Group-theoretical foundations of classical and quantum mechanics. II. Elementary systems[a)]

L. Martínez Alonso

Departamento de Métodos Matemáticos, Universidad Complutense, Madrid-3, Spain
(Received 18 October 1977; revised manuscript received 21 March 1978)

Classical and quantum elementary systems are investigated from the point of view of invariance under a connected Lie group. The classification and characterization of elementary systems are considered in a unified way. The representation theory of symmetric and enveloping algebras is used as a tool in order to characterize the observables physically and also to analyze the analogies between classical and quantum mechanics. The results obtained are applied to the Galilei, Poincaré, and Weyl Lie groups.

1. INTRODUCTION

In a previous article[1] we have explored the algebraic structures associated with Lie algebras and their importance in the analysis of the classical and quantum observables which emerge from a given Lie algebra. It was a first step in order to test the physical relevance of the formulation of invariance principles under general connected Lie groups. This paper is concerned with the group theoretical concept of elementary systems. Since the work of Wigner[2] it has been a key notion in the mathematical development of classical and quantum mechanics. Then, the way Wigner defined quantum elementary systems by means of the projective unitary irreducible representations of the Poincaré group has become familiar for a wide class of physicists. At present Wigner's ideas have been extended successfully to other invariance groups[3] and also to classical mechanics,[4,5] but the diversity of the mathematical tools involved makes difficult the understanding of the relations between the results obtained in the different contexts. For this reason it is important from both the mathematical and physical points of view to discuss the invariance under Lie groups in such a way that the classical and quantum elementary systems may be considered simultaneously. In the present paper we propose a general formalism for the classification and characterization of the elementary systems which allows us to analyze the mathematical and physical analogies between both concepts in a unified manner.

Section 2 begins with a brief survey of the basic concepts which arise in the formulation of invariance principles. This is necessary as the literature about this subject is not very explicit when it concerns the relation between the kinematical and the dynamical actions of invariance groups. Then, the mathematical definitions of elementary systems are introduced in terms of the kinematical realizations of the invariance group $\mathcal{G}_0$. Quantum elementary systems (QES) are identified with the projective unitary irreducible representations of $\mathcal{G}_0$; and classical elementary systems (CES) are defined to be the transitive canonical realizations of $\mathcal{G}_0$. This characterization of CES follows from the Souriau work,[4] and it is the most convenient from the mathematical point of view. There is a different approach to the concept of CES which has been developed by Sudarshan and Mukunda[5] and Pauri and Prosperi.[6] These authors characterize CES in terms of the representations of Lie algebras by functions in phase space. Their characterization is, in general, only local, but it has a more immediate interpretation in physical terms. A similar situation occurs in quantum mechanics. The physical characterization of QES requires the analysis of the representation of $\mathcal{G}_0$ in terms of the observables associated with its Lie algebra. Then, we have two different methods of investigating elementary systems, the first is based on the global actions of Lie groups and the second is based on the representations of Lie algebras. The global method is appropriated for classifying the elementary systems and the local one is convenient for characterizing them from a physical point of view. The analysis of the relation between these two approaches is the main object of this work. In particular we are interested in two aspects.

(1) How Lie group action, in classical and quantum mechanics, defines realizations of algebraic structures as the symmetric and the enveloping algebras.

(2) In what way the analysis of the algebraic structures may be used to obtain information about the properties of CES and QES.

The first question may be conveniently simplified by using the projective group[7] $\hat{\mathcal{G}}_0$ of $\mathcal{G}_0$. In Part C of Sec. 2 we prove that the CES of $\mathcal{G}_0$ are described by the transitive strict canonical realizations of $\hat{\mathcal{G}}_0$ (Theorem 1). It allows us to work with representations of Lie algebras, since the infinitesimal exponents of $\mathcal{G}_0$ associated with their CES and included in the Lie group structure of $\hat{\mathcal{G}}_0$. Thus, this result extends to classical mechanics the Bargmann[8] analysis of projective unitary representations. Moreover, it is formulated in terms of a unique central extension of $\mathcal{G}_0$ given by the projective group $\hat{\mathcal{G}}_0$. This considerably simplifies the formalism of Souriau[4] and constitutes the version in classical mechanics of the result of Cariñena and Santander[7] about projective irreducible unitary representations. Therefore, we have that both CES and QES of $\mathcal{G}_0$ define representations of the Lie algebra of $\hat{\mathcal{G}}_0$. This provides the bridge between global and local methods in order to analyze elementary systems. On the other hand, it allows us to apply the analysis of the observables associated with Lie algebras[1] to the characterization of the elementary systems. We use the representation theory of enveloping algebras to study, by means of the adjoint action, the transformation properties of quantum observables and to classify QES

a) Supported by the Junta de Energía Nuclear, Madrid, Spain.

in terms of the quantum numbers provided by the Casimir invariants. Nevertheless, a similar tool has not been considered rigorously in classical mechanics. In this paper we show how the representation theory of symmetric algebras may be used in this context. Thus, we prove (Theorem 3) that the transformation properties of classical observables may be analyzed by means of adjoint action. Moreover, we obtain that the invariants of the symmetric algebra play the same role in classical mechanics as the Casimir invariants in quantum mechanics. In particular, this justifies the formal use of the adjoint action implicit in the formalism of Sudarshan and Mukunda (see, for example, Ref. 5, p. 226).

Question (2) is investigated in Sec. 3 which is devoted to the coadjoint action of Lie groups. This action is fundamental to constructing the CES of a Lie group.[4,9,10] We investigate how the characteristic dimensions[1] of a Lie algebra are related to the properties of the orbits under the coadjoint action. Two results are obtained (Proposition 4 and 5) which in physical terms mean that the characteristic dimensions of $\hat{G}_0$ represent the number of degrees of freedom and the number of labeling parameters of the generic CES of $\hat{\mathcal{G}}_0$. These conclusions seem to hold also for QES. It is part of the empirical analogies between the orbits under the coadjoint action and the unitary irreducible representations of Lie groups. These analogies are the origin of the "orbit method" of Kirillov[11] and the geometric quantization program of Kostant[12] and Souriau.[4]

Section 4 is devoted to the applications of the formalism to the Galilei, Poincare and Weyl invariances. We emphasize, in particular, the analysis of the existence and the properties of such observables as the position and the spin for different elementary systems. This analysis is carried out by using the representation theory of the algebraic structures and the adjoint action. In this manner, we obtain the well-known results about localizability in quantum mechanics. Moreover, we show that the conclusions for classical mechanics turn out to be completely similar to the ones obtained at the quantum level. Especially interesting is the analysis of Weyl invariance, since the Weyl group provides a manifiestly covariant description of elementary systems which admits an observable describing the "age" of the system.

2. INVARIANCE GROUPS

A. Dynamical systems

In both classical and quantum mechanics, the notion of "state" appears as the initial condition which determines the solutions of the evolution equation of the system. These two concepts, state and evolution law, are the fundamental ones in the mathematical description of the physical systems. The next definition will be useful in discussing the common aspects arising in the dynamical formalism of classical and quantum mechanics.

Definition 1: By a dynamical system we shall mean the pair $(S, e(\cdot,\cdot))$ formed by a nonempty set S and a two parameter family of bijective maps $e(t_1, t_2)$ $(t_1, t_2 \in \mathbb{R})$ of S onto S with the following properties:

(i) $e(t,t)$ is the identity map on S for all $t \in \mathbb{R}$,

(ii) $e(t_1, t_2) = e(t_1, t_3)\, e(t_3, t_2)$ for all $t_1, t_2, t_3 \in \mathbb{R}$.

