega_2 + \omega_3 \omega_3$$

is invariant under all the dynamorphisms.

Proof. It can be shown that $\{\boldsymbol{\omega} \cdot \boldsymbol{\omega}, f\} = 0$ whenever f is any one of the ten generators p_i, J_i, K_i, H of the Poincaré group. Furthermore $\boldsymbol{\omega} \cdot \boldsymbol{\omega}$ is preserved under the action, as we specified it, of time reversal and space reversal.

Select a real number $s \geqq 0$ and consider the subset $N(s)$ of $N = T_1(R^3) \times \mathfrak{G}'$ on which $\boldsymbol{\omega} \cdot \boldsymbol{\omega} = s$. This set $N(s)$ is, by 5.6, invariant under the action of $\mathscr{P}$ and thus forms a subsystem. A state in that subsystem consists of two things,

5.61 *a state for a one-particle system in* R^3;

5.62 *a direction for the spin axis.*

It is evident that the system with configuration space $N(s)$ is the nonquantum analogue of the Dirac electron.

6. **An extension to $\mathscr{P}$ of the regular representation of $E(3)$.** Let T be the left regular representation of $E(3)$. This takes place in the Hilbert space $L^2(E(3), C)$ and, for φ in $L^2(E(3), C)$ and g in $E(3)$, $T_g(\varphi)(h) = \varphi(g^{-1}h)$. Now $E(3)$ can be written as a product, $E(3) = R^3 \times O(3)$ with a multiplication suggested by 5.3. We form a partial Fourier transform $\varphi \to \int e^{-ip \cdot b} \varphi(b, F_0) \, db$ where $p = (p_1, p_2, p_3)$ and $p \cdot h = p_1 b_1 + p_2 b_2 + p_3 b_3$. We thus come to a representation U equivalent to T again in the Hilbert space $L^2(R^3 \times O(3), C)$. We will call it *the representation U of $E(3)$.*

The Lie algebra of $E(3)$ has a basis of right-invariant vector fields corresponding (in the sense explained just prior to 4.2) to the infinitesimal euclidean transformations

$$(6.1) \qquad \partial/\partial x^j, \qquad -\varepsilon_{jkm} x^k (\partial/\partial x^m) \qquad (j = 1, 2, 3).$$

In the representation U of $E(3)$, the basis elements of the Lie algebra corresponding to 6.1 are represented by the skew-adjoint operators

$$(6.11) \qquad ip_j, \qquad X_j - \varepsilon_{jkm} p_k (\partial/\partial p_m).$$

Here X_j are the vector fields on $R^3 \times O(3)$ corresponding to the second triplet in 6.1.

We want to make the analogous remarks concerning the simply-connected covering group of $E(3)$.

The simply-connected covering group of $E(3)$, which we will call $SCE(3)$, can be regarded as a cartesian product $H(2, 0) \times SU(2) \times \{-1, 1\}$ where the last factor is the two-element group consisting of -1 and $+1$. We will write the identity of $SU(2)$ also as 1 for brevity. $H(2, 0)$ is the set of 2-by-2 hermitian matrices of trace 0. The product is defined by $(h, u, \xi)(j, v, \eta) = (h + \xi u j u^*, uv, \xi \eta)$. For $(h, u, \xi) \in SCE(3)$

and j hermitian, define $(h, u, \xi)j = h + \xi u j u^*$. Thus we have $SCE(3)$ *acting* in the linear space $H(2, 0)$. By identifying each point (a_1, a_2, a_3) of R^3 with the hermitian matrix

$$(6.2) \qquad \begin{pmatrix} a_3 & a_1 - a_2^i \\ a_1 + a_2^i & -a_3 \end{pmatrix}$$

we obtain an action of $SCE(3)$ in R^3. This action preserves the euclidean metric and leads to the desired 2:1 homomorphism of $SCE(3)$ onto $E(3)$. Space inversion i_s is provided by the action of $(0, 1, -1)$.

The regular representation T of $SCE(3)$ takes place in $L^2(SCE(3), C)$ analogously to that of $E(3)$. We form a partial Fourier transform $\varphi \to \int e^{-ip \cdot h} \varphi(h, u, \xi)\, dh$ where p is as before but $p \cdot h$ for h as in 6.2 is $p_1 a_1 + \cdots + p_3 a_3$. This leads us to the *representation U of $SCE(3)$*. In this representation, U assigns to the infinitesimal generators of $SCE(3)$, the same operators 6.11 except that now the X_j are the right-invariant vector field on $SU(2) \times \{-1, 1\}$ corresponding to the second triplet in 6.1.

The Poincaré group $\mathscr{P}$ contains $E(3)$ as remarked earlier. We have also to embed $SCE(3)$ in a simply-connected covering group of $\mathscr{P}$.

Let us form a group $SC\mathscr{P}$ whose elements are the elements of $H(2) \times SL(2, C) \times \{-1, 1\} \times \{-1, 1\}$. Here $H(2)$ is the set of hermitian matrices (cf. 6.2).

$$(6.3) \qquad \begin{pmatrix} d+c & a-bi \\ a+bi & d-c \end{pmatrix}$$

(which can be identified with points of R^4 in such a way that the determinant is the Minkowski form). The multiplication is defined as

$$(h, u, \xi, \sigma)(j, v, \eta, \tau) = (h + [\xi u j u^*]^\sigma, uv, \xi\eta, \sigma\tau)$$

where the notation $[\ ]^\sigma$ applied to a hermitian matrix, say 6.3, is that matrix itself if σ is $+1$ and merely reverses the sign of d if σ is -1.

This does define a group. As before, we can let it act on R^4 and this gives a 2:1 homomorphism π into $\mathscr{P}$. The element $(0, 1, 1, -1)$ provides time reversal i_t. The element $(0, 1, -1, -1)$ provides space inversion. The element $N = (0, -1, 1, 1)$ commutes with every element of $SC\mathscr{P}$ and is the only generator of the kernel of π.

The group $SC\mathscr{P}$ contains $SCE(3)$ as a well-defined subgroup. Let us denote the set of (h, u, ξ, σ) for which σ is $+1$ by $SC\mathscr{P} \uparrow$.

6.4 THEOREM. *The representation U of $E(3)$ can be extended to a unitary representation of $\mathscr{P} \uparrow$ in a manner assigning to the infinitesimal Lorentz transformations (cf. 4.26) $-x^j(\partial/\partial x^0) - x^0(\partial/\partial x^j)$ $(j = 1, 2, 3)$ the operators*

$$(6.41) \qquad S\frac{\partial}{\partial p_j} + \frac{p_j}{2S} + \frac{\varepsilon_{jkm} p_k X_m}{S_0 + S}$$

and to time translation by τ, the multiplication operator $\exp i\tau S$ where S_0 is the positive square root of $-\theta(X_1 \circ X_1 + X_2 \circ X_2 + X_3 \circ X_3)$, S is the positive square root

of $S_0^2 + p_1^2 + p_2^2 + p_3^2$, and θ is a nonnegative real number. This representation can be extended to be a representation of $\mathscr{P}$ by defining $U(i_t)$, where i_t is time reversal, to mean "take the complex conjugate and replace p by $-p$".

6.42 THEOREM. *The Theorem 6.4 remains true if all groups are simultaneously replaced by their simply-connected covering groups.*

The next two theorems are about another representation, but they are analogous to 6.4 and 6.42. They are easier to prove and will enable us to prove 6.4 and 6.42.

6.43 THEOREM. *Replace S_0 in 6.4 by any positive scalar E_0 and replace S by E, the positive square root of $E_0^2 + p_1^2 + p_2^2 + p_3^2$. Then the theorem remains true.*

The representation of $\mathscr{P}\uparrow$ will not be the same (unless $\theta = 0$ and $E_0 = 1$) and we will call it W.

6.44 THEOREM. *The analogue of 6.42 holds for W also.*

To prove 6.43 we will introduce an "action" of L in $E(3)$. A matrix of the form (cf. 3.53)

$$(6.5) \qquad \begin{bmatrix} F & \mathbf{a} \\ 0 & 1 \end{bmatrix}$$

is an element of $E(3)$. Here F is in $O(3)$. When $\mathbf{a} = 0$ we will say that 6.5 is in $O(3)$, although it is still a 4-by-4 matrix. Be that as it may, 6.5 is also a homogeneous orthochronous Galilean transformation and this suggests a way of mapping $E(3)$ onto the orthochronous Lorentz group $L\uparrow$. This mapping is to be called "Rect" (for "rectification") and will have the properties

6.6 Rect *maps $E(3)$ onto $L\uparrow$,*

6.61 h Rect $(g) = $ Rect (hg) *if g is in $O(3)$.*

We will use other properties of Rect which follow from the definition now to be given. Let g be the matrix 6.5. Form the unit vector (in the Minkowski sense)

$$(6.62) \qquad (a_1, a_2, a_3, (a_1^2 + a_2^2 + a_3^2 + 1)^{1/2}).$$

Consider the two-dimensional subspace containing $(0, 0, 0, 0)$, $(0, 0, 0, 1)$ and 6.62. There is a unique Lorentz transformation in $L^+\uparrow$ which sends $(0, 0, 0, 1)$ into 6.62 and leaves invariant all vectors Minkowski-orthogonal to the vectors $(0, 0, 0, 1)$ and 6.62. We apply this transformation to the Lorentz frame whose vectors are the columns of

$$\begin{bmatrix} F & 0 \\ 0 & 1 \end{bmatrix}.$$

This gives us a new Lorentz frame

$$(6.63) \qquad \begin{bmatrix} T_a F & \mathbf{a} \\ \mathbf{a}^{\sim} F & b \end{bmatrix} = \text{Rect } (g)$$

where b is $(a_1^2 + \cdots + a_3^2 + 1)^{1/2}$, $\mathbf{a}^\sim$ is the transpose of F and T_a is the 3-by-3 matrix

$$(6.64) \qquad\qquad\qquad 1 + (1+b)^{-1} \mathbf{a} \mathbf{a}^\sim .$$

Rect (g) is defined by 6.63.

In the interest of saving space we merely assure the reader that the construction geometrically described does take a given element g of $E(3)$ into Rect (g). The analytic form of 6.64 and 6.63 is going to be of some consequence.

Now 6.61 holds because of the rotation-invariant definition of the process. Finally, let h belong to $L \uparrow$ and have a certain well-known normal form. Then a g can be formed such that Rect $(g) = h$. From here, 6.61 leads to 6.6.

