International Journal of Theoretical Physics, Vol. 32, No. 10, 1993

Axiomatics of Particle Interactions

George W. Mackey[1]

Received April 30, 1993

1. INTRODUCTION

The main point of this presentation is to explain a connection between the now celebrated work of Yang and Mills (1954) on gauge fields and some work of my own which goes back to the same period. The work of Yang and Mills was published in the *Physical Review* in 1954 and the main results were independently found by Shaw and included in his 1954 Ph.D. thesis written under the supervision of A. Salam. My work was done in the early 1950s, but was first published in 1963 in a long survey article on group representations. This article was based on lectures given in 1961 and appeared in the *Bulletin of the American Mathematical Society* (Mackey, 1963*b*). My work did *not* overlap with that of Yang, Mills, and Shaw and was done in a completely different spirit. The connection between the two will be easier to explain later in this paper. For the moment it will suffice to say that the so-called "Yang–Mills trick" emerges naturally from a development of my old work on the axiomatics of quantum mechanical free particles and in such a way as to make at least one *ad hoc* definition appear naturally.

2. TWO OLD RESULTS ON THE AXIOMATIZATION OF QUANTUM MECHANICS

Between 35 and 40 years ago I made two discoveries which in quite different ways helped me to understand why the quantum mechanics of particles takes the form that it does. One of these was a partial answer to the question: Why do we have the von Neumann formalism in which observables are associated with self-adjoint operators in a Hilbert space

[1]Mathematics Department, Harvard University, Cambridge, Massachusetts 02138.

0020-7748/93/1000-1643$07.00/0

and pure states with unit vectors in the same Hilbert space? The other result was a partial answer to a more advanced question: Given the von Neumann formulation and given a system of particles, why is the Hilbert space chosen as it is and why are the self-adjoint operators in this Hilbert space which correspond to the coordinate momentum and energy observables chosen as they are? In particular, why do the famous Heisenberg commutations relations hold? In the particular case of a single free particle, the answer I was able to give to this second question was gratifyingly complete. Moreover, it was almost a corollary of a theorem about unitary group representations which I discovered over 40 years ago in 1949 and published in outline form the same year. This theorem, known as the imprimitivity theorem, plays a major role in the purely mathematical theory of unitary group representations in at least two ways. On one hand, it may be regarded as an abstract characterization of those unitary representations which may be obtained from unitary representations of closed subgroups by the process of inducing. On the other hand, it has as a corollary a general procedure for constructing *all* of the irreducible unitary representations of certain large classes of groups.

Of my two results, the first is by far the better known to this audience (Mackey, 1957, 1963*a*) and I will say little about it here except to recall that it depends in part on looking at the spectral theorem of Hilbert, Stone, and von Neumann "backward". We obtained the association between real-valued observables and self-adjoint operators by first obtaining an association between such observables and projection-valued measures on the real line and then using the spectral theorem to change this into an association between real observables and self-adjoint operators. Non-real-valued observables do *not* correspond to self-adjoint operators. In particular, angle-valued observables correspond to unitary operators.

Because of its relative unfamiliarity and its direct relevance to what I want to say here, I will now describe in outline my result on the axiomatization of the quantum mechanics of free particles.

First recall the standard description of the nonrelativistic quantum mechanics of a free spinless particle. In the model one finds in most elementary textbooks on quantum mechanics, the Hilbert space is the vector space $\mathscr{L}^2(x, y, z)$ of all complex-valued functions $x, y, z \to \psi(x, y, z)$ of the three real variables x, y, and z which are measurable and such that $\int\int\int |\psi(x, y, z)|^2 \, dx \, dy \, dz < \infty$. The self-adjoint operators X, Y, and Z associated with the x, y, and z coordinates of the particles are the multiplication operators $\psi(x, y, z) \to x\psi(x, y, z)$, $\psi(x, y, z) \to y\psi(x, y, z)$, and $\psi(x, y, z) \to z\psi(x, y, z)$, respectively. The self-adjoint operators M_x, M_y, and M_z associated with the x, y, and z components of the linear momenta of the particle are the differentiation operators

$$\psi(x, y, z) \to \frac{\hbar}{i} \frac{\partial \psi}{\partial x}(x, y, z)$$

$$\psi(x, y, z) \to \frac{\hbar}{i} \frac{\partial \psi}{\partial y}(x, y, z)$$

$$\psi(x, y, z) \to \frac{\hbar}{i} \frac{\partial \psi}{\partial z}(x, y, z)$$

respectively.

We shall now reformulate this model in a manner which is independent of our choice of coordinate system, more generalizable, more precise, and better adapted to axiomatic treatment. However, its physical content will be completely unchanged. Specifically, we shall replace the operators X, Y, and Z by a single projection-valued measure P defined not on the real line, but on a model for physical space. Moreover, we shall replace the operators M_x, M_y, and M_z by a unitary representation U of the group $\mathscr{E}$ generated by the translations and rotations of space. These new objects P and U will be defined in terms of X, Y, Z, M_x, M_y, and M_z, and conversely, given P and U and given a rectangular coordinate system in space, we may recontruct X, Y, Z, M_x, M_y, and M_z.