The set S and the family $\{e(t_1, t_2): t_1, t_2 \in \mathbb{R}\}$ will be called state space and evolution law respectively. By one evolution of the dynamical system we shall mean a curve $\pi = (t, x(t))$ in $\mathbb{R} \times S$ verifying $x(t) = e(t, t')\, x(t')$ for all $t, t' \in \mathbb{R}$. It is clear that every point $(t_0, x_0) \in \mathbb{R} \times S$ belongs to one and only one evolution π given by $\pi = (t, e(t_1 t_0) x_0)$. We shall denote by $E(S)$ the set of evolutions of $(S, e(\cdot,\cdot))$. If the evolution law $e(t_1, t_2)$ depends only upon the difference $t_1 - t_2$, i.e., $e(t_1, t_2) = e(t_1 - t_2)$ then $e(t)$ $(t \in \mathbb{R})$ becomes a one-parameter group of bijective maps of S onto S. In this case we shall say that the dynamical system is conservative.

The cornerstone of the formulation of the invariance principles is the notion of group action; let us briefly define what we mean by this term.

Definition 2: Let X be a nonempty set and let $\mathcal{G}$ be a group. By an action $(R, \mathcal{G}, S)$ of $\mathcal{G}$ over X we shall mean a map

$$\mathcal{G} \times X \to X \quad (g, p) \to R(g)\, p,$$

which verifies

(i) $R(g)$ is a bijective map of X onto X for all $g \in \mathcal{G}$,

(ii) $R(e)$ is the identify map on X,

(iii) $R(g_1 g_2) = R(g_1) R(g_2)$ for all $g_1, g_2 \in \mathcal{G}$.

Definition 3: Let $(S, e(\cdot,\cdot))$ be a dynamical system and let $\mathcal{G}$ be a group. By a kinematical action of $\mathcal{G}$ on this dynamical system we shall mean an action $(R, \mathcal{G}, S)$ of $\mathcal{G}$ on the state space S. On the other hand, by a dynamical action of $\mathcal{G}$ we shall mean an action $(R, \mathcal{G}, E(S))$ of $\mathcal{G}$ on the set of evolutions $E(S)$.

Both classes of actions of groups on dynamical systems are related as shows the next proposition which follows at once from the above definitions,

Proposition 1: Every kinematical action $(R^{(k)}, \mathcal{G}, S)$ defines a dynamical action $(R^{(d)}, \mathcal{G}, E(S))$ by means of the map

$$\pi = (t, x(t)) \to R^{(d)}(g)\, \pi = (t, e(t,0)\, R^{(k)}(g)\, e(0,t)\, x(t)),$$
$$\pi \in E(S), \quad g \in \mathcal{G}.$$

Moreover this map determines a bijective correspondence between the sets of kinematical and dinamical actions of $\mathcal{G}$.

If $\mathcal{G}_0$ is the group associated with an invariance principle, the invariance of the physical laws under changes of reference frame related by $\mathcal{G}_0$ requires the existence of a dynamical action of $\mathcal{G}_0$ over the dynamical systems which describe the physical phenomena. In this context, it is generally assumed that the invariance under the group $\mathcal{T} = \{g(b) : b \in \mathbb{R}\}$ of time translations is accomplished by means of the map:

$$\mathcal{T} \times E(S) \to E(S), \quad \pi = (t, x(t)) \to T^{(d)}(b)\pi \equiv (t, x(t-b)), \quad b \in \mathbb{R}. \qquad (1)$$

This postulate is not trivial, in fact it is easy to prove.

Proposition 2: The map (1) defines a dynamical action of $\mathcal{T}$ over $(S, e(\cdot,\cdot))$ if and only if $(S, e(\cdot,\cdot))$ is conservative. Moreover, in this case the evolution law verifies $e(t) = T^{(k)}(-t)(t\in\mathbb{R})$, where $T^{(k)}$ is the kinematical action of $\mathcal{T}$ associated with the dynamical action $T^{(d)}$.

We are interested in groups $\mathcal{G}_0$ which contain $\mathcal{T}$ as a subgroup. From Propositions 1 and 2 we see that the analysis of the consequences derived from the invariance under these groups $\mathcal{G}_0$ may be carried out in terms of the kinematical actions of $\mathcal{G}_0$. Thus the dynamical systems compatible with the invariance group $\mathcal{G}_0$ are in correspondence with the kinematical actions $(R^{(k)}, \mathcal{G}_0, S)$, the evolution law being determined by the action of the subgroup $\mathcal{T}$ of time translations.

B. Elementary systems

Henceforth, we assume that $\mathcal{G}_0$ is a connected Lie group which contains as a subgroup the group $\mathcal{T}$ of time translations. The Lie algebra of $\mathcal{G}_0$ will be denoted by G_0 and we shall fix a basis $B_0 = \{A_i; i = 1, \ldots, N_0\}$ with the commutation relations $[A_i, A_j] = \sum_k c_{ij}^k A_k$.

In classical mechanics the state space of a dynamical system is a phase space with coordinates and canonically conjugate momenta. It is described mathematically by a symplectic manifold,[13] that is a differentiable manifold M equipped with a Poisson bracket $\{\ ,\ \}$. The kinematical action of the invariance group $\mathcal{G}_0$ is given by a canonical realization[10] $(r, \mathcal{G}_0, M)$ of $\mathcal{G}_0$ over M. These actions of $\mathcal{G}_0$ have an important property: They are locally Hamiltonian.[13] Then, in a neighborhood N of each point $x_0\in M$ there is a map $A\in G_0 \to \tilde{A} \in C^\infty(N)$ where $\tilde{A}$ is unique up to additive constants and verifies

$$\{\tilde{A}, f\}(x) = \frac{d}{dt} f[r(\exp(-tA))\, x]\Big|_{t=0}, \tag{2}$$

for all $x\in N$ and $f\in C^\infty(N)$. One finds that the functions corresponding to the elements of the basis B_0 of G_0 verify the Poisson bracket relations

$$\{\tilde{A}_i, \tilde{A}_j\} = \sum_k c_{ij}^k \tilde{A}_k + \eta(A_i, A_j), \tag{3}$$

where $\eta(A_i, A_j)$ are real constants which define an equivalence class of infinitesimal exponents[5] of G_0. If this class is the trivial one we shall say that the canonical realization is "strict." In general, the maps $A\in G_0 \to \tilde{A} \in C^\infty(N)$ are not extended to the whole M. When it holds, the canonical realization is said to be Hamiltonian. It is known[4] that this is the case if M is simply connected or if G_0 verifies $[G_0, G_0] = G_0$.

From the point of view of Lie group action, the irreducible objects are those realizations $(r, \mathcal{G}_0, M)$ such that for all $x, y\in M$ there is $g\in\mathcal{G}_0$ verifying $r(g)x = y$. They are called transitive canonical realizations. This suggests the following mathematical characterization of elementary systems in classical mechanics.

Definition 4: By a classical elementary system (CES) of $\mathcal{G}_0$ we shall mean a transitive canonical realization $(r, \mathcal{G}_0, M)$ of $\mathcal{G}_0$. Two CES $(r_i, \mathcal{G}_0, M_i)$ $(i = 1, 2)$ of $\mathcal{G}_0$ are said to be equivalent if there is a canonical diffeomorphism $\tau: M_1 \to M_2$ such that $\tau\circ r_1(g) = r_2(g)\circ\tau$ for all $g\in\mathcal{G}_0$.

In quantum mechanics the state space of a dynamical system is the set of unit rays of a complex Hilbert space H. The kinematical action of the invariance group $\mathcal{G}_0$ is given by a projective unitary representation $(U, \mathcal{G}_0, H)$ of $\mathcal{G}_0$ over H. Each $A\in G_0$ defines a self-adjoint operator $\bar{A}$ on H unique up to additive constants and satisfies

$$\bar{A}\,\psi = i\frac{d}{dt} U(\exp[tA])\psi\big|_{t=0}, \qquad \psi\in H. \tag{4}$$

The image of the basis B_0 of G_0 under this map verifies the commutation relations

$$[\bar{A}_i, \bar{A}_j] = i\sum_k c_{ij}^k \bar{A}_k + i\eta(A_i, A_j), \tag{5}$$

where $\eta(A_i, A_j)$ are real constants which determine an equivalence class of infinitesimal exponents of G_0. We now give the usual definition of elementary systems in quantum mechanics.