As a result of 6.6, we can define an *action* (compare the discussion below 5.2) of $L \uparrow$ in $E(3)$. For h in $L \uparrow$ and g in $E(3)$, let $R_h(g) = \text{Rect}^{-1}(h \text{ Rect }(g))$. Evidently $R_{h_1} \circ R_{h_2} = R_{h_1 h_2}$. From 6.61 we also see that $R_h(g) = hg$ if h is in $O(3)$. These results may be formulated as follows.

6.7 PROPOSITION. *The action R of L^+ in $E(3)$ is an extension to $L \uparrow$ of the restriction to $O(3)$ of the left-regular representation of $E(3)$.*

We come now to the main result on "Rect".

6.8 LEMMA. *The action R of $L \uparrow$ in $E(3)$ assigns to the infinitesimal Lorentz transformation* $-x^4(\partial/\partial x^j) - x^j(\partial/\partial x^4)$, *the infinitesimal transformation*

$$(6.81) \qquad\qquad (1 + \mathbf{a} \cdot \mathbf{a})^{1/2} \frac{\partial}{\partial a_j} + \frac{\varepsilon_{jkm} a_k X_m}{1 + (1 + \mathbf{a} \cdot \mathbf{a})^{1/2}}$$

in $E(3)$.

For the X_m, see 4.21. To X_m corresponds the skew-symmetric matrix $-s_m$ with entry $-\varepsilon_{mjk}$ in the jth column and kth row.

We sketch the proof for $m = 1$. The infinitesimal Lorentz transformation then involved corresponds to the matrix

$$\begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}$$

which we will call Z.

The thing to be proved is that, if g is 6.5 then

$$(6.82) \qquad\qquad\qquad R_{\exp \varepsilon Z}(g)$$

is the same as what the infinitesimal transformation 6.81, performed for ε units of time (so to speak) on g produces. This latter quantity can be estimated as follows.

The **a** in 6.5 changes in accordance with the first term of 6.81 and thus is transformed, to a suitable degree of approximation, to $\mathbf{a}+\varepsilon b\mathbf{u}$ where

$$u = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix},$$

because $m=1$. The F_0 changes in accordance with the second term of 6.81, which means it is left-multiplied by $\exp \varepsilon Y$ where the matrix

$$Y = (a_3 s_2 - a_2 s_3)/(1+b).$$

The thing to be proved is therefore that

$$\text{Rect}^{-1}\left(\exp \varepsilon Z \,\text{Rect}\,(g)\right) = \begin{bmatrix} (\exp \varepsilon Y)F_0 & \mathbf{a}+\varepsilon b\mathbf{u} \\ 0 & 1 \end{bmatrix},$$

or that

$$\text{Rect} \begin{bmatrix} (\exp \varepsilon Y)F_0 & \mathbf{a}+\varepsilon b\mathbf{u} \\ 0 & 1 \end{bmatrix} = \exp \varepsilon Z \,\text{Rect}\,(g).$$

Now we must appeal to 6.63 and 6.64, and compute these matrices. It turns out that we must verify four things, two of which are obvious while the other two are

(6.83) $$T_{\mathbf{a}+\varepsilon b\mathbf{u}}(1+\varepsilon Y) = T_{\mathbf{a}}+\varepsilon\mathbf{u}\mathbf{a}^{\sim}$$

and

$$(\mathbf{a}+\varepsilon b\mathbf{u})^{\sim}(1+\varepsilon Y) = \mathbf{a}^{\sim}+\varepsilon\mathbf{u}^{\sim}T_{\mathbf{a}}.$$

It is better to put 6.83 into the form $T_{\mathbf{a}}+\varepsilon b\mathbf{u}=(T_{\mathbf{a}}+\varepsilon\mathbf{u}\mathbf{a}^{\sim})(1-\varepsilon Y)$.

The last two equations can be verified. It takes about five pages of computation. This completes the proof of 6.8.

Let E_0 be any positive number and introduce the new variables p_1, p_2, p_3 where $p_j = E_0 a_j$. This gives us a result more versatile than 6.8.

6.84 LEMMA. *There is an action of $L\uparrow$ in $E(3)$ which is an extension to $L\uparrow$ of the restriction to $O(3)$ of the left-regular representation of $E(3)$ and which assigns to $-x^4(\partial/\partial x^j)-x^j(\partial/\partial x^4)$ the vector field*

(6.85) $$E(\partial/\partial p_j)+\varepsilon_{jkm}p_k X_m/(E_0+E)$$

in $E(3)$. Here $E=(E_0+\mathbf{p}\cdot\mathbf{p})^{1/2}$.

6.86 PROPOSITION. *This Lemma 6.84 also holds if all three groups are replaced by their simply-connected covering groups.*

Denote by π the 2:1 homomorphism of $SCE(3)$ on $E(3)$. For an element h near the identity e of $\mathscr{P}\uparrow$ and each g in $SCE(3)$, define $R'_h(g)$ by three conditions: (1) $\pi(R'_h(g))=R_h(\pi(g))$, (2) $R'_h(g)$ depends continuously on h, (3) $R'_e(g)=g$. This defines a *local* homomorphism of $\mathscr{P}\uparrow$ into the group of homeomorphisms of

$SCE(3)$, which lifts to a homomorphism R'' of the component of the identity of $SC\mathscr{P}\uparrow$, into that group of homeomorphisms. For h in the other component, write $h=ji$ where i is space-inversion, and define $R''_h(g)=R''_j(ig)$.

This proves 6.86.

As pointed out in [11, 100–101], when a group acts in a manifold $\mathscr{E}$ it gives rise to a unitary representation of that group in the intrinsic Hilbert space $\mathscr{H}$ of that manifold. If we choose a measure m for $\mathscr{E}$, then $\mathscr{H}(\mathscr{E})$ is unitary-equivalent to $L^2(\mathscr{E}, m)$ and this representation takes the concrete form of substitution (of the inverse) and multiplying by a power of the Radon-Nikodym derivative.

Suppose a one-parameter group $V(t)$, acting in $\mathscr{E}$ has the infinitesimal transformation

$$(6.87) \qquad X^1(\partial/\partial x^1)+\cdots+X^n(\partial/\partial x^n).$$

Then the unitary representation $W(t)$ to which $V(t)$ gives rise has a skew-adjoint generator in $L^2(\mathscr{E}, m)$. This operator is obtained by adding to 6.87 some *function* calculated to produce a formally skew-adjoint operator [11, 101].

We apply this idea to $\mathscr{E}=E(3)$, selecting for m the Haar measure for $E(3)$. This is the product of ordinary Lebesgue measure in R^3 and Haar measure in $O(3)$. The vector fields X_m are already skew-adjoint, since they are the generators of the rotations. The X_m commutes with the p's and hence the second term in 6.85 is skew-adjoint. However, the first is not and has to be replaced by

$$\frac{1}{2}\left[E\frac{\partial}{\partial p_j}-\left(E\frac{\partial}{\partial p_j}\right)^*\right] = \frac{1}{2}\left[E\frac{\partial}{\partial p_j}-\left(\frac{\partial}{\partial p_j}\right)^*E\right]$$

$$= \frac{1}{2}\left[E\frac{\partial}{\partial p_j}+\frac{\partial}{\partial p_j}E\right] = E\frac{\partial}{\partial p_j}+\frac{p_j}{2E}.$$

Thus we have arrived at the following.

6.9 LEMMA. *There is a unitary representation W of $L\uparrow$ in $L^2(E(3))$ which is an extension of the restriction to $O(3)$ of the left-regular unitary representation of $E(3)$ in $L^2(E(3))$ and which assigns to $-x^4(\partial/\partial x^j)-x^i(\partial/\partial x^4)$ the skew-adjoint operator*

$$(6.91) \qquad E\frac{\partial}{\partial p_j}+\frac{p_j}{2E}+\frac{\varepsilon_{jkm}p_kX_m}{E_0+E}.$$

The next step is to extend this representation to all of $\mathscr{P}\uparrow$. Denote by α the translation defined by 6.3. (The d is then the amount of time translation.) Denote $p_1a+p_2b+p_3c+Ed$ by αp. We will define

$$(6.92) \qquad W(\alpha) = e^{i\alpha p}.$$

This is a multiplication operator. It will be appreciated that this combines with the W in 6.9 to give a representation of $\mathscr{P}\uparrow$ if

$$(6.93) \qquad \exp(i(h\alpha h^{-1})p) = W(h)e^{i\alpha p}W(h)^{-1},$$

for each h in $L\uparrow$. In checking 6.93 we can forget about the Radon-Nikodym factors in the $W(h)$ and $W(h)^{-1}$ because they cancel. Hence $W(h)$ becomes a substitution operator, as a matter of fact, substitution of $(R_h)^{-1}$. Thus the validity of 6.93 follows from

$$(h\alpha h^{-1})p = \alpha p \circ (R_h)^{-1} = \alpha p \circ R_h{}^{-1},$$

and this can be seen as follows. In general αp evaluated for any g, say g as in 6.5, is the dot product (in the Minkowski sense) of the fourth vector in the Lorentz frame Rect (g) with the vector $(a, b, c, -d)$. In the process of rectification this gets operated on by h^{-1}. Thus $\alpha p(R_h{}^{-1}(g))$ is the Minkowski dot product of $(a, b, c, -d)$ and the image under h^{-1} of this fourth vector; while $(h\alpha h^{-1})p(g)$ is the fourth vector dotted with the vector $(a', b', c', -d')$ corresponding to the translation $h\alpha h^{-1}$. Computation shows they are equal. The result is the following.

6.94 THEOREM. *One may replace $L\uparrow$ in 6.9 by $\mathscr{P}\uparrow$.*

One can extend the representation W in 6.9 to all of $\mathscr{P}$ by defining, for time-reversal i_t,

$$(6.95) \qquad\qquad (W(i_t)\varphi)(F, p) = \overline{\varphi(F, -p)}.$$

Here we are regarding $E(3)$ as the product $O(3) \times R^3$.

Thus 6.43 is now proved.

Just as 6.94 is based on 6.8, we can build instead on 6.86, obtaining the following.

6.96 COROLLARY. *One may replace, either in 6.9 or in 6.94, all groups simultaneously by their simply-connected covering groups.*

The added fact about $W(N)$ follows from 6.87.