To obtain P from X, Y, and Z, one introduces first the three projection-valued measures on the real line P^x, P^y, and P^z associated with X, Y, and Z by the spectral theorem. These may be written down explicitly. If E is a measurable subset of the real line, then P_E^x is the projection operator $\psi \to \varphi_E^x \psi$, where $\varphi_E^x(x, y, z)$ is one or zero according as $x \in E$ or $x \notin E$. The measures P_E^y and P_E^z are defined analogously. These three projection-valued measures on the real line may be combined by defining $P_{E_1 \times E_2 \times E_3}$ to be $P_{E_1}^x P_{E_2}^y P_{E_3}^z$, for all product sets $E_1 \times E_2 \times E_3$ where $E_1 \times E_2 \times E_3$ means the set of all x, y, and z with $x \in E_1$, $y \in E_2$, and $z \in E_3$. One can prove the existence of a unique projection-valued measure $E \to P_E$ defined on all measurable subsets E of the set of triples of real numbers such that $P_E = P_{E_1}^x P_{E_2}^y P_{E_3}^z$ whenever E is of the form $E_1 \times E_2 \times E_3$. Moreover, P can be defined directly by the formula $P_E(\psi) = \psi'$, where $\psi'(x, y, z) = \psi(x, y, z)$ when $x, y, z \in E$ and $\psi' = 0$ when $x, y, z \notin E$. Conversely, given P, one can define P^x, P^y, and P^z by applying P to sets of the form $E_1 \times E_2 \times E_3$ where two of the E_j are the whole real line.

It follows from the above that giving the three operators X, Y, and Z is the same as giving the projection-valued measure P. But with respect to any choice of coordinates the points of physical space correspond one to one to the triples x, y, z of real numbers and using this correspondence, one can make P into a projection-valued measure on a model for physical space. Thus, *giving the operators X, Y, and Z corresponding to the* *coordinate observables is equivalent to giving a certain projection-valued measure defined on the subsets of physical space.*

Before going on to the momentum observables M_x, M_y,M_z, we make two observations about P.

(a) Given P, we can at once find the self-adjoint operator corresponding to $f(x, y, z)$, where f is any real-valued measurable function of three real variables. It is the unique self-adjoint operator whose associated projection-valued measure is $E \to P_{f^{-1}(E)}$.

(b) Each operator P_E has a direct physical interpretation. It is the operator corresponding to the observable that is one when the particle is observed to be in E and zero when it is observed to be in the complement of E.

Our invariant description of the momentum observables is of a rather different character. Let $\mathscr{E}$ denote the group of all transformations of the set of all triples x, y, z of real numbers generated by those of the form $x, y, z \to x + a, y + b, z + c$ (translations) and those of the form $x, y, z \to b_{11}x + b_{12}y + b_{13}z,\ b_{21}x + b_{22}y + b_{23}z,\ b_{31}x + b_{32}y + b_{33}z$, where the matrix

$$\begin{pmatrix} b_{11} & b_{12} & b_{13} \\ b_{11} & b_{12} & b_{13} \\ b_{11} & b_{12} & b_{13} \end{pmatrix}$$

is a real orthogonal matrix of determinant one (rotations about 0, 0, 0). Because of the mapping of physical space on triples x, y, z defined by a rectangular coordinate system, one can identify $\mathscr{E}$ with the group of transformations in physical space generated by the translations and rotations. Now, if α is any member of $\mathscr{E}$, the linear transformation $f \to f(\alpha^{-1}(x, y, z)) = U_\alpha$ is a unitary transformation of $\mathscr{L}^2(x, y, z)$ onto itself since one checks easily that $U_{\alpha\beta} = U_\alpha U_\beta$ and that U is a so-called unitary representation of the group $\mathscr{E}$ whose Hilbert space is the Hilbert space $\mathscr{L}^2(x, y, z)$ of our one-particle system. Now, given this unitary representation U of $\mathscr{E}$, we may immediately recover the operators M_x, M_y, and M_z as follows. Consider the restriction of U to the subgroup of all translations in the x direction. This is the subgroup of all transformations of the form $x, y, z \to x + a, y, z$ and is isomorphic to the additive group of the real line. Now by a celebrated theorem proved by M. H. Stone in 1930, for every (strongly continuous) unitary representation V of the additive group of the real line there is a unique self-adjoint operator A such that $V_a = e^{iaA}$ for all real a. Conversely, for every self-adjoint operator A, $V_a = e^{iaA}$ is well defined for all a and $a \to V_a$ is a strongly continuous unitary representation of the additive group of the real line. iA is called the *infinitesimal generator* of V because there is a sense in which $(d/da)V_a = iA$ when $a = 0$. One can show that the infinitesimal generator of the restriction of U to the subgroup of translations in the x direction is d/dx, so that M_x is just $\hbar/i$ times this infinitesimal generator. The obvious analogs for M_y and M_z also hold. Conversely, knowing M_x, M_y, and M_z, one can immediately write down U_α for any member α of the translation subgroup of $\mathscr{E}$.

Summing up, we have shown that giving X, Y, Z and M_x, M_y, M_z is the same as giving a projection-valued measure on physical space and a unitary representation of the group of spatial translations. This replacement has three advantages.

(a) It is coordinate free.

(b) Since all operators concerned are bounded, it makes it possible to avoid the tricky domain questions involved in the rigorous definition of unbounded self-adjoint operators.

(c) It sets the stage for the axiomatic definition of a particle which we propose to present below.

The projection-valued measure P and the unitary representation U satisfy a simple relationship which we may readily verify by direct computation. It reads

$$U_\alpha P_E U_{\alpha^{-1}} = P_{(E)\alpha^{-1}} \tag{1}$$

where $(E)\alpha$ is the transformation of the set E by the isometry α. When restricted to the spatial translations it is equivalent to the Heisenberg commutation relations between X, Y, Z, M_x, M_y, and M_z.