Definition 5: By a quantum elementary system (QES) of $\mathcal{G}_0$ we shall mean an irreducible projective unitary representation $(U, \mathcal{G}_0, H)$ of $\mathcal{G}_0$. Two QES $(U_i, \mathcal{G}_0, H_i)$ $(i = 1, 2)$ of $\mathcal{G}_0$ are said to be equivalent if there is a unitary transformation $V: H_1 \to H_2$ such that $VU_1(g) = U_2(g)V$ for all $g\in\mathcal{G}_0$.

From (3) and (5) we see that both CES and QES define representations up to a factor of the Lie algebra G_0. At this point, it is convenient to introduce the projective group[7] $\hat{\mathcal{G}}_0$ of $\mathcal{G}_0$. This group is defined in terms of a basis $\{\eta_r; r = 1, \ldots, m\}$ of infinitesimal exponents of G_0 as the simply connected Lie group with Lie algebra $\hat{G}_0$ verifying

$$\begin{aligned}
&[A_i, A_j] = \sum_k c_{ij}^k A_k + \sum_r \eta_r(A_i, A_j)\, M_r,\\
&[A_i, M_r] = [M_r, M_s] = 0.
\end{aligned} \tag{6}$$

If we denote by $\mathcal{G}_0^*$ the universal covering group of $\mathcal{G}_0$, we have[7] that $\hat{\mathcal{G}}_0$ is a central extension of $\mathcal{G}_0^*$ by the Abelian group $\mathbb{R}^m$. Then, there is a exact sequence

$$0 \to \mathbb{R}^m \xrightarrow{t} \hat{\mathcal{G}}_0 \xrightarrow{q} \mathcal{G}_0^* \to 0. \tag{7}$$

The composition $\hat{q} \equiv p\circ q$ of $q: \hat{\mathcal{G}}_0 \to \mathcal{G}_0^*$ with the covering homomorphism $p: \mathcal{G}_0^* \to \mathcal{G}_0$ is called the projective homomorphism of $\hat{\mathcal{G}}_0$ onto $\mathcal{G}_0$.

Now, we shall prove that the CES of $\mathcal{G}_0$ may be identified with strict canonical realizations of $\hat{\mathcal{G}}_0$.

Theorem 1: Let $(\hat{r}, \hat{\mathcal{G}}_0, M)$ be a transitive strict canonical realization of $\hat{\mathcal{G}}_0$ such that the kernel of the projective homomorphism $\hat{q}: \hat{\mathcal{G}}_0 \to \mathcal{G}_0$ acts identically on M. Then, the map

$$\mathcal{G}_0 \times M \to M, \quad (g, x) \to \hat{r}(\hat{q}^{-1}(g))\, x \tag{8}$$

defines a CES of $\mathcal{G}_0$. Moreover each CES of $\mathcal{G}_0$ is equivalent to one of this form.

Proof: Given $(\hat{r}, \hat{\mathcal{G}}_0, M)$ let us denote $r(g) \equiv \hat{r}(\hat{q}^{-1}(g))$, $g\in\mathcal{G}_0$. Since Ker$\hat{q}$ acts trivially, it is clear that $r(g)$ is a canonical transformation on M for all $g\in\mathcal{G}_0$. Moreover, it follows easily that $r(g_1 g_2) = r(g_1)\, r(g_2)$ and that $r(e)$ is the identity map on M. The unique nontrivial point, in order to conclude that $(r, \mathcal{G}_0, M)$ is a CES of $\mathcal{G}_0$, is to prove that $(g, x)\in\mathcal{G}_0\times M \to r(g)x\in M$ is a C^∞ map. From (7) we have that $q: \hat{\mathcal{G}}_0 \to \mathcal{G}_0^*$ is a C^∞ homomorphism with central kernel in $\hat{\mathcal{G}}_0$. A result due to Hochschild[14] shows the existence of a C^∞ section $c:\mathcal{G}_0^* \to \hat{\mathcal{G}}_0$ verifying $q(c(g^*))=g^*$ for all $g^*\in\mathcal{G}_0^*$. On the other hand, since $p:\mathcal{G}_0^* \to \mathcal{G}_0$ is a local isomorphism, there is a neighborhood U of the identity element in $\mathcal{G}_0$ such that p^{-1} is a diffeomorphism of $U\in\mathcal{G}_0$ onto $p^{-1}(U)\subset\mathcal{G}_0^*$. Then $(g,x)\in U\times M \to r(g)x\in M$ is a C^∞ map since it is the composition of the following C^∞ maps:

$$U\times M \longrightarrow \mathcal{G}_0^*\times M \longrightarrow \hat{\mathcal{G}}_0\times M \longrightarrow M,$$

$$(g,x) \longrightarrow (p^{-1}(g) \longrightarrow [c(p^{-1}(g)),x] \longrightarrow r[c(p^{-1}(g))]x.$$

Now, for each $g_0\in\mathcal{G}_0$ the set g_0U is a neighborhood of g_0 and $(g,x)\in g_0U\times M\to r(g)x\in M$ is a C^∞ map since it is the the composition of the following C^∞ maps:

$$g_0U\times M \longrightarrow U\times M \longrightarrow M \longrightarrow M,$$

$$(g,x) \longrightarrow (g_0^{-1}g,x) \longrightarrow r(g_0^{-1}g)x \longrightarrow r(g_0)\,r(g_0^{-1}g)x.$$

Then, we conclude that $(g,x)\in\mathcal{G}_0 x M\to r(g)x\in M$ is a C^∞ map and therefore $(r,\mathcal{G}_0,M)$ is a CES.

On the other hand, given a CES $(r,\mathcal{G}_0,M)$ of $\mathcal{G}_0$, it follows at once that $\hat{r}\equiv r\circ\hat{q}$ defines a transitive canonical realization of $\hat{\mathcal{G}}_0$. In order to prove that $(\hat{r},\hat{\mathcal{G}}_0,M)$ is a strict realization, let us suppose that η is the infinitesimal exponent of G_0 defined by the map $A\in G_0 \to \tilde{A}\in C^\infty(N)$ in the neighborhood of some point of M. We may redefine this correspondence in such a form that there is a linear dependence $\eta=\sum_r m_r\eta_r$. Then, from (5) and (8) we deduce that $(\hat{r},\hat{\mathcal{G}}_0,M)$ is an strict realization. This proves the assertion. (Q.E.D.)

Similarly, we have the following result[7] about the reduction of the QES of $\mathcal{G}_0$ to unitary representations of $\hat{\mathcal{G}}_0$.

Theorem 2: Let $(\hat{U},\hat{\mathcal{G}}_0,U)$ be an irreducible unitary representation of $\hat{\mathcal{G}}_0$ mapping into $U(1)$ the kernel of the projective homomorphism $\hat{q}:\hat{\mathcal{G}}_0\to\mathcal{G}_0$. Then, the map

$$\mathcal{G}_0\times H\to\mathcal{G}, \quad (g,\psi)\to\hat{U}(\hat{q}^{-1}(g))\psi, \tag{9}$$

defines a QES of $\mathcal{G}_0$. Moreover each QES of $\mathcal{G}_0$ is equivalent to one of this form.

From Theorems 1 and 2 we have that both CES and QES define representations of the Lie algebra $\hat{G}_0$. Let us notice that for a CES of $\mathcal{G}_0$ described by a transitive strict canonical realization $(\hat{r},\hat{\mathcal{G}}_0,M)$ of $\hat{\mathcal{G}}_0$, the representation of $\hat{G}_0$ is only locally defined in a neighborhood of each element of M. However, all of these realizations of $\hat{\mathcal{G}}_0$ may be constructed in terms of transitive strict Hamiltonian realizations of $\hat{\mathcal{G}}_0$.[11]

C. Classical and quantum observables

One of the most important aspects which should be analyzed for the characterization of observables is their behavior under the action of the invariance group. This is fundamental in order to assign an appropriate mathematical object with a concrete physical observable.