We can improve 6.96 by extending the representation to all of $SC\mathscr{P}$ by defining, for time-reversal,

$$(6.97) \qquad\qquad (W(i_t)\varphi)(p, \mu, \xi) = \overline{\varphi(-p, \mu, \xi)}.$$

This is the substance of 6.44 which is therefore also proved. From 6.43 we now prove 6.4. The Hilbert space in question, we may call $\mathscr{H}$. Let u be an irreducible matrix representation of the compact group $O(3)$. Let $\mathscr{H}(u)$ denote the closed linear subspace generated by those elements φ for which there exist functions f^k on R^3 such that

$$(6.98) \qquad\qquad \varphi(p, F) = f^k_j(p)u^j_k(F) \qquad (F \text{ in } O(3)).$$

Now $\mathscr{H}$ is the orthogonal sum of these $\mathscr{H}(u)$, if u runs over a complete collection of representatives, one from each equivalence class. If a representation having the desired infinitesimal form (e.g., 6.41) exists in each $\mathscr{H}(u)$, then 6.4 is proved.

Now in $\mathscr{H}(u)$ the operator $X_1 \circ X_1 + X_2 \circ X_2 + X_3 \circ X_3$ is indeed a scalar, a negative scalar $-\eta$ at that. We let $E_0 = (1 + \theta\eta)^{1/2}$ and consider the representation

W of 6.43. This representation leaves $\mathscr{H}(u)$ invariant, and its infinitesimal genera-
tors, in $\mathscr{H}(u)$, coincide with those given in 6.4 when restricted to $\mathscr{H}(u)$. We can
therefore adopt W on $\mathscr{H}(u)$ to define U on $\mathscr{H}(u)$. We have only to check that they
agree on $E(3)$, where U is already given. 6.94 shows that they agree on $O(3)$. The
definition for translations (6.92 with $d=0$) shows that W agrees with U there,
and thus agrees with U on all of $E(3)$. Thus 6.4 is proved. To prove 6.42, we use
6.44 (instead of 6.43) and 6.96 (instead of 6.94). The representation u in this case
is again a representation of the compact subgroup, in this case $SU(2) \times \{-1, 1\}$.

7. Poincaré invariant quantum dynamics of a rigid body.

We adhere to the
G. Birkhoff-von Neumann concept of a quantum system as simplified by Gleason
and extended by Mackey [11, 135–136]. We have, however, to extend the definition
before it defines a dynamical system in the sense of this paper. Accordingly, we
shall call a dynamical system (Δ, k) a *quantum* dynamical system if K is the union
of disjoint sets $K_1, K_2, \ldots$ where each K_j is the projective space $P(\mathscr{H}_j)$ of (complex)
one-dimensional subspaces of some infinite-dimensional Hilbert space $\mathscr{H}_j$; and $\Delta_{\mathbf{x}}^{\mathbf{y}}$
maps each K_j onto itself in such a manner as to be a physical correspondence
[16, 203], which is to say if ξ and ξ' are elements of K_j and $[\xi : \xi']$ is defined as
$|(\varphi; \varphi')|^2$ for unit vectors φ in ξ and φ' in ξ', then $[\Delta_{\mathbf{x}}^{\mathbf{y}}\xi : \Delta_{\mathbf{x}}^{\mathbf{y}}\xi'] = [\xi : \xi']$.

In view of a theorem of Wigner [16, 204] a mathematically equivalent definition
is as follows. A dynamical system (Δ, H) is a quantum dynamical system if H is the
union of disjoint infinite-dimensional Hilbert spaces $\mathscr{H}_1, \mathscr{H}_2, \ldots$ and each $\Delta_{\mathbf{x}}^{\mathbf{y}}$, when
restricted to any $\mathscr{H}_j$ is

7.1 *a unitary or an antiunitary map of* $\mathscr{H}_j$ *onto* $\mathscr{H}_j$.

However, we must understand that with this revised definition we have to relax
2.4 and 2.41 as follows.

7.11 *On each* $\mathscr{H}_j$, $\Delta_{\mathbf{y}}^{\mathbf{z}} \circ \Delta_{\mathbf{x}}^{\mathbf{y}} = \lambda(\mathbf{x}, \mathbf{y}, \mathbf{z}, j)\Delta_{\mathbf{x}}^{\mathbf{z}}$, *where* $\sigma(\mathbf{x}, \mathbf{y}, \mathbf{z}, j)$ *is a complex
number.*

7.12 *In each* $\mathscr{H}_j$, $\Delta_{\mathbf{x}}^{\mathbf{x}} = \lambda(\mathbf{x}, j)$ *where* $\lambda(\mathbf{x}, j)$ *is a complex scalar.*

Finally, let (Δ, H) and (Θ, K) be two quantum dynamical systems with a 1:1
mapping of H onto K which is a unitary or antiunitary U_j when restricted to each
$\mathscr{H}_j$ and such that

$$(7.2) \qquad\qquad \Theta_{\mathbf{x}}^{\mathbf{y}} \circ U_j = \mu(\mathbf{x}, \mathbf{y}, j)U_j \circ \Delta^{\mathbf{y}},$$

then we must call these two systems *isomorphic*.

All these complex numbers will have modulus 1.

Finally (cf. 2.5), S will be called a quantum dynamical equivalence of a given
system if

7.21 *on each* $\mathscr{H}_j$, $\Delta_{\mathbf{x}}^{S \circ \mathbf{x}} = \nu(\mathbf{x}, \mathbf{y}, j)\Delta_{\mathbf{y}}^{S \circ \mathbf{y}}$ *where* $\nu(\mathbf{x}, \mathbf{y}, j)$ *is a complex number.*

Then instead of 2.53, we obtain [11, 126],

7.23 *if* S *and* T *are quantum dynamical equivalences then, on* $\mathscr{H}_j$ *(see 2.5),*

$$U_{\mathbf{x}}(S) \circ U_{\mathbf{x}}(T) = \nu(x, T \circ x, j)\lambda(x, T \circ x, S \circ T \circ x, j)U_{\mathbf{x}}(S \circ T).$$

It is natural in quantizing a classical system with configuration space Q to use the intrinsic Hilbert space $\mathcal{H}(Q)$. Thus we will call a quantum dynamical system (Δ, H) a quantum dynamical system *of a rigid body* if $H = L^2(E(3), m)$ where m is the Haar measure of the Euclidean group $E(3)$, and (compare 4.1 and 4.11) whenever S is Euclidean, and $\mathbf{x}$ is any coordinator, then there is a complex number $\sigma(\mathbf{x}, S)$ such that, for u in $L^2(E(3))$ and g in $E(3)$,

$$(7.3) \qquad\qquad (U_{\mathbf{x}}(S)\varphi)(g) = \sigma(\mathbf{x}, S)\varphi(S^{-1}g).$$

The first Poincaré invariant example is given by the representation of 6.4. One defines $\Delta_{\mathbf{x}}^{S \circ \mathbf{x}}$ as $U(S)$ for every S in $\mathcal{P}$.

We analyze this representation, for another purpose, in §7. It appears that in this system, the total spin has only the values $0, 1, 2, \ldots$. The same is true if we use 6.43, but it does not "correspond" in a sense to be described in §8, to the "classical" system described in §4.

To obtain another example we turn to 6.42. Here the Hilbert space is $L^2(SCE(3))$, but the representation is one of $SC\mathcal{P}$ and not $\mathcal{P}$. It is not hard to see that if we have a representation U of $SC\mathcal{P}$ by unitary and antiunitary operators in a Hilbert space $\mathcal{H}$ of such a sort that $U(g)$ and $U(g')$ produce the same effect in $P(\mathcal{H})$ whenever $\pi(g) = \pi(g')$ in $\mathcal{P}$, then $U(N)$ must be either $+1$ on all of $\mathcal{H}$ or -1 on all of $\mathcal{H}$. Since this is not true of $U(N)$ in $L^2(SCE(3))$ we consider the subspaces $\mathcal{H}(+)$ and $\mathcal{H}(-)$ on which the unitary $U(N)$ is $+1$ and -1 respectively.

We can map $L^2(E(3))$ into $L^2(SCE(3))$ by sending φ into $\varphi \circ \pi$. The image of $L^2(E(3))$ here is exactly $\mathcal{H}(+)$. Therefore $\mathcal{H}(+)$ yields a system equivalent to the previous example. The orthogonal complement $\mathcal{H}(-)$ consists of those φ in $L^2(SCE(3))$ for which $\varphi(Ng) = -\varphi(g)$. There are several ways of mapping $L^2(E(3))$ unitarily on $\mathcal{H}(-)$. By using the upper and lower hemispheres of $SU(2)$ for example, one can divide $SCE(3)$ into two measurable sets A and B such that $B = NA = AN$. For φ in $L^2(E(3))$, there is exactly one ψ in $\mathcal{H}(-)$ such that ψ and $\varphi \circ \pi$ have the same restriction to A (and hence are negatives on B). We can thus carry our representation U of $SC\mathcal{P}$ back into $L^2(E(3))$ as 7.3 requires. We denote this representation by V.

7.4 THEOREM. *For S in $\mathcal{P}$ take g in A such that $\pi(g) = S$ and define*

$$\Delta_{\mathbf{x}}^{S \circ \mathbf{x}} = V(g).$$

This defines a quantum dynamical system of a rigid body. With varying choices of A and B one obtains isomorphic systems.

The proof is simple enough to be omitted. The various multipliers in 7.11–7.3 are of course all ± 1.

7.5 THEOREM. *In the case of each of the two systems defined above there is in $L^2(E(3))$ a projective representation [11, 126] of $O(3)$ by isomorphisms of the system. The kinematic significance is the following: if ξ and η are related by one of these*

isomorphisms then in these two states the body has the same location and momentum of the centroid and the same angular momentum, but the body itself will have undergone an orthogonal transformation.

Proof. It will suffice to deal only with the more complicated case arising from $\mathscr{H}(-)$. First we employ an isomorphism to replace $L^2(E(3))$ by $\mathscr{H}(-)$.