We now start anew and assume only the von Neumann formulation in the abstract. We do not make any assumption about the Hilbert space other than separability and we do not make *any* assumption about the special form of the position and momentum operators. In fact, in line with the above, we do not seek coordinate operators at all, but instead seek the self-adjoint operators P_E which for each measurable subset E of space correspond to the observable that is one when the particle is in E and zero when it is not. The P_E are necessarily projection operators (since these are the only self-adjoint operators taking on only zero and one) and we make the physically plausible assumption that they commute with one another and satisfy the conditions defining a projection-valued measure:

$$P_E P_F = P_{E\cap F}, \qquad P_{E_1 \cup E_a \cup \cdots} = P_{E_1} + P_{E_2} + \cdots \tag{2}$$

when the E_j are *disjoint* sets, $P_S = I$, $P_\varnothing = 0$, where $\varnothing$ is the empty set. Instead of seeking momentum operators, we invoke symmetry considerations and assume that the laws of physics are invariant under spatial translations and rotations; that is, under the natural action of $\mathscr{E}$ on physical space. Given that every symmetry of the Hilbert space of states is implemented by either a unitary or an antiunitary operator, that the square of every antiunitary operator is unitary, and that every element of $\mathscr{E}$ is the square of another element, we can almost conclude that there must exist a unitary representation U of $\mathscr{E}$ in the Hilbert space which implements the symmetry. However, since two unitary operators define the *same* automorphism of the states whenever one is a constant of absolute value one tunes the other, we can only conclude that $U_{xy} = U_x U_y \sigma(x, y)$, where $\sigma(x, y)$ is some function of x and y of absolute value one. This means that U is a *projective* or *ray representation* of $\mathscr{E}$, which may or may not be an ordinary unitary representation.

In sum, then, we assume that given a free particle, there must be some projection-valued measure $E \to P_E$ defined on the measurable subsets of space (from which all coordinate observables can then be derived) and there must be some (possibly projective) unitary representation U of $\mathscr{E}$ in the same Hilbert space which implements the symmetry of physics under translation and rotation. At this point, P could be an arbitrary projection-valued measure defined on the measurable subsets of space and U could be an arbitrary (possibly projective) unitary representation of $\mathscr{E}$ and the possibilities for the pairs U, P would be enormous. However, we have not taken account of the fact that U implements the fact that physical laws are independent of orientation and position in space and that $\mathscr{E}$ acts on P through its action on space. It does not take much reflection to reach the conclusion that these facts demand that U and P must be related as follows:

$$U_\alpha P_E U_{\alpha^{-1}} = P_{[E]\alpha^{-1}} \qquad (*)$$

for all α in $\mathscr{E}$. This is just (1) above, which, as already indicated, can be derived by calculation in the classical model: Now, however, it appears, not as a property of a particular model, but as a consequence of physically plausible general principles.

The interest of this relationship for us now is that it puts extremely strong constraints on the pairs P, U which satisfy it. There is a very general theorem in the theory of the unitary representations of locally compact groups which tells us that (to within unitary equivalence) there is precisely one solution of (*) for each (possibly projective) unitary representation of the subgroup of $\mathscr{E}$ consisting of all α which take some fixed origin of space into itself; that is, the group of all proper rotations $SO(3)$. These representations were among the first to be classified when I. Schur and H. Weyl extended the representation theory of finite groups to compact Lie groups. Thus, applying this theorem (called the imprimitivity theorem) allows us to find *all* concrete models satisfying our axioms for a particle. It turns out that $SO(3)$ has essentially only one projective multiplier σ other than the one which is identically equal to 1 and that with this multiplier it has just one irreducible unitary projective representation of every even dimension. One denotes the one of dimension $2l+1$ by D^l, $l = 1/2, 3/2, 5/2$. Also, it has an ordinary irreducible unitary representation of every odd dimension and denotes the one of dimension $2l+1$ by D^l, where $l = 0, 1, 2, 3$. The simplest case is that in which $l = 0$, so that D^l is the trivial representation. The corresponding concrete model satisfying our axioms is the model for the free spinless particle with which we started. The next simplest case is that in which $l = 1/2$ and the corresponding model is the standard model for a free particle of "spin 1/2." Other irreducible cases correspond to the standard model for particles of higher spin. We emphasize the natural way in which spin comes into the picture without any relativistic assumptions or any need to explain spectroscopic data. There is a great deal more to be said, but it would throw this paper out of balance to do so. For example, one can discuss the case in which the underlying representation of $SO(3)$ is *not* irreducible and so make connection with particle multiplets, one can discuss dynamics and be led naturally to the Dirac equations in the spin-1/2 case, etc. For further details we refer the reader to Mackey (1968) and Chapter 18 of Mackey (1978).

3. THE AXIOMATIZATION OF PARTICLE INTERACTIONS

There is no problem in passing from the axiomatization of a single free particle to that of a finite number of such particles—provided that the particles move independently and do not interact with one another. However, particles which do not interact are not very interesting in physics and the essential next step is to study particle interactions from the same axiomatic point of view as we studied free particles above. I did some work in this direction in the fall of 1965 and included it in a course of lectures given at Harvard at that time and again in a course of lectures given at Oxford in the academic year 1966–1967. These lectures were written up, typed, and mimeographed and in 1978, together with 9000 words of "notes and references," were published in book form (Mackey, 1978). The relevant part is Chapter 19.

In this work I considered only the case in which the particles are all distinct and the system is Galilean invariant. I made the plausible assumption that the operators defining the velocity observables of any one particle commute with the position observables of all other particles and that the operators defining the position observables are the same as they are when the particles do not interact. The operators defining the velocity components of the jth particle are the commutators $i(HX_j - X_jH)$, $i(HY_j - Y_jH)$, and $i(HZ_j - Z_jH)$, where H is the self-adjoint operator such that $t \to e^{-iHt}$ is the one-parameter unitary group defining the time evolution of the system and X_j, Y_j, and Z_j are the operators defining the position observables. The reason for this relationship between velocity operators and the "dynamical operator" H is explained in Chapter 18 of Mackey (1978). On the other hand, the extremely natural and plausible assumption that a system of n distinct noninteracting particles is described by just taking the tensor products of the representations defining the individual particles combined with the results already described about the quantum mechanisms of single particles gives us the X_j, Y_j, and Z_j.