Let $(\hat{r},\hat{\mathcal{G}}_0,M)$ be a CES of $\mathcal{G}_0$. The observables of this classical system are described by real functions f defined over the phase space M, and the action of $\hat{\mathcal{G}}_0$ on them is given by:

$$\hat{r}(\hat{g})(f)=f\circ\hat{r}(\hat{g})^{-1}, \quad \hat{g}\in\hat{\mathcal{G}}_0. \tag{10}$$

Analogously, given a QES $(\hat{U},\hat{\mathcal{G}}_0,H)$ of $\mathcal{G}_0$, the quantum observables are described by the self-adjoint operators F on H and the action of $\hat{\mathcal{G}}_0$ on them is given by

$$\hat{U}(\hat{g})(F)=\hat{U}(\hat{g})\,F\hat{U}(\hat{g})^{-1}, \quad \hat{g}\in\hat{\mathcal{G}}_0. \tag{11}$$

The analysis of the actions (10) and (11) are particularly easy for the observables coming from the algebraic structures associated with $\hat{G}_0$. This will be clear from the following discussion.

Let $\mathcal{G}$ be a connected Lie group with Lie algebra G and let S be the symmetric algebra of G.[1] Given a basis $B=\{A_\alpha:\alpha=1,\ldots,N\}$ of G we may identify S with the polynomial ring $\mathbb{C}[a_1,\ldots,a_N]$ in N variables equipped with a Lie algebra structure defined by

$$\{p_1,p_2\}=\sum_{\alpha,\beta,\nu}c^\nu_{\alpha\beta}a_\nu\frac{\partial p_1}{\partial a_\alpha}\frac{\partial p_2}{\partial a_\beta}, \quad p_1,p_2\in S, \tag{12}$$

where $\{c^\nu_{\alpha\beta}:\alpha,\beta,\nu=1,\ldots,N\}$ are the structure constants of G in the basis B.

Let $(r,\mathcal{G},M)$ be a strict Hamiltonian realization of $\mathcal{G}$. Since the associated infinitesimal exponent η is the trivial one, we can choose the map $A\in G\to\tilde{A}\in C^\infty(M)$ such that it defines a Lie algebra homomorphism. This map may be extended to the symmetric algebra defining $\tilde{p}\equiv p(\tilde{A}_1,\ldots,\tilde{A}_N)$ for all $p=p(a_1,\ldots,a_N)$ in S. We have:

Theorem 3: The map $p\in S\to\tilde{p}\in C^\infty(M)$ verifies,

(i) It is a representation of S,

(ii) For each $p\in S$, $r(g)(\tilde{p})=[\mathrm{ad}g(p)]^\sim$, where $\mathrm{ad}g$ $(g\in\mathcal{G})$ denotes the adjoint action of $\mathcal{G}$ over S.

Proof: (i) By its definition it is obvious that $p\to\tilde{p}$ is a linear map which verifies $(p_1p_2)^\sim=\tilde{p}_1\tilde{p}_2$. Moreover from (12) we deduce $\{p_1,p_2\}^\sim=\{\tilde{p}_1,\tilde{p}_2\}$. Then the conclusion follows.

(ii) It is sufficient to prove that it is true for the Lie algebra G. Given $A\in G$, we have

$$\{r(g)(\tilde{A}),f\}=\{\tilde{A}\circ r(g^{-1}),f\}=\{\tilde{A},f\circ r(g)\}\circ r(g^{-1})$$

for all $f\in C^\infty(M)$. On the other hand, we obtain

$$\begin{aligned}\{[\mathrm{ad}g(A)]^\sim,f\}(x)&=\frac{d}{dt}f[r(\exp[-t\,\mathrm{ad}g(A)])x]\Big|_{t=0}\\&=\frac{d}{dt}f(r(g)r[\exp(-tA)]r(g^{-1})x)\Big|_{t=0}\\&=\{\tilde{A},f\circ r(g)\}(r(g^{-1})x).\end{aligned}$$

Hence $r(g)(\tilde{A})=[\mathrm{ad}g(A)]^\sim$ and the conclusion follows.

Q.E.D.

Let S^I be the set of invariants in S under the adjoint action. From part (ii) of Theorem 3, every $p\in S^I$ verifies $r(g)(\tilde{p})=\tilde{p}$ for all $g\in\mathcal{G}$. Then $\tilde{p}$ is an invariant function on M under the action of $\mathcal{G}$. Hence if $(r,\mathcal{G},M)$ is a transitive action we conclude that the invariants in S are represented by constant functions over M.

These results are similar to the well-known properties about the representations of enveloping algebras[1] induced by unitary representations in Hilbert spaces.[15] Given a unitary representation $(U,\mathcal{G},H)$ there exists a dense domain H^∞ (the Gärding domain) in H which is left invariant under the action of $\mathcal{G}$ and also by the operators associated with the Lie algebra G. Thus, we get a representation of G by operators defined on H^∞ and we can associate with every element u of the enveloping algebra U of G an operator $\tilde{u}$ defined on H^∞. If we denote by $O^\infty(H)$ the set of operators of H^∞ into itself, we can write the result analogous to Theorem 3 in terms of enveloping algebras in the following way:

Theorem 4: The map $u\in U\rightarrow \tilde{u}\in O^\infty(H)$ verifies,

(i) It is a representation of U,

(ii) For each $u\in U$, $U(g)\,(\tilde{u})=[\mathrm{ad}g(u)]^\sim$ where $\mathrm{ad}g(g\in\mathcal{G})$ denotes the adjoint action of $\mathcal{G}$ over U.

Similarly, when $(U,\mathcal{G},H)$ is irreducible the elements of the set U^I of invariants in U under the adjoint action are represented by constant operators.

It is not possible to extend the representations of S and U to all the elements of their quotient fields $D(S)$ and $D(U)$, respectively. Indeed given $h=p_1/p_2\in D(S)$ $(p_1,p_2\in S)$, the function $\tilde{h}=\tilde{p}_1/\tilde{p}_2$ is not always well defined on the whole M. However within the definition domain of $\tilde{h}$ it holds $r(g)(\tilde{h})=[\mathrm{ad}g(h)]^\sim$ where $\mathrm{ad}g(g\in\mathcal{G})$ is the adjoint action of $\mathcal{G}$ over $D(S)$. The same comments may be translated to the elements of $D\,(U)$

With respect to the application of Theorems 3 and 4 in the formalism of classical and quantum mechanics, it is clear from Theorems 1 and 2 that the relevant algebraic structures of observables are those associated with the Lie algebra of the projective group. We emphasize the fundamental role played by the adjoint action of the projective group in the transformation properties of the observables. In this way a wide class of observables may be studied in terms of the adjoint action of $\hat{\mathcal{G}}_0$ over the algebraic structures associated with its Lie algebra. It is known[1] that several physically interesting observables are in the quotient structures $D(S)$ and $D\,(U)$ or more general ones. Then they may not be defined for all the elementary systems, since only S and U admit always a well-defined representation. Thus, the analysis of the representations of the algebraic structures enables us to know whether a given observable is or is not admissible for an elementary system.