Now we define a new action of $SU(2) \times \{-1, 1\} = G$. For (v, η) in G and (h, u, ξ) in $SCE(3)$ define $(h, u, \xi)(v, \eta) = (h, uv, \xi\eta)$. Computation reveals that this action commutes with left multiplication in $SCE(3)$. Hence it gives rise to a representation of G in $L^2(SCE(3))$, which commutes with the regular representation and in particular with N. Thus it preserves $\mathscr{H}(-)$. In this representation of G the operator corresponding to N is the same as for the regular representation: -1 on $\mathscr{H}(-)$. In an obvious way this generates a projective representation of $O(3)$. The kinematic significance cannot be demonstrated until the observables involved have been defined.

When this is done in the conventional way, it turns out to be true. For example, observables depending only on the *momentum* of the centroid and the *position* of the body relative to the centroid are operators which merely multiply φ in $L^2(E(3))$ by some function defined on $E(3)$. The position operator is $-i\nabla_p$. All of these are related to the representation in the way required by our assertion.

We obtain only isomorphisms in this projective representation of $O(3)$ because the action of G commutes not only with the left-regular action of $E(3)$ on itself but also with the action of $L\uparrow$ on $E(3)$; and the representation of G commutes with the operations defined to extend W to $\mathscr{P}$. Thus 7.5 is proved.

Let $\mathscr{U}(\pm)$ be the class of irreducible representations u of $SU(2) \times \{-1, 1\}$ for which $u(N)$ is the scalar operator ± 1. Let $\mathscr{U}(\pm)$ be a set of matrix forms, one for each equivalence class. For each u in $\mathscr{U}(\pm)$, let $\mathscr{H}(u)$ be the subspace of those functions in $L^2(SCE(3))$ of the form $\psi_j^k(p)u_k^m$. Then $\mathscr{H}(u)$ lies in $\mathscr{H}(\pm)$ and (cf. [4, 2.46])

7.6 $\mathscr{H}(\pm)$ *is the direct sum of the* $\mathscr{H}(u)$ *for* $\mathscr{U}$ *in* $u(\pm)$.

8. **The quotient concept for quantum systems.** We turn to a quantum analogue of 5.5. The question is, if a quantum dynamical system (Δ, H) has a group G of geometric symmetries (5.2) is the quotient system also a quantum dynamical system? For simplicity we take H to consist of only one Hilbert space $\mathscr{H}$ and we suppose([9]) that (cf. 7.2, 7.5)

([9]) We do not include in our general concept of a system any specification of which functions on K are to be observables. A *system with observables* is therefore a more specific thing than a system in our sense. Mackey systems [11] are systems with observables, evidently. As Mackey does, one could specify a *logic* as the collection of certain subsets of K. One could require the action of G to define an automorphism of the logic. This would imply 8.1. The idea of the quotient of a system with a logic is in [4]. However, the full set of dynamorphisms is not considered in [4].

8.1 *G is represented by isomorphisms of* (Δ, H).

We will consider *for simplicity* a situation slightly more general than that presented by 7.5. Let P be a space with a measure and let G be a compact group. Let u be a continuous matrix representation of G which is equivalent to its complex conjugate

$$(8.11) \qquad\qquad \bar{u} = \sigma u \sigma^{-1}.$$

We have a natural measure in $P \times G$ and can form $L^2(P \times G, C)$ and single out therein the closed linear subspace $\mathcal{H}(u)$ generated by the functions of the form $\psi_k^j(p)u_j^k(g)$ where the summation goes (in each case) from 1 to d, the degree of u.

We will suppose that we have a system $(\mathcal{H}, \Delta)$ in which $\mathcal{H}$ is $\mathcal{H}(u)$. We suppose that the way in which G acts is by *right transformation*, i.e., for φ in $\mathcal{H}$, and a in G, $(R_a\varphi)(p, g) = \varphi(p, ga)$ for each (p, g) in $P \times G$.

We can form $L^2(P, C^d)$ thinking of an element ψ of this space as a *column* of d complex-valued functions defined on P. Let * as usual denote taking the complex-conjugate transpose of the matrix to which it is applied. Then ψ^* is a row (of functions) and ψ^*u is a row of functions (defined on $P \times G$) because u is a d-by-d matrix. It would obviously be neater here to take ψ as a row-matrix and leave off the * but we have to use columns if we want to compare our results with others in the physics literature.

Denote by $\Gamma(\psi)$ the subspace of $L^2(P \times G)$ generated by the d functions in the row-matrix ψ^*u.

8.2 PROPOSITION. $\Gamma(\psi)$ *is in* $\mathcal{H}(u)$. *If* $\psi \neq 0$, *then* $\Gamma(\psi)$ *is d-dimensional and G-invariant and is minimal with respect to being G-invariant. Every minimal G-invariant subspace of* $\mathcal{H}(u)$ *is of the form* $\Gamma(\psi)$ *for some* ψ *and this* ψ *is unique up to a scalar multiple.*

A proof of this is easy. The fact that every minimal G-invariant subspace is of the form $\Gamma(\psi)$ can be made up along the lines of [4, 356], but there are a few misprints on that page 356. Every reference to 2.43 should be to 2.42 and every reference to 2.42 should be to 2.43. Also (2.41) is not the (usual) definition of representation used here.

Now suppose D is one of the dynamorphisms of the system. Presumably, because R_a is an isomorphism of the system, we have D is unitary or antiunitary and

$$(8.21) \qquad\qquad D \circ R_a = \mu(a)R_a \circ D$$

where $\mu(a)$ is a scalar. Then we can show that

8.22 *there is a unitary or anitunitary (resp.) operator* $\hat{D}$ *in* $L^2(P, C^d)$ *such that* $\Gamma(\hat{D}\psi) = D(\Gamma(\psi))$.

In greater detail (with $\nu = 1/\mu$),

8.23 $\hat{D}\psi = [\nu D(\psi^*u)_e]^*$ *if D is linear;*

8.24 $\hat{D}\psi = [\nu D(\psi^*u)_e\sigma]^*$ *if D is antilinear.*

Here ψ^*u is a row of elements of $L^2(P \times G)$ and $D(\psi^*u)$ is meant to be the row whose elements are the images under D of the elements of ψ^*u. The subscript e is

intended to indicate that the result is evaluated at the identity element of G. We verify the sufficiency of 8.24. Insert ψ^*u into 8.21 and obtain

$$D(R_a\psi^*u) = \mu(a)R_a[D(\psi^*u)].$$

Now $R_au=uu(a)$, so $\mu(a)R_a[D(\psi^*u)]=D(\psi^*u)\bar{u}(a)=D(\psi^*u)\sigma u(a)\sigma^{-1}$. Evaluation at e (this does not mean set $a=e$!) gives $\mu(a)R_a[D(\psi^*u)_e]=D(\psi^*u)_e\sigma u(a)\sigma^{-1}$ from which one obtains $\mu D(\psi^*u)=D(\psi^*u)_e\sigma u\sigma^{-1}$, or $\nu D(\psi^*u)_e\sigma u=D(\psi^*u)\sigma$. Now let λ be any column matrix of d scalars. As λ ranges over all possibilities, $D(\psi^*u)\sigma\lambda$ takes on all possible values accessible to $D(\psi^*u)\lambda'$ where λ' is any column matrix of scalars, as σ is nonsingular. This latter set of values is the same as the set of values of $D(\psi^*u\lambda'')$ and hence the same as the subspace $D(\Gamma(\psi))$. This is thus equal to the set of values of $\nu D(\psi^*u)_e\sigma u\lambda$, i.e., $\Gamma((\nu D(\psi^*u)_e\sigma)^*)$. This, in view of 8.24, establishes 8.22 for the antilinear case. The linear case is easier: just omit the σ.

We have thus proved the following

8.3 THEOREM. *The quotient system for* $(\mathscr{H}(u), \Delta)$ *is the quantum dynamical system* $(L^2(P, \boldsymbol{C}^d), \hat{\Delta})$ *where the dynamorphisms are obtained from those of* $(\mathscr{H}(u), \Delta)$ *by the formulas of* 8.22.

Now let $\mathscr{U}'$ be a class of mutually equivalent continuous irreducible representations of G, satisfying 8.11 (with a σ for each u, of course). Suppose we have a quantum dynamics Δ in the Hilbert space $\mathscr{H}(\mathscr{U}')$ generated in $L^2(P \times G)$ by the subspaces $\mathscr{H}(u)$ where each $\mathscr{H}(u)$ is invariant under the dynamorphisms. Also assume t at G acts by right transformation in $\mathscr{H}(u)$ in a manner defining isomorphisms of $(\mathscr{H}(\mathscr{U}'), \Delta)$, meaning 8.21 holds for each dynamorphism. We obtain the following result.

8.4 THEOREM. *The quotient system is the quantum dynamical system* $(H, \hat{\Delta})$ *where* H *is union of the Hilbert spaces* $\mathscr{H}_u$ *where each* $\mathscr{H}_u$ *is* $L^2(P, \boldsymbol{C}^d)$, $d=$ *degree of* u, *and the dynamorphism is defined in each* $\mathscr{H}_u$ *by* 8.22.

There is an interesting thing here. Whereas for the original system there was only one Hilbert space, for the quotient system there are several, i.e., a "superselection rule" operates [11, 136].

These ideas apply to the systems for the rigid body described in §7. The group of geometric symmetries is $O(3)$. Let us consider the system involving $\mathscr{H}(-)$, and study the equivalent form given in 7.5. Thus the Hilbert space is $\mathscr{H}(-)$, the set of functions in $L^2(R^3 \times SU(2) \times \{-1, 1\})$ for which $\varphi(p, -g, \xi)=-\varphi(p, g, \xi)$. Here the group of isomorphisms to be divided out is the group of right transformations by $G=SU(2) \times \{-1, 1\}$.

Take $\mathscr{U}'$ to be those continuous matrix representations u of G for which $u(N)=-u(e)$, where $e=(1, 1)$ is the identity of G. We form $\mathscr{H}(\mathscr{U}')$ as explained above. Then obviously $\mathscr{U}'$ is the $\mathscr{U}(-)$ defined in the last section and

(8.5) $\mathscr{H}(\mathscr{U}')$ *is* $\mathscr{H}(-)$.

Moreover,

8.51 *each $\mathcal{H}(u)$ is invariant under the dynamorphisms.*

Also (cf. 8.11)

8.52 *for u in $\mathcal{U}'$, $\bar{u}$ is equivalent to u.*

The reason is this. If G were merely $SU(2)$, then u would be equivalent to $\bar{u}$ because they have the same degree and for $SU(2)$ there is only one equivalence class for each degree. For our G this is not the case: it is easy to see that for each representation of $SU(2)$ there are two of G. But only one is in $\mathcal{U}'$.