From the above considerations and assumptions I was able to show that for a Galilean invariant system of n distinct particles in quantum mechanics, the underlying Hilbert space may be chosen to be the space of all square summable functions from Euclidean $3n$ space to a certain auxiliary Hilbert space $\mathscr{H}_0$ in such a manner that the X_j, Y_j, and Z_j are the usual multiplicative operators $f \to x_j f$, $f \to y_j f$, $f \to z_j f$ and when this is done the dynamical operator H has the following form:

$$H = \frac{1}{2} \sum_{j=1}^{n} \frac{1}{\mu_j} \left(\frac{1}{i}\frac{\partial}{\partial x_j} + A_j'\right)^2 + \left(\frac{1}{i}\frac{\partial}{\partial y_j} + B_j'\right)^2 + \left(\frac{1}{i}\frac{\partial}{\partial z_j} + C_j'\right)^2 + V' \quad (2)$$

where A_j', B_j', C_j', V' are measurable functions from Euclidean 3-space to the self-adjoint operators in the Hilbert space $\mathscr{H}_0$ and the μ_j are constant multiples of the particle masses. The $3n + 1$ functions A_j', B_j', C_j', V' determine the nature of the interaction and are all identically zero for noninteracting particles. One computes that the operators $i(HX_j - X_jH)$, $i(Hy_j - Y_jH)$, and $i(HZ_j - Z_jH)$ defining the velocity observables are

$$\frac{1}{\mu_j}\left(\frac{1}{i}\frac{\partial}{\partial x_j} + A_j'\right), \qquad \frac{1}{\mu_j}\left(\frac{1}{i}\frac{\partial}{\partial y_j} + B_j'\right), \qquad \frac{1}{\mu_j}\left(\frac{1}{i}\frac{\partial}{\partial z_j} + C_j'\right)$$

Thus the assumption that the velocity operators are the same as in the noninteracting case implies that $A_j' = B_j' = C_j' = 0$ for all j and that the interaction is completely described by the single function V'. If one looks at the special case in which all particles are spinless, then V' is a real-valued function and one computes that in the classical limit one has precisely the classical theory of n particles moving under the influence of the potential function V'. It is noteworthy that we obtain this result without starting from a classical model. This suggests that certain apparently arbitrary features of classical mechanics are "quantum effects" in the sense that they are consequences of the fact that classical mechanics must be a limiting form of quantum mechanics. See Chapter 19 of Mackey (1978) for further details.

The consequences of (2) were explored further in the case of general A'_j, B'_j, C'_j but spin zero for all particles and in the case $n = 2$ with general spin for both particles but with $A'_j, B'_j, C'_j \equiv 0$. In the one case the classical limit exhibited "velocity-dependent forces" of the sort one finds when a particle moves in a magnetic field, and in the other case one is led to interactions of the sort found by "guessing" in early work on nuclear physics. Neither case was explored in depth.

I returned to the subject 22 years later in a talk given in the summer of 1987 at a conference in Como and entitled "Weyl's program and modern physics." Section VII of my paper in the *Proceedings* of this conference (Mackey, 1988) is entitled "Particle interactions and the gauge principle." In it I review some of the above and remark on the probability of a connection with the work of Yang and Mills. The last paragraph of Section VII begins as follows: "The author has not yet done the detailed calculations required to check this but would be very surprised if pursuing this line of thought did not lead to something very close to, and possibly more general than, the modern theory of non Abelian gauge fields which originated in the celebrated work of Yang and Mills. If so, it is interesting that one could have been led to it by systematically seeking to extend Weyl's program in the manner indicated in this section."

The main result of the present paper is that detailed calculations have now been carried out which confirm the conjecture just described. The author confesses that at the time the text of Mackey (1978) was written he was not aware of the Yang–Mills paper and the subsequent development of gauge theories by Gell-Mann, Salam, Ward, Glashow, Bludman, and others. If he had been, he might have explored much earlier the case in which the A_j, B_j, and C_j are not zero *and* the particles are not spinless. In his ignorance at the time, this case seemed to be of purely mathematical interest.

In all of this work, as well as in that which follows, we have worked purely formally, ignoring the fact that our self-adjoint operators are mainly unbounded and hence not everywhere defined. In particular, we have written

$$\left[\frac{1}{i}\frac{\partial}{\partial x}, \tilde{A}\right] = \frac{1}{i}\widetilde{\frac{\partial A}{\partial x}}$$

when A is not known to be differentiable anywhere. This indicates that the commutator of two densely defined self-adjoint operators may be nowhere defined. Evidently, in order to make sense of our argument, H must be subjected to regularity restrictions in addition to being required to satisfy the indicated commutation relations. Just what these are remains to be worked out. In this sense our work on the axiomatics of particle interactions is less complete than our work on the axiomatics of free particles. Nevertheless, it seems to us to be interesting and worth recording.

4. CONNECTIONS WITH GAUGE FIELDS

Let us simplify the exposition, as was done in Section VII of Mackey (1988), by restricting ourselves to two particles and replacing the problem of two interacting particles by the equivalent one of a single particle moving in a "force field." This will have the advantage of making the comparison with the Yang–Mills paper more direct and immediate.