3. THE COADJOINT ACTION

A. The orbits under the coadjoint action

Let $\mathcal{G}$ be a connected Lie group with Lie algebra G and let G^* be the dual vector space of G. The coadjoint action (cad, $\mathcal{G}$, G^*) of $\mathcal{G}$ on G^* is given by[11,16]

$$\langle \mathrm{cad}(g)a,A\rangle \equiv \langle a,\mathrm{ad}(g)^{-1}A\rangle,\quad g\in\mathcal{G},\quad a\in G^*, A\in G. \quad (13)$$

Let $B=\{A_\alpha:\alpha=1,\ \dots,\ N\}$ be a basis of G with commutation relations $[A_\alpha,A_\beta]=\sum_\nu c^\nu_{\alpha\beta}A_\nu$. We shall denote by $B^*=\{a^*_\alpha;\alpha=1,\ \dots,\ N\}$ the dual basis of B (i.e., $\langle a^*_\alpha,A_\beta\rangle=\delta_{\alpha\beta}$) and by $\{a_\alpha:\alpha=1,\ \dots,\ N\}$ the coordinate functions in G^* associated with B^*. It follows immediately that (cad, $\mathcal{G}$, G^*) is a linear action of $\mathcal{G}$ with matrix representation relative to B^* verifying

$$(\mathrm{cad}(g))_{\alpha\beta}=(\mathrm{ad}(g)^{-1})_{\beta\alpha}, \quad (14)$$

where $(\mathrm{ad}(g)^{-1})$ is the matrix associated to $\mathrm{ad}(g)^{-1}$ in the basis B.

Example: Consider the SU(2) Lie group. Its elements may be written as

$$A(\mathbf{n},\theta)=\cos\theta/2-i(\mathbf{n}\sigma)\sin\theta/2,\ |\mathbf{n}|=1, \quad (15)$$

where $\sigma=(\sigma_1,\sigma_2,\sigma_3)$ are the standard Pauli matrices. A basis of G is given by $B=\{\boldsymbol{\zeta}=(1/2i)\sigma\}$. The adjoint action of SU(2) is of the form

$$\mathrm{ad}A(\mathbf{n},\theta)(\mathbf{r}\boldsymbol{\zeta})=(R(\mathbf{n},\theta)\mathbf{r})\cdot\boldsymbol{\zeta},$$

with

$$R(\mathbf{n}\theta)\mathbf{r}=\cos\theta\mathbf{r}+(1-\cos\theta)(\mathbf{nr})\mathbf{n}+\sin\theta(\mathbf{n}\times\mathbf{r}).$$

Since $R(\mathbf{n},\theta)\in\mathrm{SO}(3)$, then $R(\mathbf{n},\theta)_{ij}=[R(\mathbf{n},\theta)^{-1}]_{ji}$. Therefore, the coadjoint action of SU(2) relative to the dual basis $B^*=\{\mathbf{s}^*\}$ is

$$\mathrm{cad}A(\mathbf{n},\theta)(\mathbf{s}\circ\mathbf{s}^*)=(R(\mathbf{n},\theta)\mathbf{s})\cdot\mathbf{s}^*.$$

The coadjoint action of $\mathcal{G}$ is generated by a linear action of its Lie algebra G defined by

$$\langle\mathrm{cad}(A)a,A'\rangle=\langle a,[A',A]\rangle,\ \ A,A'\in G,\ a\in G^*. \quad (16)$$

The right member of this equation coincides with the expression of the so-called "Kirillov form"[11]

$$B_a(A',A)\equiv\langle a,[A',A]\rangle,\ \ A,A'\in G,\ a\in G^*, \quad (17)$$

which defines a skew-symmetric bilinear form over G for every $a\in G^*$.

With every $a\in G^*$ we associate the set $O_a \equiv\{\mathrm{cad}(g)\,a;g\in\mathcal{G}\}$ called the orbit of a, and the closed subgroup $\mathcal{G}_a\equiv\{g\in\mathcal{G};\mathrm{cad}(g)\,a=a\}$ called the isotropy subgroup at a. From the theory of Lie group action over differentiable manifolds,[17] it follows that O_a has a submanifold structure of G^* diffeomorphic to the quotient manifold $\mathcal{G}/\mathcal{G}_a$. Using this result we obtain,

Proposition 3: The dimension of O_a equals the rank of the Kirillov form B_a at a.

Proof: Since O_a is a submanifold of G^* diffeomorphic to $\mathcal{G}/\mathcal{G}_a$, then dim $O_a=$ dim $\mathcal{G}$ − dim $\mathcal{G}_a=$ dim G − dim G_a, where G_a is the Lie algebra of the subgroup $\mathcal{G}_a$. On the other hand, from (16) and (17) we deduce that $G_a=\{A\in G:B_a(A,A')=0\ \forall A'\in G\}$; this implies that dim $G_a=$ dim G-rank B_a. Therefore, we conclude dim $O_a=$ rank B_a. (Q.E.D.)

There is a deep influence of the algebraic properties of the Lie algebra G over the structure of the orbits O_a. Let us remind[1] ourselves that at the algebraic level G is characterized by two nonnegative integers defined in terms of the matrix function $M_G(a)_{\alpha\beta}=\sum_\nu c^\nu_{\alpha\beta}a_\nu$ by

$$n(G)\equiv\tfrac{1}{2}\sup_{(a_1,\dots,a_N)}\ \mathrm{rank}M_G(a),\quad S(G)\equiv\dim G-2n(G). \quad (18)$$

We have called them canonical and central dimensions respectively of G. The role of these integers is fundamental in the analysis of the "generic orbits," that is, the orbits in G^* of maximal dimension. Let $G^*_{\max}$ be the set of points $a\in G^*$ such that dim O_a is maximal; we have,

Proposition 4: $a\in G^*_{\max}\Rightarrow$ dim $O_a=2n(G)$. (19)

Proof: By (17) we have $B_a(A_\alpha,A_\beta)=\sum_\nu c^\nu_{\alpha\beta}a_\nu$, where $\{a_\nu:\nu=1,\ \dots,\ N\}$ are the coordinates of the point $a\in G^*$ with respect to the dual basis B^*, so that we obtain

$$\operatorname{rank} M_{\mathcal{G}}(a) = \operatorname{rank} B_a. \tag{20}$$

Therefore, if $a \in G^*_{\max}$, from Proposition 3 and (18) we deduce that $\dim \mathcal{O}_a = 2n(G)$. (Q. E. D.)

In order to find the orbits it is very important to know the invariant functions under the action of $\mathcal{G}$. A function f differentiable in a neighborhood of a in G^* is said to be invariant at a if there is a neighborhood U of (e, a) in $\mathcal{G} \times G^*$ on which $f(\mathrm{cad}(g)a') = f(a')$ for all $(g, a') \in U$. The set of these functions will be denoted by $C^I(a, G^*)$. In addition we shall denote by $C^I(G^*)$ the set of differentiable functions f on G^* such that $f[\mathrm{cad}g(a)] = f(a)$ for all $g \in \mathcal{G}$ and $a \in G^*$; they are the global invariant functions. It follows easily from (16) that the action of G over the functions f defined in G^* may be written:

$$(\mathrm{cad}(A_\alpha)f)(a) = \sum_{\beta,\nu} c^\nu_{\alpha\beta} a_\nu \frac{\partial f}{\partial a_\beta}. \tag{21}$$

Since $\mathcal{G}$ is a connected Lie group, then $f \in C^I(a, G^*)$ if and only if $\mathrm{cad}(A_\alpha)f$ vanishes on some neighborhood of a for all $\alpha = 1, \ldots, N$. Similarly $f \in C^I(G^*)$ if and only if $\mathrm{cad}(A_\alpha)f$ vanish on G^* for all $\alpha = 1, \ldots, N$. At this point we find an important relation with the theory of Casimir invariants, since from (21) the set $C^I(a, G^*)$ coincides with the set of "formal invariants" of G[18] defined at the point a. Moreover we have the following result about the maximal number of functionally independent elements in $C^I(a, G^*)$:

Proposition 5: $a \in G^*_{\max} \Rightarrow \dim C^I(a, G^*) = s(G)$. (22)

Proof: If $a \in G^*_{\max}$, Frobenius Theorem[19] implies $\dim C^I(a, G^*) = \dim G^* - \dim \mathcal{O}_a$. Therefore, from (18) and Proposition 4 the conclusion follows. (Q. E. D.)