Thus 8.4 holds for the quantum dynamical system for a rigid body given by $\mathcal{H}(-)$. (It also holds for $\mathcal{H}(+)$.)

We now take a closer look at the quotient system given by 8.3 for the single representation u_0 of G where u_0 is that representation in $\mathcal{U}'$ which reduces on $SU(2)$ to the identity representation of $SU(2)$, i.e.,

$$(8.6) \qquad u_0(g, \pm 1) = \pm g \quad \text{for each g in } SU(2).$$

One can verify (cf. 8.11) that

$$(8.61) \qquad \bar{u}_0 = \sigma_2 u_0 \sigma_2^{-1} \quad \text{where } \sigma_2 = \begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix}.$$

Time-reversal in the system 7.5 and in $\mathcal{H}(u_0)$ is given by

$$\varphi(p, g, \xi) \to \overline{\varphi(-p, g, \xi)}.$$

We want to get the corresponding operator in $\mathcal{H}_{u_0} = L^2(R^3, C^2)$. The μ in 8.21 for this is clearly 1. Calculating on the basis of 8.24, we obtain

8.62 *time-reversal in $\mathcal{H}_{u_0}$ sends*

$$\begin{bmatrix} \psi_1(p) \\ \psi_2(p) \end{bmatrix} = \psi(p)$$

into $\sigma_2 \overline{\psi(-p)}$, the entries of which are $i\overline{\psi_2(-p)}$ and $-i\overline{\psi_1(-p)}$.

We could just as well omit the i. Either way, the square is minus 1.

Space-inversion in 7.5 and in $\mathcal{H}(u_0)$ is given by the operator $U(i_s)$ calculated from 7.3 for S being space inversion. Actually, we are working in $L^2(SCE(3))$, and not $L^2(E(3))$, as 7.3 envisions, and according to the nature of the equivalence set up in 7.5 we calculate $U(0, \pm 1, -1)$ where U is the representation in 6.42. The operator corresponds to left-multiplication by $(0, \pm 1, -1)$. Thus in $\mathcal{H}(u_0)$, space-inversion sends $\varphi(p, g, \xi)$ into $\varphi(-p, \pm g, -\xi)$ and since we are in $\mathcal{H}(-1)$, these are $\pm \varphi(-p, g, -\xi)$. We had, of course, the *right* to expect these two vectors to be linearly dependent. The μ in 8.22 is 1 in this case also (as it is for every dynamorphism in this system). Using 8.23 for the operator $D: \varphi \to \varphi(-p, g, -\xi)$ gives $\hat{D}$ which sends ψ into $-\psi(-p)$. In any case,

8.63 *space-inversion in $\mathcal{H}_{u_0}$ sends $\psi(p)$ into $\pm \psi(-p)$.*

We must now observe that if σ_1, σ_2, σ_3 are the Pauli matrices and y^1, y^2, y^3 real then exp $[y^j\sigma_j/2i]$ is in $SU(2)$ and its image in $SO(3)$ is the rotation whose matrix is exp $(y^j s_j)$ where s_j is the matrix which has ε_{jkm} in the kth column and the mth row (cf. [4, A.5]). This matrix corresponds to the $\varepsilon_{jkm}x^k(\partial/\partial x^m)$, as can be shown by going back to 4.2. From this we can discover the operator $\hat{D}$ for the case of $D = X_m$, the X_m being the one defined in 6.42 and thus a vector field on $SU(2)$. From 8.23 it follows that $\hat{X}_m\psi = (X_m u_0)_e^*\psi$. The matrix $(X_m u_0)_e$ is the rate of change of u_0 at e when one follows a one-parameter subgroup whose tangent is $(X_m)_e$. The matrix corresponding to X_m in $O(3)$ is $-s_m$, as already remarked just after 6.81. Thus we have to differentiate the unitary matrix corresponding to exp $(-\varepsilon s_m)$, namely exp $[-\varepsilon\sigma_m/2i]$, at $\varepsilon = 0$. Therefore (since s_m is self-adjoint)

$$(8.64) \qquad\qquad\qquad \hat{X}_m = -i\sigma_m/2.$$

Looking at this and at 6.41, but thinking about 6.42, we see that
8.65 *the operator corresponding to* $-x^j(\partial/\partial x^0) - x^0(\partial/\partial x^j)$ *is*([10])

$$S\frac{\partial}{\partial p_j} + \frac{p_j}{2S} - \frac{i\varepsilon_{jkm}p_k\sigma_m}{2(S_0 + S)};$$

8.66 *the operators corresponding to 6.1 are* ip_j *and* $-i\sigma_j/2 - \varepsilon_{jkm}p_k(\partial/\partial p_m)$, *and*
8.67 *the operator corresponding to* $\partial/\partial x^0$, *thus giving the infinitesimal dynamorphism for infinitesimal time translation is the multiplication operator* iS.

8.7 THEOREM. *The quotient system of the quantum dynamical rigid-body system 7.4 when (that quotient system is) restricted to* $\mathcal{H}(u_0)$, *is isomorphic to the Foldy-Wouthuysen system restricted to the*
8.71 *negative-energy subspace when* $m = S_0$ *and*
8.72 *only orthochronous changes of observer are considered.*

The truth of this consists in comparing 8.65–8.67 with 9.5 and 8.63 with 9.6.
We can easily obtain a system in which 8.71 is replaced by
8.73 *positive-energy subspace.*
To do so we replace the system 7.4 by a system based not on the representation U given by 6.42 but on the representation U' defined by $U'(g) = T^{-1}U(g)T$ where T is the time-reversal operator given in 8.62,

$$\psi \to \sigma_2\overline{\psi(-p)}.$$

This changes the signs in 8.65 and 8.67.

([10]) Time-reversal anticommutes with the infinitesimal Lorentz transformation in 8.65. Hence the composite of the operator in 8.65 with 8.62 should be at worst a scalar multiple of the composite in the opposite order. One may wonder how that can be, when σ_2 occurs in 8.62 and the σ_m appear symmetrically in 8.65. The resolution of this paradox is that σ_m is real for $m = 1, 3$ and pure-imaginary for $m = 2$.

We may form the direct sum $U' + U$ of these two representations and obtain a representation equivalent to the F.-W. representation, restricted as before to $\mathscr{P}\uparrow$.

The system 4.3 after which to some extent 7.4 was patterned, turns out really to be more closely related to the system involving U' which we just mentioned. The reason for our somewhat *gauche* discussion is that the "Rectification" idea is more immediately applicable to 7.4 than its reversed-energy form.

We must remark that when observables are properly assigned to the quotient system here discussed (8.7) that it does not represent an elementary particle unless the constant θ appearing in 6.4 and 6.42 is 0. For if θ is positive, then the velocity of the centroid has an upper bound less than 1, and 1 is the speed of light here.

Setting $\theta = 0$ makes the mass equal to 1. This is not essential. A separate parameter can be included from the beginning to make the mass any desired value.

9. The Foldy-Wouthuysen representation.

The theory of the Dirac equation provides us with a unitary representation of the "inhomogeneous $SL(2, C)$" [15, p. 14]. This group is the simply-connected covering group of $\mathscr{P}\uparrow_{+}$, the restricted Poincaré group. The homomorphism is 2:1. The Lie algebra (of infinitesimal transformations) for $\mathscr{P}$ is isomorphic to that of the inhomogeneous $SL(2, C)$. Hence we may describe a unitary representation of the inhomogeneous $SL(2, C)$ by telling what skew-adjoint operators it assigns to the infinitesimal Poincaré transformations (4.26).

Foldy and Wouthuysen transformed the Dirac representation to a form making evident the spectrum of the Dirac Hamiltonian. It is possible to compute the transforms of all the operators assigned by the Dirac representation to the infinitesimal Poincaré transformation with the following result in which m is a positive number.

9.1 THEOREM. *The Foldy-Wouthuysen representation takes place in the Hilbert space $L^2(R^3, C^4)$ of square-integrable functions on ordinary Cartesian 3-space, whose values are 4-component complex column vectors. Let the Cartesian coordinates in R^3 be p_1, p_2, p_3 and let $E = (m^2 + p_1^2 + p_2^2 + p_3^2)^{1/2}$. Then the skew-adjoint operators corresponding to the ten infinitesimal Poincaré transformations (4.26) are*

$$(9.2) \qquad -i\beta E, \quad ip_1, \quad ip_2, \quad ip_3$$

$$(9.3) \qquad \tfrac{1}{4}\varepsilon_{jkm}\gamma^k\gamma^m - \varepsilon_{jkm}p_k(\partial/\partial p_m)$$

$$(9.4) \qquad \begin{aligned} & -\beta\left[E(\partial/\partial p_1) + p_1/2E + \frac{\gamma^1(\gamma^2 p_2 + \gamma^3 p_3)}{2(m+E)}\right] \\ & \quad .\quad .\quad .\quad .\quad .\quad .\quad .\quad .\quad .\quad .\quad . \\ & -\beta\left[E(\partial/\partial p_3) + p_3/2E + \frac{\gamma^3(\gamma^1 p_1 + \gamma^2 p^2)}{2(m+E)}\right]. \end{aligned}$$

Here the β's and γ's are exactly those defined in [13, p. 69, equations (24), (26), (27)]. The first operator in 9.11 is of course the one calculated by Foldy-Wouthuysen, by the method explained in [13]. There is a minor difference: we prefer to

deal with skew-adjoint operators, because in either convention, commutators are skew-adjoint and thus we need no patching up with i's. There is also a difference in sign when 9.2 is compared with [13, p. 93, equation (183)], but that is merely because H' is the *negative* of the generator of time translations. The calculations resulting in 9.3 begin with the observation of [13, p. 78, (80)]. They are very tedious and hence are omitted here, as are those resulting in 9.13 which have their beginning in [13, p. 78, (82)]. In every case one takes the appropriate operator for the Dirac representation and performs $e^{iS} \cdots e^{-iS}$ as directed by Foldy-Wouthuysen.

Actually, [13] is concerned with functions of four variables and 9.1 talks about functions on R^3. The connection is that a function of four variables, presumably a solution of Dirac's equation is associated with its restriction to the three-space $t = 0$. This association does not help us with *time*-reversal. Accordingly we do not presume to assert what the F.-W. representation does about that.