In other words, let us consider a Hilbert space $\mathscr{H}$ of the form $\mathscr{L}^2(x, y, z, \mathscr{H}_0)$, where $\mathscr{H}_0$ is some other Hilbert space, and let X, Y, Z denote the "multiplication operators" $f \to xf$, $f \to yf$, $f \to zf$, where f is an element of $\mathscr{H}$. These are then the position operators for a single particle defining the motion of a pair of particles with respect to their center of gravity. The motion itself is described by a one-parameter unitary group $\mathscr{A} \to V_t = e^{-iHt}$, where H is some self-adjoint operator depending on the interaction between the particles. The axioms of Section 3 then translate into the following axioms restricting H. Let $X' = i(HX - XH)$, $Y' = i(HY - YH)$, and $Z' = i(HZ - ZH)$. Then

$$XX' - X'X = YY' - Y'Y = ZZ' - Z'Z = \frac{i}{\mu} I \tag{3}$$

where μ is some positive real number, I is the identity, and

$$\begin{aligned} XY' - Y'X &= YZ' - Z'Y = ZX' - X'Z = XZ' - Z'X \\ &= YX' - X'Y = ZY' - Y'Z = 0 \end{aligned} \tag{4}$$

The simplest H such that X, Y, Z, X', Y', and Z' satisfy (3) and (4) is

$$H_0 = -\frac{1}{2\mu}\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}\right)$$

which corresponds to the case of two noninteracting particles or of a single particle moving freely. μ is proportional to the (effective) mass of the particle. The problem of finding the most general interaction between the particles subject to the above axioms reduces to finding the most general H satisfying (3) and (4).

The problem of finding the most general such H is quite easy and is solved by an obvious adaptation of the arguments of Chapter 19 of Mackey (1978). We observe first that replacing X' by $(1/\mu)\,\partial/\partial x$, Y' by $(1/\mu)\,\partial/\partial y$, and Z' by $(1/\mu)\,\partial/\partial z$ leaves (3) and (4) unaltered. Hence

$$X' - \frac{1}{i\mu}\frac{\partial}{\partial x}, \qquad Y' - \frac{1}{i\mu}\frac{\partial}{\partial y}, \qquad \text{and} \qquad Z' - \frac{1}{i\mu}\frac{\partial}{\partial z}$$

all commute with X, Y, and Z. Now it is a well-known theorem in real variable theory that the bounded linear operators in $\mathscr{L}^2(x, y, z, \mathscr{H}_0)$ which commute with X, Y, and Z are precisely the operators $f \rightarrow f'$, where $f'(x, y, z) = G(x, y, z)(f(x, y, z))$ and where G is some bounded measurable function of x, y, and z having values which are bounded linear operators in $\mathscr{H}_0$. Here $G(x, y, z)(f(x, y, z))$ denotes the result of acting on the vector $f(x, y, z)$ in $\mathscr{H}_0$ with the operator $G(x, y, z)$. It follows easily that the (not necessarily bounded) self-adjoint operators which commute with X, Y, and Z are of the same form but with G an arbitrary measurable function whose values are *self-adjoint* operators in $\mathscr{H}_0$. We conclude that there must exist measurable self-adjoint operator-valued functions A, B, and C of x, y, and z such that

$$X' = \frac{1}{\mu}\left(\frac{1}{i}\frac{\partial}{\partial x} + \tilde{A}\right) \qquad Y' = \frac{1}{\mu}\left(\frac{1}{i}\frac{\partial}{\partial y} + \tilde{B}\right), \qquad Z' = \frac{1}{\mu}\left(\frac{1}{i}\frac{\partial}{\partial z} + \tilde{C}\right)$$

where $\tilde{A}$, $\tilde{B}$, and $\tilde{C}$ are the operators defined by the functions A, B, and C as indicated above.

Now let H' be any self-adjoint operator for which

$$i(H'X - XH') = X', \qquad i(H'Y - YH') = Y', \qquad i(H'Z - ZH') = Z'$$

Then $H - H'$ must commute with X, Y, and Z and so, by the argument given above, must be $\tilde{v}$ for a fourth measurable function v of x, y, z and having self-adjoint operators in $\mathscr{H}_0$ as values. Hence our problem will be solved if we can find a single self-adjoint operator H' which satisfies the three equations at the beginning of the paragraph. But a straightforward calculation shows that the operator

$$\frac{1}{2\mu}\left[\left(\frac{1}{i}\frac{\partial}{\partial x} + \tilde{A}\right)^2 + \left(\frac{1}{i}\frac{\partial}{\partial y} + \tilde{B}\right)^2 + \left(\frac{1}{i}\frac{\partial}{\partial z} + \tilde{C}\right)^2\right]$$

has the required properties. In short, the most general self-adjoint operator satisfying our axioms (3) and (4) is

$$H = \frac{1}{2\mu}\left[\left(\frac{1}{i}\frac{\partial}{\partial x} + \tilde{A}\right)^2 + \left(\frac{1}{i}\frac{\partial}{\partial y} + \tilde{B}\right)^2 + \left(\frac{1}{i}\frac{\partial}{\partial z} + \tilde{C}\right)^2\right] + \tilde{v} \qquad (5)$$

where A, B, C, and v are measurable functions of x, y, z taking values in the space of all self-adjoint operators in $\mathscr{H}_0$.