The main property of the orbits under the coadjoint action is that they have the structure of sympletic manifolds. Moreover, the restriction $[(\mathrm{cad})^\sim, \mathcal{G}, \mathcal{O}]$ of the coadjoint action over a given orbit $\mathcal{O}$ is a transitive strict Hamiltonian realization of $\mathcal{G}$.[10,11] Since the dimension of a symplectic manifold is twice the number of pairs of canonically conjugate coordinates, Proposition 4 implies that the canonical dimension $n(G)$ of G is just this number for the orbits of maximal dimension.

It is known[11] that each transitive strict canonical realization of $\mathcal{G}$ is locally equivalent to a realization $[(\mathrm{cad})^\sim, \mathcal{G}, \mathcal{O}]$. The manifolds which are locally diffeomorphic to a given orbit $\mathcal{O}$ can be completely classified. Indeed, they are the covering manifolds of $\mathcal{O}$.[20] Moreover, if $\mathcal{O}$ is simply connected, its locally equivalent realizations are in fact equivalent to $[(\mathrm{cad})^\sim, \mathcal{G}, \mathcal{O}]$. The same is true[4] when the isotropy subgroup at the orbit $\mathcal{O}$ is connected. These results allow us to construct the CES of an invariance group $\mathcal{G}_0$ in terms of the realizations $[(\mathrm{cad})^\sim, \hat{\mathcal{G}}_0, \mathcal{O}]$ of its projective group $\hat{\mathcal{G}}_0$. From Propositions 4 and 5 we see the important role played by the characteristic dimensions in the context of CES. Thus, $n(\hat{G}_0)$ is the number of pairs of canonically conjugate variables of the CES of $\mathcal{G}_0$ with maximal dimension. On the other hand, $s(\hat{G}_0)$ coincides with the number of invariant functions associated with the CES of maximal dimension.

B. The sets of classical and quantum elementary systems

One of the most useful tools in the formulation of quantum mechanics has been the use of quantization rules to construct quantum observables. In this context, the Hamiltonian formalism play a central role. This is so since observables are constructed in both classical and quantum mechanics in terms of canonical variables. Then, there is a wide class of quantum observables which have a classical analog. Nevertheless, the spin observable has been considered as an exception in this correspondence. It fact, it was usually assumed that the spin is a consequence of relativistic quantum mechanics. But the analysis of the Galilean invariance[3] shows that the spin is also implicit in nonrelativistic quantum mechanics. On the other hand, the application of Lie group theory to classical mechanics[5,6] shows how the spin may be also considered as a classical observable. From this it may be expected that the analogies between classical and quantum systems have a group-theoretical origin. In particular, it must be implicit in the relation between Lie group actions in classical and quantum mechanics.

It is an empirical fact that there exist strong analogies between the set $\mathcal{E}(\mathcal{G})$ of realizations $(\mathrm{cad}, \mathcal{G}, \mathcal{O})$ and the set $\mathcal{R}(\mathcal{G})$ of equivalence classes of irreducible unitary representations $(U, \mathcal{G}, \mathcal{H})$. It is well known from the Kirillov's work[11] that there is a bijective correspondence between the sets $\mathcal{E}(\mathcal{G})$ and $\mathcal{R}(\mathcal{G})$ for simply connected nilpotent Lie groups. Unfortunately, the correspondence between orbits and irreducible unitary representations it is not so clear in other cases. However, there seems to exist a general rule to break $\mathcal{E}(\mathcal{G})$ and $\mathcal{R}(\mathcal{G})$ according to types of elements and to define a bijective correspondence between them. Let us see an example to illustrate these comments.

Example: Let E(2) be the bidimensional Euclidean group. The elements of E(2) are labeled by $(\mathbf{a}, R(\theta))$ with $\mathbf{a} \in \mathbb{R}^2$, $\theta \in [0, 2\pi)$. Its Lie algebra is generated by $\{P_1, P_2, L\}$ with commutations rules

$$[L, P_1] = P_2, \quad [L, P_2] = -P_1, \quad [P_1, P_2] = 0. \tag{23}$$

The coadjoint action is given by

$$\mathbf{p}' = R(\theta)\mathbf{p}, \quad l' = l + a_1 p_2' - a_2 p_1'. \tag{24}$$

Hence we find two types of orbits,

(I) $\{p\}$, $p \neq 0$. The orbit is $C_p(\mathbf{p}) \times \mathbb{R}(l)$ where C_p is the circle $|\mathbf{p}| = p$.

(II) $\{l\}$, $l \in \mathbb{R}$. The orbit is the point $(0, l)$.

Similarly there are two types of classes of irreducible unitary representations.

(I)′ $[p]$, $p \neq 0$. The group acts over the square integrable functions on C_p by

$$[U(\mathbf{a}, R)\psi](\mathbf{p}) = \exp(i\mathbf{p}\mathbf{a})\psi(R^{-1}\mathbf{p}). \tag{25}$$

(II)′ $[l]$, l positive integer. The one-dimensional representation

$$U(\mathbf{a}, R(\theta)) = \exp(il\theta). \tag{26}$$

It seems natural to associate (I) → (I)′, (II) → (II)′.

Thus, we are faced with two questions. The first question is to give a meaning to the term "type" in $\mathcal{E}(\mathcal{G})$ and $\mathcal{R}(\mathcal{G})$. The second is to define the correspondence between types of $\mathcal{E}(\mathcal{G})$ and $\mathcal{R}(\mathcal{G})$. Both are closely related with the classification of the action of $\mathcal{G}$ over the functions and the operators of the orbits and the unitary representations. In physical terms, if we think of $\mathcal{E}(\mathcal{G})$ and $\mathcal{R}(\mathcal{G})$ as sets of elementary systems, two classical or quantum elementary systems would be of the same type if they admit the same observables, and the correspondence between classical and quantum types would be the quantization.

In $\mathcal{E}(\mathcal{G})$ the appropriate definition of type turns out to be the concept of "stratum,"[21] i.e., two realizations $[(\text{cad})^{\sim}, \mathcal{G}, \mathcal{O}_i]$ $(i=1,2)$ are said to be in the same stratum when they have conjugate isotropy subgroups. As a consequence we easily obtain that there is a diffeomorphism $b: \mathcal{O}_1 \to \mathcal{O}_2$ which commutes with the coadjoint action of $\mathcal{G}$ over the orbits. Moreover, if we define a map $f \in C^{\infty}(\mathcal{O}_1) \to f' \in C^{\infty}(\mathcal{O}_2)$ by $f'(x) = f(b^{-1}x)$, we have that it also commutes with the action of $\mathcal{G}$ over the differentiable functions. Therefore $C^{\infty}(\mathcal{O}_1)$ and $C^{\infty}(\mathcal{O}_2)$ are identified from the point of view of the action of $\mathcal{G}$. This means physically that the CES described by the realizations $[(\text{cad})^{\sim}, \mathcal{G}, \mathcal{O}_i]$ $(i=1,2)$ admit the same observables. With respect to $\mathcal{R}(\mathcal{G})$, we do not know a definition which describes the types in a precise way. However, in practice they may be identified in terms of the invariants under the group action. Each type in $\mathcal{E}(\mathcal{G})$ defines a characteristic set of elements $\{h_1, \ldots, h_n\}$ in the classical algebraic structures $(S, D(S), \ldots)$, which are invariants under the action of the group over all the orbits of the type and hence each one of these orbits may be labeled by the constant values $\{\tilde{h}_1, \tilde{h}_2, \ldots, \tilde{h}_n\}$. This parametrization is useful to distinguish the orbits in a given type. Thus, for algebraic Lie groups the orbits of maximal dimension are almost completely distinguished[11] by the constant values of rational invariants. In many cases one is able to construct the quantized version $\{H_1, \ldots, H_n\}$ in the quantum algebraic structures $(\mathcal{U}, D(\mathcal{U}), \cdots)$ of the elements $\{h_1, \ldots, h_n\}$. The set of irreducible unitary representations on which $\{H_1, \ldots, H_n\}$ are invariants defines a type in $\mathcal{R}(\mathcal{G})$ which is the quantum analog of the type in $\mathcal{E}(\mathcal{G})$ associated with $\{h_1, \ldots, h_n\}$.