Let us write down 9.2–9.4 for the subspace of functions $(0, 0, \psi_3, \psi_4)$ where β is -1. Calculations give

$$iE, \quad ip_1, \ldots, ip_3, \quad -i\sigma_j/2 - \varepsilon_{jkm}p_k(\partial/\partial p_m),$$

(9.5) $\quad E(\partial/\partial p_1) + p_1/2E - i(p_2\sigma_3 - p_3\sigma_2)/2(m+E), \ldots, E(\partial/\partial p_3)$
$$+ p_3/2E - i(p_1\sigma_2 - p_2\sigma_1)/2(m+E).$$

When space inversion is represented by the choice made as in [13, 107], the transform $e^{iS} \cdots e^{-iS}$ yields (apart from a factor i which is of no consequence for projective representations)

9.6 *space inversion sends ψ to $\pm\beta\psi(-p)$.*

10. **Quantizing a classical system.** By a classical system (K, Δ) we will mean a completely Hamiltonian system (3.2–3.24) where K is in fact $T_1(Q)$ where Q is a finite dimensional manifold. This terminology is justified because such systems are the usual point of departure when a quantum system is to be constructed. When a quantum alternative to the given system is proposed, it invariably contains a positive parameter h (the exact value of which is left perhaps to experimental determination). When it is argued that a quantum system (H, Θ) has a classical system (K, Δ) as its correspondent, the argument usually involves letting the h tend to zero. It is thus reasonable only to define when a sequence or *one-parameter family* of quantum systems corresponds to a given classical system.

Let (H, Θ) be a one-parameter family of quantum systems (the parameter h being not expressly indicated) where H consists of only one Hilbert space which is (for every value of h) the intrinsic Hilbert space $\mathscr{H}_Q$ of some manifold Q which is the configuration space for a classical system (K, Δ). Thus (for each value of h) and each pair of coordinators $\mathbf{z}, \mathbf{y}, \Theta_{\mathbf{z}}^{\mathbf{y}}$ is unitary or antiunitary in $\mathscr{H}_Q$. We suppose that, for S near the identity in the space-time group, the operator $U_{\mathbf{z}}(S)$ (see 2.5) is in fact unitary and that, moreover,

10.1 *for each continuous one-parameter group $T(s)$ there is a skew-adjoint operator* $\Theta'_{T,z}$ *in* $\mathcal{H}_Q$, *such that for* φ *in the domain of that operator,*

$$\frac{d}{ds} U_z(T(s))(\varphi)\Big|_{s=0} = \Theta'_{T,z}(\varphi).$$

This is simply a more precise form of what we have already required in 2.63 for every dynamical system.

Given the quantum system (H, Θ) we can construct a new system $(H, \overline{\Theta})$ where each dynamorphism $\overline{\Theta}^x_x$ is $U(i_t) \circ \Theta^x_x \circ U(i_t)$. We have met this construction before, in passing from 8.71 to 8.73. The infinitesimal dynamorphisms Θ' figure in 10.21 below.

We will assume that each complex valued function φ on Q which is infinitely differentiable and compactly supported is in the domain of each $\overline{\Theta}'_{T,z}$.

The classical system (K, Δ) also has a time-reversed variant $(K, \overline{\Delta})$. We will say that

10.2 *the one-parameter family* (H, Δ) *is a quantization of the classical system* (K, Δ) *if the following is true.*

Let z be a coordinator and T a one-parameter subgroup of the space-time group. Let $\bar{g}_{T,z} = \bar{g}$ be the generator (3.22) corresponding to T, z in the time-reversed system $(K, \overline{\Delta})$. Let $\overline{\Theta}'_{T,z} = \overline{\Theta}$ be the infinitesimal dynamorphism assured by 10.1 for the system $(H, \overline{\Theta})$. Then for φ, ψ, and w any infinitely differentiable functions on Q with compact support, and w being *real*, the limit

(10.21) $$\lim_{h \to 0} (-ih\overline{\Theta}[\exp(-iw/h)\varphi]; \exp(-iw/h)\psi)$$

exists and has the value $([\bar{g} \circ dw]\varphi; \psi)$. The small circle here means substitution. We present first a very simple example.

Schrödinger's equation has the form $ih(\partial\varphi/\partial t) = E\varphi$. Let us take $Q = R$. For such system, $\Theta'_{\partial/\partial t, x}$ is[11] $-E/ih$, as pointed out at the end of §8. Now suppose E is a polynomial, with real coefficients, in the operator $-ih(\partial/\partial x)$. It will suffice to treat the case $E = (-ih(\partial/\partial x))^n$ for $n = 0, 1, 2, \ldots$. For the time-reversed system we clearly have $\overline{\Theta} = (ih)^{-1}(ih(\partial/\partial x))^n$.

Expression 10.21 requires us to calculate

$$-(-(h/i)\partial/\partial x)^n[\exp(-iw/h)\varphi].$$

It is evident that the only term of this which does not tend to 0 as $h \to 0$ is $-(w')^n \exp(-iw/h)\varphi$. Thus the limit 10.21 is $(-(w')^n\varphi; \psi)$. Let $\bar{g}$ be the (real!) function on the phase space $T_1(R)$ defined by $-p^n$ where p as usual is the function for which $p(bdx) = b$. Then $\bar{g}(dw(a)) = -[p(dw(a))]^n = [p(w'(a)dx)]^n = -[w'(a)]^n$. Thus $\bar{g} \circ dw = -(w')^n$. Now $\bar{g}$ is presumably the generator of time translation for the

---

[11] We will abbreviate this to Θ.

system $(T_1(R), \bar{\Delta})$. Thus the generator of time translations for $(T_1(R), \Delta)$ is $-\bar{g} = p^n$. We thus see that if the Schrödinger equation is

$$\frac{h}{i}\frac{\partial \varphi}{\partial t} + \left[\frac{1}{2}\left(\frac{h}{i}\frac{\partial}{\partial x}\right)^2 + V\right]\varphi = 0,$$

then the Hamiltonian of the classical limit is $\frac{1}{2}p^2 + V$.

Spatial dynamorphisms have also to be examined. For *all* such simple quantum systems the (skew-adjoint) generator for translation in the positive x direction is $\partial/\partial x$, independent of the dynamics. Thus Θ and $\bar{\Theta}$ are $\partial/\partial x$. For the classical systems $(T_1(R), \Delta)$ and $(T_1(R), \bar{\Delta})$ the generating function is given by 3.7 and is $-p$ in each case, of course. The truth of 10.21 is easily verified.

The criterion 10.2 suggests a way of proceeding from a classical system $(T_1(Q), \Delta)$ to a (one-parameter) quantum system $(\mathcal{H}_Q, \Theta)$ which frequently does have $(T_1(Q), \Delta)$ as its limit. The rule is as follows.

10.3 *Corresponding to a basis for the Lie algebra of the space-time group, write down the list of generating functions for the given system.*

10.31 *Modify these appropriately to serve for the time-reversed (still classical) system.*

10.32 *Replace each p_j in the expressions of 10.31 by $ih(\partial/\partial x^j)$ and arrange these to form formally self-adjoint operators.*

10.33 *Divide these operators by $-ih$, and use them to define infinitesimal dynamorphisms for a quantum system $(\mathcal{H}_Q, \bar{\Theta})$.*

10.34 *Define time-reversal for $(\mathcal{H}_Q, \bar{\Theta})$ by $\varphi \to \bar{\varphi}$.*

10.35 *Conjugate the dynamorphisms of $(\mathcal{H}_Q, \bar{\Theta})$ with time-reversal to produce the desired system $(\mathcal{H}_Q, \Theta)$.*

This rule (10.3 etc.) preserves bracket relations to a considerable extent, although of course[12] not perfectly. The carrying out of step 10.32 involves so much guessing that a mathematical, analytic, rather than formal, check such as 10.2 provides is surely indispensable.

We will apply our rule (10.3 etc.) to the system 4.3. We first define a Hermitean operator S (cf. 6.4).

10.4 *S_0 is the positive square root of $1 - \theta h^2(X_1 \circ X_1 + \cdots + X_3 \circ X_3)$; and S is the positive square root of $S_0^2 - h^2[(\partial/\partial x^1)^2 + \cdots + (\partial/\partial x^3)^2]$.*

The precise mathematical specification of these operators, and those given below is to be along the lines of 6.4, the operators of which are the generators of certain one-parameter groups whose existence was proved in §6.

---

(12) See [3]. In [3] the classical bracket is defined in the *original* way (Enzyklopädie der Math. Wiss., Band II, 1. Teil, 1. Haelfte; S. 333 (63)), which is the negative of the "modern" way followed in the present paper. The original way is, as a matter of fact, more elegant. It would avoid the minus sign in 3.7, thus allowing $P = p$ in 4.5; and avoid prefixing opposite signs to the two occurrences of ih in 10.32 and 10.33.

The application of 10.3 (etc.) to 4.3 leads to the following formulae for the infinitesimal dynamorphisms to be assigned to the infinitesimal Poincaré transformations in the list (except for a change of indices i to j) 4.26, and in the order listed.

(10.41) $$-S/ih,$$

(10.42) $$\partial/\partial x^j,$$

(10.43) $$X_j + \varepsilon^{jkm} x^m (\partial/\partial x^k),$$

(10.44) $$-i(x^j S + S x^j)/2h - ih(S + S_0)^{-1} \varepsilon^{jkm} X_m (\partial/\partial x^k).$$

To allow a comparison with 6.4 we apply the partial Fourier transform, which makes $\partial/\partial x \to ip$ and $x \to i(\partial/\partial p)$. (We do not allow the parameter h to play a role in the definition of Fourier transform!)

This leads to

10.45 $S_0 \to S_0$ but $S \to$ *the positive square root of* $S_0^2 + h^2 \mathbf{p} \cdot \mathbf{p}$.

For the operators 10.31–10.34 we obtain, respectively,

(10.46) $$iS/h,$$

(10.47) $$ip_j,$$

(10.48) $$X_j - \varepsilon^{jkm} p_k (\partial/\partial p_m),$$

(10.49) $$\frac{1}{2h} \left(S \frac{\partial}{\partial p_j} + \frac{\partial}{\partial p_j} S \right) + \frac{h \varepsilon^{jkm} p_k X_m}{S_0 + S}.$$

This evidently agrees with 6.4 for $h = 1$. For $h \neq 1$ we can still fall back on 6.4 to prove the existence of a system satisfying 10.46–10.49. We have only to make a suitable change of variables.