Our next remark is that the physics defined by the operator (5) depends only on the unitary equivalence class of the operator quadruple H, X, Y, Z. In particular, H_1 and H_2 defined by (5) with A, B, C, v replaced by A_1, B_1, C_1, v_1 and A_2, B_2, C_2, v_2, respectively, will define the same physics whenever there exists a unitary operator U which commutes with X, Y, and Z and is such that $UH_1U^{-1} = H_2$. We now seek to express the condition that such a U exists directly on the functions A_1, B_2, C_1, v_1 and A_2, B_2, C_2, v_2. From earlier arguments we deduce at once that the most general U which commutes with X, Y, and Z is $\tilde{u}$, where u is a measurable function of x, y, and z whose values are unitary operators in $\mathscr{H}_0$. Given such a u, let us compute UH_1U^{-1}. This reduces at once to computing $U(\partial/\partial w)U^{-1}$ for $w = x$, y and z and $U\tilde{G}U^{-1}$ for $G = A_1, B_1, C_1, v_1$. Evidently, $U\tilde{G}U^{-1} = \widetilde{uGU^{-1}}$ and it is easy to show that

$$U\tilde{G}U^{-1} = \frac{\partial}{\partial w} - \frac{\partial u}{\partial W} u^{-1}$$

Using these facts, it is straightforward to compute that

$$A_2 = uA_1u^{-1} - \frac{1}{i}\frac{\partial u}{\partial x}u^{-1}$$

$$B_2 = uB_1u^{-1} - \frac{1}{i}\frac{\partial u}{\partial y}u^{-1}$$

$$C_2 = uC_1u^{-1} - \frac{1}{i}\frac{\partial u}{\partial z}u^{-1}$$

$$v_2 = uv_1u^{-1}$$

The special case in which $\mathscr{H}_0$ is one-dimensional is the easiest to analyze further, for in that case the operators of the form $\tilde{G}$ all commute with one another and are multiplication operators by real-valued functions. In particular, the relations between A_1, B_1, C_1, v_1 and A_2, B_2, C_2, v_2 when U exists become much simpler, reducing to

$$A_2 = A_1 - \frac{1}{i}\frac{1}{u}\frac{\partial u}{\partial x}$$

$$B_2 = B_1 - \frac{1}{i}\frac{1}{u}\frac{\partial u}{\partial y}$$

$$C_2 = C_1 - \frac{1}{i}\frac{1}{u}\frac{\partial u}{\partial z}$$

$$v_2 = v_1$$

Writing $u = e^{i\lambda}$, where λ is a real-valued function of x, y, and z, we get

$$A_2 = A_1 - \frac{\partial \lambda}{\partial x}$$

$$B_2 = B_1 - \frac{\partial \lambda}{\partial y}$$

$$C_2 = C_1 - \frac{\partial \lambda}{\partial z}$$

$$v_2 = v_1$$

In short, the scalar v is uniquely determined, but the vector A, B, C is only determined up to the addition of the gradient of a scalar λ.

Before going on to the general case, where the situation is more complicated, let us look briefly at the classical limit of the case at hand, where there is an interesting connection with electromagnetism. The technique for passing to the classical limit is discussed (in the case of spinless particles) in Chapter 19 (Mackey, 1978). It is based on computing the operators associated with "acceleration observables"—these being defined as the commutators $i(HX' - X'H)$, $i(HY' - Y'H)$, and $i(HZ' - Z'H)$, where X', Y', and Z' are the operators associated with the velocity observables. Equivalently these operators are the double commutators

$$-H(HX - XH) + (HX - XH)H$$

$$-H(HY - YH) + (HY - YH)H$$

$$-H(HZ - ZH) + (HZ - ZH)H$$

and are of course invariants of the system consisting of H and the three position operators X, Y, and Z. When H is as above and $\mathscr{H}_0$ is one-dimensional, these double commutators can be computed quite easily and turn out to be the three operators

$$-\frac{1}{\mu}\left(\widetilde{\frac{\partial v}{\partial x}}\right) + \frac{1}{2\mu}\left\{Y', \left(\widetilde{\frac{\partial A}{\partial y} - \frac{\partial B}{\partial x}}\right)\right\} + \frac{1}{2\mu}\left\{Z', \left(\widetilde{\frac{\partial A}{\partial z} - \frac{\partial C}{\partial x}}\right)\right\} \tag{6}$$

$$-\frac{1}{\mu}\left(\widetilde{\frac{\partial v}{\partial y}}\right) + \frac{1}{2\mu}\left\{Z', \left(\widetilde{\frac{\partial B}{\partial z} - \frac{\partial C}{\partial y}}\right)\right\} + \frac{1}{2\mu}\left\{X', \left(\widetilde{\frac{\partial B}{\partial x} - \frac{\partial A}{\partial y}}\right)\right\} \tag{7}$$

$$-\frac{1}{\mu}\left(\widetilde{\frac{\partial v}{\partial z}}\right) + \frac{1}{2\mu}\left\{X', \left(\widetilde{\frac{\partial C}{\partial x} - \frac{\partial A}{\partial z}}\right)\right\} + \frac{1}{2\mu}\left\{Y', \left(\widetilde{\frac{\partial C}{\partial y} - \frac{\partial B}{\partial z}}\right)\right\} \tag{8}$$

where $\{\cdot\}$ denotes the *anti*commutator of the two operator arguments. From this, as in Chapter 19 of Mackey (1978), one deduces that in the classical limit the particle satisfies the ordinary differential equations

$$\mu \frac{d^2x}{dt^2} = -\frac{\partial v}{\partial x} + \frac{dy}{dt}\left(\frac{\partial A}{\partial y} - \frac{\partial B}{\partial x}\right) - \frac{dz}{dt}\left(\frac{\partial C}{\partial x} - \frac{\partial A}{\partial z}\right)$$

$$\mu \frac{d^2y}{dt^2} = -\frac{\partial v}{\partial y} + \frac{dz}{dt}\left(\frac{\partial B}{\partial z} - \frac{\partial C}{\partial y}\right) - \frac{dx}{dt}\left(\frac{\partial A}{\partial y} - \frac{\partial B}{\partial x}\right)$$

$$\mu \frac{d^2z}{dt^2} = -\frac{\partial v}{\partial z} + \frac{dx}{dt}\left(\frac{\partial C}{\partial x} - \frac{\partial A}{\partial z}\right) - \frac{dy}{dt}\left(\frac{\partial B}{\partial z} - \frac{\partial C}{\partial y}\right)$$

But these are just the equations of motion of a particle of charge e and mass μ/e in an electric field, $-\operatorname{grad} v$, and a magnetic field, $\operatorname{curl}(A, B, C)$. Once again we see that physical effects depend only on $\operatorname{curl}(A, B, C)$ and not on A, B, C itself.