Another interesting feature of the analysis of types in $\mathcal{E}(\mathcal{G})$ and $\mathcal{R}(\mathcal{G})$ is their connection with the representations of Lie algebras. Thus, every type is associated with a particular realization of the Lie algebra G in terms of canonical variables. This fact may be observed in the formalism of Pauri and Prosperi[6] of the classification of CES, and also it seems to hold for QES. It is also related to the canonical properties of the algebraic structures associated with G^1. It seems to be generally valid that these structures can be constructed in terms of $n(G)$ pairs of canonical variables and $s(G)$ invariants. In particular, for algebraic Lie algebras this feature appears at the quotient field level. With the terminology of Pauri and Prosperi a set $\{q_i, p_i; x_r;$ $i=1, \ldots, n(G); r=1, \ldots, s(G)\}$ of canonical variables and invariants generating the classical algebraic structures of G is called a "classical regular schema" and it corresponds (see Propositions 4 and 5) to generic orbits. In practice it is not difficult to find the quantum analogue set $\{Q_i, P_i, X_r; i=1, \ldots, n(G); r=1, \ldots s(G)\}$ which defines an associated "quantum regular schema" and corresponds to a generic class of irreducible unitary representations. Therefore, the canonical and central dimensions $n(G)$ and $s(G)$ may be interpreted as the number of degrees of freedom and the number of labeling parameters of the generic elementary systems.

4. APPLICATIONS

A. Nonrelativistic elementary systems

Let $\mathcal{G}_0$ be the Galilei group; its projective group $\hat{\mathcal{G}}_0$ is constituted by the elements

$$g = (\tau, b, \mathbf{a}, \mathbf{v}, A), \quad \tau, b \in \mathbb{R}, \quad \mathbf{a}, \mathbf{v} \in \mathbb{R}^3, \quad A \in SU(2), \tag{27}$$

with the composition law

$$g_1 g_2 = (\tau_1 + \tau_2 + \omega_{12}, b_1 + b_2, \mathbf{a}_1 + R(A_1)\mathbf{v}_2 + b_2\mathbf{v}_1, \mathbf{v}_1 + R(A_1)\mathbf{v}_2, A_1 A_2), \tag{28}$$

where $\omega_{12} = v_1^2 b_2/2 + \mathbf{v}_1 \cdot R(A_1)\mathbf{a}_2$ and $R(A)$ is the image in SO(3) of $A \in SU(2)$ under the covering homomorphism. The projective homomorphism $\hat{q}: \hat{\mathcal{G}}_0 \to \mathcal{G}_0$ is given by

$$(\tau, b, \mathbf{a}, \mathbf{v}, A) \longrightarrow (b, \mathbf{a}, \mathbf{v}, R(A)), \tag{29}$$

and evidently $\operatorname{Ker} \hat{q} = \{(\tau, 0, 0, 0, \pm \mathbb{1}) : \tau \in \mathbb{R}\}$ is a central subgroup of $\hat{\mathcal{G}}_0$. A basis $B = \{\mathcal{M}, \mathcal{H}, \mathcal{P}, \mathcal{K}, \mathcal{J}\}$ of the Lie algebra $\hat{G}_0$ is fixed by the following relations:

$$\begin{gathered}(\tau, 0, 0, 0, \mathbb{1}) = \exp(-\tau\mathcal{M}), \quad (0, b, 0, 0, \mathbb{1}) = \exp(-b\mathcal{H}),\\ (0, 0, \mathbf{a}, 0, \mathbb{1}) = \exp(\mathbf{a}\mathcal{K}), \quad (0, 0, 0, \mathbf{v}, \mathbb{1}) = \exp(\mathbf{v}\mathcal{K}),\\ (0, 0, 0, 0, A(\mathbf{n}, \theta)) = \exp(\theta \mathbf{n}\mathcal{J}),\end{gathered} \tag{30}$$

and it verifies

$$\begin{gathered}[\mathcal{J}_i, \mathcal{K}_j] = \epsilon_{ijk}\mathcal{K}_k, \quad [\mathcal{J}_i, \mathcal{P}_j] = \epsilon_{ijk}\mathcal{P}_k, \quad [\mathcal{J}_i, \mathcal{J}_j] = \epsilon_{ijk}\mathcal{J}_k,\\ [\mathcal{K}_i, \mathcal{P}_j] = -\delta_{ij}\mathcal{M}, \quad [\mathcal{K}_i, \mathcal{H}] = -\mathcal{P}_i.\end{gathered} \tag{31}$$

All other commutators are zero. The characteristic dimensions of $\hat{G}_0$ are $n(\hat{G}_0) = 4$ and $s(\hat{G}_0) = 3$.

In the coordinate system $\{m, h, \mathbf{p}, \mathbf{k}, \mathbf{j}\}$ of the dual basis B^* in $\hat{G}_0^*$ the coadjoint action may be written as follows:

$$\begin{aligned} m' &= m,\\ h' &= h + \tfrac{1}{2}mv^2 + (R(A)\mathbf{p}) \cdot \mathbf{v},\\ \mathbf{p}' &= R(A)\mathbf{p} + m\mathbf{v},\\ \mathbf{k}' &= R(A)\mathbf{k} + bR(A)\mathbf{p} + bm\mathbf{v} - m\mathbf{a},\\ \mathbf{j}' &= R(A)\mathbf{j} + \mathbf{v} \times R(A)\mathbf{k} + \mathbf{a} \times R(A)\mathbf{p} + m\mathbf{a} \times \mathbf{v},\end{aligned} \tag{32}$$

and determines three independent invariants

$$m, \; 2mh - \mathbf{p}^2, \; (m\mathbf{j} + \mathbf{k} \times \mathbf{p})^2. \tag{33}$$

The analysis of the algebraic structure of $D(S)$ suggests[1] the use of the coordinate system $\{m, u, \mathbf{q}, \mathbf{p}, \mathbf{s}\}$ in $\hat{G}_0^*$, where

$$\mathbf{q} = -\mathbf{k}/m, \; \mathbf{s} = \mathbf{j} - \mathbf{q} \times \mathbf{p}, \; u = h - \mathbf{p}^2/2m \tag{34}$$

have the physical meaning of position, spin, and internal energy, respectively. Let us note that these coordinates are defined only at points with $m \neq 0$, and in terms of them the coadjoint action reads:

$$m' = m, \quad u' = u,$$
$$\mathbf{q}' = R(A)\left(\mathbf{q} - \frac{b}{m}\mathbf{p}\right) - b\mathbf{v} + \mathbf{a}, \tag{35}$$
$$\mathbf{p}' = R(A)\mathbf{p} + m\mathbf{v}, \quad \mathbf{s}' = R(A)\mathbf{s}.$$

In this way, one finds the following physically interesting types of CES.