It remains, however, to prove that 10.41–10.44 really does provide a quantization of 4.3. The very complexity of the proof that it does, shows how inadequate is 10.3 (etc.) and how vital is 10.2.

We will limit the discussion to the operators 10.41 and 10.44, since 10.21 holds for *differential* operators formed from polynomials in the p's by the rule 10.3 etc. Moreover, the methods we will use for 10.44 can easily be seen to work for 10.41.

The first term of 10.49 is evidently the same as $h^{-1} S(\partial/\partial p_j) + h(2S)^{-1} p_j$. Here the latter term is a bounded operator times h and hence contributes 0 to the limit 10.21. Therefore we may replace 10.44 by

(10.5) $$-ix^j S/h - ih(S + S_0)^{-1} \varepsilon^{jkm} (\partial/\partial x^k) X_m.$$

(The operators X_m, $\partial/\partial x^k$, S_0, and S commute with each other.) The Lorentz transformation associated with 10.44, as well as the generating function K_j, change sign under time-reversal. The $\overline{\Theta}$ to be inserted into 10.21 is the negative of the

operator 10.5; and the $\bar{g}$ is $-K_j$. Let us, however, study only the second term of 10.5. If we define

(10.51) $$I_h(\varphi) = -h^2 \varepsilon^{jkm}(S + S_0)^{-1} X_m(\partial/\partial x^k)$$

and let $g = \varepsilon^{jkm} p_k w_m (H + H_0)^{-1}$, then it will suffice (for that second term) to prove

(10.52) $$\lim (I_h[\exp(-iw/h)\varphi]; \exp(-iw/h)\psi) = ([g \circ dw]\varphi; \psi).$$

We proceed to prove 10.52. Let P_h be the differential operator

$$\varepsilon^{jkm} \frac{h}{i} \frac{\partial}{\partial x^k} \frac{h}{i} X_m S^6$$

which has obviously been obtained from $g(H + H_0)H^6$ by the process 10.32. Then $I_h = P_h A_h$ where $A_h = (S + S_0)^{-1} S^{-6}$. The reason for the exponent 6 is that 6 is even and large enough to allow A_b to have the following integral representation.

10.6 LEMMA. *For each z in E(3), and each φ in $L_2(E(3))$,*

$$A_h(\varphi)(z) = \iiint\iiint \varphi\, [\exp(-h\mathbf{s}\cdot\mathbf{W}) \exp(-h\mathbf{t}\cdot\mathbf{X})z]k(\mathbf{s}, \mathbf{t})\, d\mathbf{s}\, d\mathbf{t},$$

where each integration goes from $-\infty$ to $+\infty$. Here

$$\mathbf{s} = (s_1, s_2, s_3), \qquad \mathbf{W} = (\partial/\partial x^1, \partial/\partial x^2, \partial/\partial x^3)$$

where these three are the part of the basis of the Lie algebra of E(3) corresponding to the infinitesimal translations. For each Y in the Lie algebra, $\exp(Y)$ is the element of E(3) generated by Y [14, 223]. Similarly, $\mathbf{t}\cdot\mathbf{X} = t_1 X_1 + \cdots + t_3 X_3$ where X_1, X_2, X_3 are the generators (4.21). As to the kernel,

(10.61) $$K(s, t) = \frac{2}{h\tau} \sin\frac{h\tau}{2} (2\pi)^{-6} \iiint\iiint e^{k\mathbf{s}\cdot\mathbf{a} + it\cdot\mathbf{b}} \lambda\left(\frac{\beta}{h} - \frac{1}{2}, \frac{\mathbf{a}}{h}\right) d\mathbf{a}\, d\mathbf{b}$$

where $\beta = [(b_1)^2 + \cdots + (b_3)^2]^{1/2}$,

(10.62) $$\lambda(j, \mathbf{c}) = [H(j, \mathbf{c}) + H(j, 0)]^{-1} H(j, \mathbf{c})^{-6},$$

and $H(j, \mathbf{c}) = [1 + h^2\mathbf{c}\cdot\mathbf{c} + h^2\theta j(j+1)]^{1/2}$.

Proof. The first thing to observe is that K is indeed absolutely integrable on R^6, This follows from the fact that the function λ is essentially the reciprocal of the 7th power of the distance to the origin in R^6.

We now observe that A_h commutes with right multiplication in $E(3)$ and so does the operator defined by the left side of the formula proposed. Hence it suffices to prove it for $z =$ the identity element of $E(3)$.

Let χ_j $(j = 0, 1/2, 1, \ldots)$ be a character of the irreducible representation of $O(3)$ which has degree $2j + 1$ [13, 25]. Let ψ be any continuous function on R^3 with

compact support. Then $\psi\chi_j = \varphi$ is a function defined on $E(3)$ because of the presentation of $E(3)$ as a product $R^3 \times O(3)$. Moreover, if 10.6 holds for all such products, then it holds in general.

The reader must verify that for $z = xy$, x in R^3 and y in $O(3)$, we have (actually, by definition of A_h!)

$$A(\psi\chi_j)(xy) = (2\pi)^{-3}\chi_j(y) \iiint \lambda(j, \mathbf{c})e^{i\mathbf{c}\cdot x}\Psi(\mathbf{c}) \, d\mathbf{c}$$

where Ψ is the Fourier transform of ψ. We now take $x = 0$, $y = 1$ and equate it to the right-hand side of 10.6. To evaluate the latter, one must use the fact that

$$\psi[\exp(-h\mathbf{s}\cdot\mathbf{W})] = (2\pi)^{-3} \iiint e^{-h\mathbf{s}\cdot\mathbf{c}}\Psi(\mathbf{c}) \, d\mathbf{c}$$

and that

$$\chi_j[\exp(-h\mathbf{t}\cdot\boldsymbol{\chi})] = \frac{\sin h\tau(j+\frac{1}{2})}{\sin h\tau/2}$$

where $\tau = [(t_1)^2 + \cdots + (t_3)^2]$. This fact about χ_j results from the observation that χ_j is invariant under inner automorphisms and hence the rotation $\exp(-h\mathbf{t}\cdot\boldsymbol{\chi})$ might as well be about the z-axis. The rest of the argument is routine Fourier-transform theory.

It is rather obvious that A_h has a representation $A_h(\varphi)(z) = \int \varphi(w^{-1}z)K(w) \, dw$ where this integration is over $E(3)$. However, such a representation would not allow us to see what happens as $h \to 0$. (The reader should try to prove 10.52 on the unit circle with $I_h = [1 - h^2(\partial/\partial\theta)^2]^{1/2}$ with such a representation!)

The six-fold integral in 10.61 is almost independent of h. The only reason it is not is that $H(j, \mathbf{c})$ does not quite depend on $h^2 j^2$ and $h^2 c^2$ alone. It is easy to see that, as $h \to 0$, $K(s, t)$ tends to a limiting function

$$K_0(s, t) = (2\pi)^{-6} \iiiint\!\!\!\int \exp(i\mathbf{s}\cdot\mathbf{a} + i\mathbf{t}\cdot\mathbf{b})k(\mathbf{b}, \mathbf{a}) \, d\mathbf{a} \, d\mathbf{b}$$

where $k(\mathbf{b}, \mathbf{a}) = [S(\mathbf{a}, \mathbf{b}) + S(0, \mathbf{b})]^{-1}S(\mathbf{a}, \mathbf{b})^{-6}$ where $S(\mathbf{a}, \mathbf{b}) = (1 + \mathbf{a}\cdot\mathbf{a} + \theta\mathbf{b}\cdot\mathbf{b})^{1/2}$. The limiting kernel K_0 is also absolutely integrable and moreover, K_0 and the K (for sufficiently small b) are dominated by a summable function M.

The remainder of the proof of 10.52 uses merely this dominated convergence. The presentation will be easier to follow if we give the argument for the case in which the configuration space is the circle([13]) group $SO(2)$ rather than $E(3)$. We would then be knowing that

$$(10.63) \qquad A_h(\varphi)(z) = \int_{-\infty}^{\infty} \varphi[\exp(-htX)z]K(t) \, dt$$

([13]) The circle group is more instructive than the case of the real line in that it is remarkable that in the former case the integral in 10.63 is not extended over the group.

where X is the infinitesimal rotation which makes $\exp(-htX) = e^{-iht}$, $|z| = 1$, and $K \to K_0$ under the domination of some summable function M. Moreover,

$$K_0(t) = (2\pi)^{-1} \int e^{ias} k(a) \, da,$$

where $A_h = k((h/i)\partial/\partial y)$, y being the angular coordinate in $SO(2)$. Thus $z = e^{iy}$!

Let φ, ψ, and w be functions qualified as in 10.2 and let us form (compare 10.52) the inner product $(I_h[\exp(-iw/h)\varphi]; \exp(-iw/h)\psi)$. This is the same as

(10.64) $(A_h[\exp(-is/h)\varphi]; P_h[\exp(-iw/h)\psi])$.

Here $P_h = p(-(h/i)\partial/\partial y)$ for some polynomial whose coefficients are real and independent of h. Clearly,

$$P_h[\exp(-iw/h)\psi] = \exp(-iw/h)[p(\partial w/\partial y)\psi + \omega_h]$$

where ω_h is a linear combination of continuous functions (with compact support) with coefficients that vanish for $n = 0$.

Expression 10.64 is therefore a linear combination of several integrals of the form

(10.65) $\int_0^{2\pi} e^{-iw/h} A_h[\exp(-iw/h)\varphi(e^{iy})] f(y) \, dy$

with coefficients that vanish when h does, except for one term whose coefficient is 1, in which

$$f(y) = p(w') \overline{\psi(e^{iy})}.$$

The expression $A_h[\cdots]$ in 10.65 is, according to 10.63,

$$\int_{-\infty}^{\infty} \exp\left[-\frac{i}{h} w(e^{iy-iht})\right] \varphi(e^{iy-iht}) K(t) \, dt.$$

Consequently 10.65 takes the form of a double integral

$$\int_0^{2\pi} \int_{-\infty}^{\infty} \exp\left\{-\frac{i}{h}[w(e^{iy-iht}) - w(e^{iy})]\right\} \varphi(e^{iy-iht}) K(t) f(y) \, dt \, dy.$$

The entire integrand is dominated by $BM(t)|f(y)|$ where B is an upper bound for $|\varphi|$. The integrand has a limit as $h \to 0$, namely

$$\exp\{it(\partial w/\partial t)(e^{iy})\} \varphi(e^{iy}) K_0(t) f(y) \, dt \, dy.$$

Carrying out first the integration over t we see that the limit of 10.65 exists and equals

$$\int_0^{2\pi} k\left(-\frac{\partial w}{\partial y}\right) \varphi(e^{iy}) f(y) \, dy.$$

Hence the limit of 10.64 is

$$\int_0^{2\pi} k(-w') \varphi(e^{iy}) p(w') \overline{\psi}(e^{iy}) \, dy.$$

Recall that $I_h = k((h/i)\partial/\partial y) p(-(h/i)\partial/\partial y)$. Thus 10.52 is proved.