It is interesting that we are able to derive the equations of motion of a particle in an electromagnetic field from general principles of quantum mechanics as applied to spinless particles.

We pause to mention a connection between the derivation just given and a private informal lecture given by the late R. P. Feynman to Freeman Dyson in 1948. This lecture was briefly summarized in the lower left-hand portion of p. 37 of Dyson's retrospective article about Feynman in *Physics Today* (Dyson, 1989). On reading Dyson's account of this (by then) 40-year-old unpublished lecture, the author was immediately reminded of Section VII of Mackey (1988) and sent Dyson a copy with a request for comments and further particulars. Dyson obliged by sending a copy of his own (reconstructed) notes on Feynman's lecture and soon thereafter published these notes in the *American Journal of Physics* (Dyson, 1990).

Feynman's analysis is like ours in two respects: (1) He, too, makes the assumption that $x_j\dot{x}_i - \dot{x}_i x_j$ is a constant times δ_i^j and of course that the x_j commute. (2) He deduces some of the features of electromagnetism. He goes further than we did in that he considered the time-dependent case. When his fields are time independent the equations he deduces, namely $\operatorname{div} H = 0$ and $\partial H/\partial t + \operatorname{curl} E = 0$, become $\operatorname{div} H = 0$ and $\operatorname{curl} E = 0$, which of course are implied by the equations $E = -\operatorname{grad} V$ and $H = \operatorname{curl}(A, B, C)$ which we obtain. On the other hand, Feynman's point of view and motivation are quite different. In particular, (according to Dyson) he was not trying to axiomatize particle interactions, but to find an alternative to quantum mechanics.

Now let us return to the general case in which no restriction is placed on the dimension of $\mathscr{H}_0$ so that the operator-valued functions A, B, C, and v do not usually reduce to real-valued functions and the operators $\hat{A}$, $\hat{B}$, $\hat{C}$, $\hat{v}$, and $\hat{u}$ do not necessarily commute with one another. In particular, the passage to the classical limit becomes rather more complicated. As above, it depends upon computing the acceleration operators, but the result of this computation is less simple and the computation is lengthier. These complications can be handled more efficiently if we change our notation and replace the coordinates x, y, and z by x_1, x_2, and x_3, the operator-valued functions A, B, and C by A_1, A_2, and A_3, and the operators X, Y, and Z by X_1, X_2, and X_3.

We wish to compute $[H, [H, X_k]]$ for $k = 1, 2, 3$, where $H = \frac{1}{2}\mu((X'_1)^2 + (X'_2)^2 + (X'_3)^2) + \tilde{v}$ and $X'_j = i[H, X_j]$. We note first that

$$[H, [H, X_k]] = \left[\sum_{j=1}^{3} (X'_j)^2, X'_k\right] + \frac{1}{i}[\tilde{v}, X'_k]$$
$$= \frac{\mu}{2i}\sum_{j=1}^{3} [(X'_j)^2, X'_k] + \frac{1}{i}[\tilde{v}, X'_k]$$

Next we apply the easily proved general theorem that for any two operators F and G the commutator $[F^2, G] = \{F, [F, G]\}$, where $\{\cdot\}$ denotes the *anti*commutator of its two arguments. We deduce that

$$[H, [H, X_k] = \frac{\mu}{2i}\sum_{j=1}^{3} \{X'_j, [X'_j, X'_k]\} + \frac{1}{i}[\tilde{v}, X'_k]$$

Now recall that

$$X'_j = \frac{1}{\mu}\left(\frac{1}{I}\frac{\partial}{\partial x_j} + \tilde{A}_j\right)$$

so that

$$[X'_j, X'_k] = \frac{1}{\mu^2}\left(0 + \left[\frac{1}{i}\frac{\partial}{\partial x_j}, \tilde{A}_k\right] + \left[\tilde{A}_j, \frac{1}{i}\frac{\partial}{\partial x_k}\right] + [\tilde{A}_j, \tilde{A}_k]\right)$$
$$= \frac{1}{\mu^2}\left(\frac{1}{i}\frac{\partial \widetilde{A}_k}{\partial x_j} - \frac{1}{i}\frac{\partial \widetilde{A}_j}{\partial x_k} + [\tilde{A}_j, \tilde{A}_k]\right)$$

At this point it will be convenient to define

$$\frac{1}{i}\frac{\partial A_k}{\partial x_j} - \frac{1}{i}\frac{\partial A_j}{\partial x_k} + [A_j, A_k]$$

as $F_{j,k}$ and notice that $F_{j,k} = -F_{k,j}$ for all j and k and $F_{j,j} = 0$ for all j. Notice that $F_{j,k}$ is an operator-valued function for all j and $k = 1, 2, 3$ and that we may now write

$$[H, [H, X_k]] = \frac{\mu}{2i}\sum_{j=1}^{3}\left\{\frac{1}{\mu}\left(\frac{1}{i}\frac{\partial}{\partial x_j} + \tilde{A}_j\right), \frac{1}{\mu^2}\tilde{F}_{j,k}\right\} + \frac{1}{i}[\tilde{v}, X'_k]$$
$$= \frac{1}{2i\mu^2}\sum_{j\neq k}\left\{\frac{1}{i}\frac{\partial}{\partial x_j}, \tilde{F}_{j,k}\right\} + \frac{1}{2\mu^2 i}\sum_{j\neq k}\{\tilde{A}_j, \tilde{F}_{j,k}\} + \frac{1}{i}[\tilde{v}, X'_k]$$

Now observe that

$$\left\{\frac{1}{i}\frac{\partial}{\partial x_j}, \tilde{F}_{j,k}\right\} = 2\tilde{F}_{j,k}\frac{1}{i}\frac{\partial}{\partial x_j} + \frac{1}{i}\frac{\partial \widetilde{F_{j,k}}}{\partial x_j}$$

and that

$$[\tilde{v}, X'_k] = \left[\tilde{v}, \frac{1}{\mu}\left(\frac{1}{i}\frac{\partial}{\partial x_k} + \tilde{A}_k\right)\right] = \frac{1}{\mu}[\tilde{v}, \tilde{A}_k] + \left[\tilde{v}, \frac{1}{\mu}\frac{1}{i}\frac{\partial}{\partial x_k}\right] = \frac{1}{\mu}[\tilde{v}, \tilde{A}_k] - \frac{1}{\mu_i}\frac{\partial \tilde{v}}{\partial x_k}$$

Substituting these evaluations into the last expression for $[H, [X, X_k]]$, we get finally

$$[H, [H, X_k]] = \frac{1}{\mu^2}\sum_{j \neq k} \tilde{F}_{j,k}\frac{1}{i}\frac{\partial}{\partial x_j} + \beta_k \tag{9}$$

where β_k is the operator-valued function

$$-\frac{1}{2\mu^2}\sum_{j \neq k}\frac{\partial F_{j,k}}{\partial x_j} + \frac{1}{2\mu_i^2}\sum_{j \neq k}\{A_j, F_{j,k}\} + \frac{1}{i\mu}[v, A_k] + \frac{1}{\mu}\frac{\partial v}{\partial x_k} \tag{10}$$

We shall not go on to write down the equations of motion in the classical limit. For our purpose it is sufficient to notice that

$$F_{j,k} = \frac{1}{i}\left(\frac{\partial A_k}{\partial x_j} - \frac{\partial A_j}{\partial x_k}\right) + [A_j, A_k]$$

appears automatically as the coefficients of $(1/i)\,\partial/\partial j$ in the expression of $[H, [H, X_k]]$ as a first-order differential operator and hence is an invariant of the operator quadruple H, X_1, X_2, X_3. Specifically, if U is any unitary operation carrying H, X_1, X_2, X_3 into H', X_1, X_2, X_3, and $F_{j,k}$ and $F'_{j,k}$ are defined from A_1, A_2, A_3 and A'_1, A'_2, A'_3, respectively, then it follows from general principles that U carries $\tilde{F}_{j,k}$ into $\tilde{F}'_{j,k}$, i.e. $\tilde{F}'_{j,k} = U^{-1}\tilde{F}_{j,k}U$.

In comparing the above results with the paper of Yang and Mills (1954) we wish to emphasize the following. Our Hamiltonian H is defined by four operator-valued functions on space A_1, A_2, A_3, and v. The "b field" of Yang and Mills is defined by four 2×2 matrix-valued functions on space-time. The condition that H, X_1, X_2, X_3 and H', X_1, X_2, X_3 be unitarily equivalent is the existence of a unitary operator-valued function on space such that

$$A'_j = uA_ju^{-1} + i\frac{\partial u}{\partial x_j}u^{-1}$$

$$v' = uvu^{-1}$$

for $j = 1, 2, 3$. Equation (3) of the Yang–Mills paper reads

$$B'_\mu = S^{-1}B_\mu S + \frac{i}{\epsilon}S^{-1}\frac{\partial S}{\partial x_\mu}$$

where B_μ and B'_μ are gauge equivalent to "b fields" and S is a 2×2 unitary operator-valued function on space-time.

Equation (4) of the Yang–Mills paper is the definition

$$F_{\mu,v} = \frac{\partial B_\mu}{\partial x_v} - \frac{\partial B_v}{\partial x_\mu} + i\epsilon(B_\mu B_v - B_v B_\mu)$$

and equation (5) is

$$F'_{\mu v} = S^{-1} F_{\mu v} S$$

which is proved by computation from equation (3). The only motivation given for making the definition of equation (4) is "Other simple functions of B than (4) do not lead to such a simple transformation property" [see (5)]. In our theory the expression $\partial A_k/\partial x_j - \partial A_j/\partial x_k + i(A_j A_k - A_k A_j)$ occurs naturally in such a way that its invariance properties follow from general principles.

REFERENCES

Dyson, F. J. (1989). Feynman at Cornell, *Physics Today*, **42**, 32–38.

Dyson, F. J. (1990). Feynman's proof of the Maxwell equations, *American Journal of Physics*, **58**, 209–211.

Mackey, G. W. (1957). Quantum mechanics and Hilbert space, *American Mathematical Monthly*, **64**, 45–57.

Mackey, G. W. (1963*a*). *The Mathematical Foundations of Quantum Mechanics*, Benjamin.

Mackey, G. W. (1963*b*). Infinite-dimensional group representations, *Bulletin of the American Mathematical Society*, **69**, 628–686.

Mackey, G. W. (1968). *Induced Representations of Groups and Quantum Mechanics*, Benjamin and Editore Boringhieri.

Mackey, G. W. (1978). *Unitary Group Representations in Physics, Probability and Number Theory*, Benjamin.

Mackey, G. W. (1988). Weyl's program and modern physics, in *Differential Geometrical Methods in Theoretical Physics*, K. Bleuler and M. Werner, eds., pp. 11–36.

Yang, C. N., and Mills, R. L. (1954). Conservation of isotopic spin and isotopic gauge invariance, *Physical Review*, **96**, 191–195.