10.7 THEOREM. *The quantum system* 10.41–10.44 *is a quantization of the classical system* 4.3.

We close with a remark about observables.

Let A be a (one-parameter family of) self-adjoint operator of the (one-parameter family of) quantum system (H, Θ). Let f be a differentiable function on $T_1(Q)$, i.e. an observable for the classical system. Then we will say

$$(10.8) \qquad\qquad\qquad\qquad\qquad A \rightarrow f$$

if for φ, ψ, and w as in 10.21,

$$(10.9) \qquad\qquad \lim_{h \to 0} (A[\exp{(iw/h)\varphi}]; \exp{(iw/h)\psi}) = ([f \circ dw]\varphi; \psi).$$

It is easily verified for the simple example discussed above that $(-ih(\partial/\partial x))^n \rightarrow p^n$. If Θ is some infinitesimal dynamorphism, then $-ih\Theta$ is a quantum observable. One might be tempted therefore to suppose that if $\Theta \rightarrow g$ in the sense of 10.2, then also $-ih\Theta \rightarrow g$ in the sense just defined. This is true for infinitesimal motions of the configuration space, but it breaks down for spatio-temporal transformations. For the infinitesimal Poincaré transformation $\partial/\partial t + \partial/\partial x$ the discrepancy is not merely an over-all change of sign.

11. Alternating structures: local theory. Let A be an alternating structure on a manifold K. By $v(A)$ we mean the linear space of all vector fields X_h (see 3.14). Evidently (from 3.15) if X_1, X_2 belong to $v(A)$, so does $[X_1, X_2]$. For each point ξ of K, let $v_\xi(A)$ be the linear space of vectors obtained from evaluating the fields of $v(A)$ at ξ. In the usual terminology (but not that of [14, 130]) we have here an involutive distribution. Let us make the assumption

11.1 *The dimension of $v_\xi(A)$ is independent of ξ.*

11.12 THEOREM. *If* 11.1 *holds, then, given a point ξ of K, there is a coordinate system $x^1, \ldots, x^n$ defined in a neighborhood of ξ such that*

$$\{f, g\} = \frac{\partial f}{\partial x^1}\frac{\partial g}{\partial x^2} - \frac{\partial f}{\partial x^2}\frac{\partial g}{\partial x^1} + \frac{\partial f}{\partial x^3}\frac{\partial g}{\partial x^4} - \frac{\partial f}{\partial x^4}\frac{\partial g}{\partial x^3} + \cdots + \frac{\partial f}{\partial x^{p-1}}\frac{\partial g}{\partial x^p} - \frac{\partial f}{\partial x^p}\frac{\partial g}{\partial x^{p-1}}$$

(and, obviously, p is even).

Suppose that dimension is p. Let ξ be a point of K. By Frobenius' theorem [14, 132] there is a coordinate system $x^1, \ldots, x^n$ defined in a neighborhood of ξ such that $v(A)$ is generated by $\partial/\partial x^1, \ldots, \partial/\partial x^p$. In terms of these coordinates $A^{ij} = 0$ if i or j is greater than p. On the other hand, the matrix $(A^{ij} : 1 \leqq i, j \leqq p)$ has rank p and hence is nonsingular. Let us define (B_{ij}) as the inverse matrix and extend the definition by making $B_{ij} = 0$ whenever i or j exceeds p. Now we assert that the differential form $B_{ij}\, dx^i \wedge dx^j = \Omega$ is closed.

To see this, observe that $B_{mj}A^{jk}$ is constant (indeed, it equals δ_m^k for $m, k \leqq p$

and is 0 otherwise), whence $B_{mj\alpha}A^{jk} + B_{mj}A^{jk}_\alpha = 0$ where the third suffix indicates a partial derivative. From this one concludes that $A^{mk}_\alpha = A^{m\beta}A^{k\gamma}B_{\beta\gamma\alpha}$.

Now insert $f = x^i$, $g = x^j$, $h = x^k$ into 3.13. The result is $SA^{i\alpha}A^{jk}_\alpha = 0$ where $S(\cdots)$ applied to any 3 index symbol, say C^{ijk}, is $C^{ijk} + C^{jki} + C^{kij}$. It follows that $S(A^{i\alpha}A^{j\beta}A^{k\gamma}B_{\beta\gamma\alpha}) = 0$. From this, $A^{i\alpha}A^{j\beta}A^{k\gamma}S(B_{\beta\gamma\alpha}) = 0$ whence $S(B_{\beta\gamma\alpha}) = 0$. This is the condition that Ω is closed. According to [14, 140] there is a coordinate system $y^1, \ldots, y^n$ which gives to Ω the form $dy^1 \wedge dy^2 + \cdots + dy^{p-1} \wedge dy^p$. Let us call this $b_{\alpha\beta}\, dy^\alpha \wedge dy^\beta$. We have thus two ways to evaluate $\langle \Omega; U, V \rangle$ for vectors U and V. Let $U = \partial/\partial y^\gamma$ where $\gamma \leq p$. Then $\langle \Omega; U, V \rangle$ is zero according to the former expression for Ω, while according to the new one it is $\pm \partial y^\gamma/\partial^q$. Thus $\partial y^\gamma/\partial x^q$ is 0 for $\gamma \leq p < \omega$. In a similar manner, we compare the two evaluations for $\langle \Omega; \partial/\partial y^q, \partial/\partial x^\gamma \rangle$. This shows that $\partial x^\gamma/\partial y^q$ is 0 for $\gamma \leq p < q$. It follows from this that we can replace $y^{p+1}, \ldots, y^n$ by $x^{p+1}, \ldots, x^n$ and still have a coordinate system on some neighborhood of ξ. This does not change the new form of Ω, so let us just say that $y^q = x^q$ for $q > p$.

Denote by $a^{\alpha\beta}$ the components of A relative to $y^1, \ldots, y^n$. From the previous assertion about $B_{mj}A^{jk}$ one can deduce that

$$(11.2) \qquad b_{\alpha\beta}a^{\beta\gamma} = \sum_{i=1}^{p} \frac{\partial x^i}{\partial y^\alpha}\frac{\partial y^\gamma}{\partial x^i} \equiv \delta^\gamma_\alpha - \sum_{i=p+1}^{n} \frac{\partial x^i}{\partial y^\alpha}\frac{\partial y^\gamma}{\partial x^i}.$$

(There is a sum on the left, too.) In the last sum here we can write $\partial y^i/\partial y^\alpha$ for $\partial x^i/\partial y^\alpha$. Hence this last sum is 0 if $\alpha \leq p$, whence $b_{\alpha\beta}a^{\beta\gamma} = \delta^\gamma_\alpha$ if $\alpha \leq p$. This tells us exactly what the $a^{\beta\gamma}$ are for $\beta, \gamma \leq p$, namely the transpose of b, as it happens.

It is also true that $a^{\alpha\beta} = A^{ij}(\partial y^\alpha/\partial x^i)\partial y^\beta/\partial x^j$. If we take $\alpha > p$ then $\partial y^\alpha/\partial x^i = \partial x^\alpha/\partial x^i = 0$ whenever $i \leq p$ while $A^{ij} = 0$ if $i > p$. Thus $a^{\alpha\beta} = 0$ for α or $\beta > p$. Thus 11.2 is essentially proved. To obtain the formula exactly, change the signs of all the coordinates $y^1, \ldots, y^n$ and call them $x^1, \ldots, x^n$.

BIBLIOGRAPHY

1. R. Abraham, *Foundations of mechanics*, Benjamin, New York, 1967.

2. R. Arens, *Differential-geometric elements of analytic dynamics*, J. Math. Anal. Appl. **9** (1964), 165–202.

3. R. Arens and D. Babbitt, *Algebraic difficulties of preserving dynamical relations when forming quantum-mechanical operators*, J. Math. and Phys. **6** (1965) 1071–1075.

4. R. Arens, *Invariant sublogics as a way from scalar, to many component, wave equations*, J. Math. Mech. **15** (1966), 344–372.

5. ———, *Laws like Newton's*, J. Math. and Phys. **7** (1966), 1341–1348.

6. ———, *Quantum-mechanization of dynamical systems*, J. Math. Anal. Appl. **19** (1967), 339–385.

7. R. Arens and D. G. Babbitt, *The geometry of relativistic n particle interactions*, Pacific J. Math. **28** (1969), 243–274.

8. D. G. Currie, *A new class of invariance transformations in classical Hamiltonian particle dynamics*, Ann. Physics **36** (1966), 104–129.

9. L. Hörmander, *Pseudo-differential operators*, Comm. Pure Appl. Math. **18** (1965), 501–517.

10. H. Leutwyler, *A no-interaction theorem in classical relativistic Hamiltonian particle mechanics*, Nuovo Cimento **37** (1965), 556–567.

11. G. W. Mackey, *Mathematical foundations of quantum mechanics*, Benjamin, New York, 1963.

12. L. Pontrjagin, *Topological groups*, Princeton Univ. Press, Princeton, N. J., 1939.

13. S. S. Schweber, *Introduction to relativistic quantum field theory*, Row-Peterson, Evanston, Ill., 1961.

14. S. Sternberg, *Lectures on differential geometry*, Prentice-Hall, Englewood Cliffs, N. J., 1964.

15. R. F. Streater and A. S. Wightman, *PCT, spin & statistics,* . . ., Benjamin, New York, 1964.

16. V. S. Varadarajan, *Geometry of quantum theory*, Indian Statistical Institute, Calcutta, 1965.

UNIVERSITY OF CALIFORNIA,
     LOS ANGELES, CALIFORNIA