(I) $\{m/s/u\}$, $m \neq 0$, $s \geq 0$. These include two types according as $s > 0$ or $s = 0$. The state space is the manifold $\mathbb{R}^6(\mathbf{q}, \mathbf{p}) \times S_s^2(\mathbf{s})$, where $S_2^2(s)$ is the sphere $|\mathbf{s}| = s$. The sympletic structure is determined by the Poisson bracket relations,

$$\{q_i, q_j\} = \{p_i, p_j\} = \{s_i, q_j\} = \{s_i, p_j\} = 0,$$
$$\{q_i, p_j\} = \delta_{ij}, \quad \{s_i, s_j\} = \epsilon_{ijk} s_k. \tag{36}$$

The kinematical action of the Galilei group is generated by the functions

$$H = \mathbf{p}^2/2m + u, \quad P = \mathbf{p}, \quad K = -m\mathbf{q}, \quad \mathcal{J} = \mathbf{q} \times \mathbf{p} + \mathbf{s}. \tag{37}$$

From (35) we obtain that the evolution law is

$$\mathbf{q}(t) = \mathbf{q}(0) + t\mathbf{p}(0)/m, \quad \mathbf{p}(t) = \mathbf{p}(0), \quad \mathbf{s}(t) = \mathbf{s}(0). \tag{38}$$

It follows at once that the dynamical action of $\mathcal{G}_0$ takes the form

$$\pi = (t; \mathbf{q}(t), \mathbf{p}(t), \mathbf{s}(t)) \to \pi' = (t'; \mathbf{q}'(t'), \mathbf{p}'(t'), \mathbf{s}'(t')),$$

where

$$\mathbf{q}'(t') = R\mathbf{q}(t) + \mathbf{v}t + \mathbf{a}, \quad t' = t + b,$$
$$\mathbf{p}'(t') = R\mathbf{p}(t) + m\mathbf{v}, \quad \mathbf{s}'(t') = R\mathbf{s}(t). \tag{39}$$

It is clear that $\{m/s/u\}$ $(m \neq 0)$ describes a nonrelativistic CES with mass m, spin s, and internal energy u. We see that these systems admits a position observable which transforms correctly under the whole Galilei group. One proves easily that two CES$\{m, s, u_i\}$ $(i = 1, 2)$, which differ only in the parameter u, are equivalent realizations of $\mathcal{G}_0$. On the other hand, since the associated orbit is simply connected, all the CES locally equivalent to $\{m/s/u\}$ are in fact equivalent.

(II) $\{0/p/\pm s\}$, $p > 0$, $s \geq 0$. They are associated with the orbits on which $m = 0$, $\mathbf{k} \times \mathbf{p} = 0$, and $\mathbf{p} \neq 0$. If we define

$$\tau = -\mathbf{k} \cdot \mathbf{p}/p^2, \quad \lambda = \mathbf{j} \cdot \mathbf{p}/p, \quad \mathbf{w} = \mathbf{j} \times \mathbf{p}, \tag{40}$$

the state space is the manifold $\mathbb{R}^2(h, \tau) \times \{(\mathbf{p}; \mathbf{w}) \in \mathbb{R}^6 : |\mathbf{p}| = p, \ \mathbf{p} \cdot \mathbf{w} = 0\}$ and the kinematical action of $\mathcal{G}_0$ is given by

$$h' = h + \mathbf{v} \cdot R\mathbf{p}, \quad \mathbf{p}' = R\mathbf{p}, \quad \tau' = \tau - b, \quad \lambda' = \lambda,$$
$$\mathbf{w}' = R\mathbf{w} + [(\mathbf{a} - \tau\mathbf{v}) \times R\mathbf{p}] \times R\mathbf{p}. \tag{41}$$

This leads to the evolution law

$$h(t) = h(0), \ \mathbf{p}(t) = \mathbf{p}(0), \ \tau(t) = \tau(0) + t, \ \lambda(t) = \lambda(0),$$
$$\mathbf{w}'(t) = \mathbf{w}(0). \tag{42}$$

We may interpret $\{0/p/\pm s\}$ as a nonrelativistic CES with zero mass and helicity $\lambda = \pm s$. It does not admit a position observable because the condition $\{q_i, p_j\} = \delta_{ij}$ implies $\{\mathbf{q}, p^2\} = 2\mathbf{p} \neq 0$, which is absurd since p^2 is constant. On the other hand, we note that τ may be interpreted as the "age" of the elementary system. The isotropy subgroup under the action of $\mathcal{G}_0$ is $\{(b, \mathbf{a}, \mathbf{v}, R): b = a_3 = v_3 = 0, \ R = \mathbb{1}\}$. By its connected character we deduce that $\{0/p/\pm s\}$ do not have inequivalent locally equivalent CES.

The characterization of the QES of $\mathcal{G}_0$ is facilitated by the analysis of the quotient field[1] $D(U)$ associated with $\hat{G}_0$. In fact $D(U)$ is generated by the set $\{M, U, \mathbf{Q}, \mathbf{P}, \mathbf{S}\}$, where

$$\mathbf{Q} = -\mathbf{K}/M, \quad \mathbf{S} = \mathbf{J} - \mathbf{Q} \times \mathbf{P}, \quad U = H - \mathbf{P}^2/2M \tag{43}$$

have the physical meaning of position, spin, and internal energy, respectively. Moreover, there are three independent Casimir invariants,

$$M, \ MH - \mathbf{P}^2, \ (M\mathbf{J} + \mathbf{K} \times \mathbf{P})^2. \tag{44}$$

The QES with physical interpretation are of the form[3]

(I)′ $[m/s/u]$, $m \neq 0$, $2s$ nonnegative integer. The state space is the Hilbert space of functions $\psi = \psi(\mathbf{p})$ with $2s + 1$ components and finite norm,

$$\|\psi\|^2 \int \psi^\dagger(\mathbf{P})\psi(\mathbf{P})\, d^3\mathbf{P}, \quad \psi^\dagger\psi = \sum_{-s \leq j \leq s} \psi_j^* \psi_j.$$

The action of $\hat{\mathcal{G}}_0$ is generated by the operators

$$M = m\mathbb{1}, \quad H = \mathbf{P}^2/2m + u\mathbb{1}, \quad \mathbf{P} = \mathbf{p}, \quad \mathbf{K} = -im\nabla_p \times \mathbf{P} + \mathbf{S}^{(s)}, \tag{45}$$

where $\mathbf{S}^{(s)}$ is the $(2s+1)$-dimensional spin operator.

The elements $\mathbf{Q}$ and $\mathbf{S}$ of $D(U)$ admit a representation as operators in the Hilbert space. Of course from (43) and (45) we get

$$\mathbf{Q} = i\nabla_p, \quad \mathbf{S} = \mathbf{S}^{(s)} \tag{46}$$

which correspond to the standard position and spin observables in nonrelativistic quantum mechanics. The kinematical action of $\mathcal{G}_0$ transform $\{\mathbf{Q}, \mathbf{P}, \mathbf{S}\}$ in $\{\mathbf{Q}', \mathbf{P}', \mathbf{S}'\} = U^{-1}(g)\{\mathbf{Q}, \mathbf{P}, \mathbf{S}\}U(g)$ according to

$$\mathbf{Q}' = R\left(\mathbf{Q} - \frac{b}{m}\mathbf{P}\right) - b\mathbf{v} + \mathbf{a},$$
$$\mathbf{P}' = R\mathbf{P} + m\mathbf{v}, \quad \mathbf{S}' = R\mathbf{S}. \tag{47}$$

From this we obtain the evolution law

$$\mathbf{Q}(t) = \mathbf{Q}(0) + t\mathbf{P}(0)/m, \quad \mathbf{P}(t) = \mathbf{P}(0), \quad \mathbf{S}(t) = \mathbf{S}(0) \tag{48}$$

and the dynamical action of $\mathcal{G}_0$ over $\{\mathbf{Q}, \mathbf{P}, \mathbf{S}\}$ becomes

$$\mathbf{Q}'(t') = R\mathbf{Q}(t) + \mathbf{V}t + \mathbf{a}, \quad t' = t + b,$$
$$\mathbf{P}'(t') = R\mathbf{P}(t) + m\mathbf{v}, \quad \mathbf{S}'(t') = R\mathbf{S}(t). \tag{49}$$

Let us note that these equations are linear in $\{\mathbf{Q}, \mathbf{P}, \mathbf{S}\}$ and hence they continue to hold if we replace the operators by their expectation values.

Clearly $[m/s/u]$ $(m > 0)$ describes a nonrelativistic QES with mass m, spin s, and internal energy u. It admits a position operator which transforms correctly under the whole Galilei group. Two QES $[m/s/u_i]$ $(i = 1, 2)$, which differ only in the parameter u, turn out to be equivalent.

(II)′ $[0/p/\pm s]$, $p > 0$, $2s$ nonnegative integer. The group acts over the square integrable functions $\psi = \psi(E, \mathbf{p})$ on the cylinder $\{\mathbf{p}^2 = p^2, \text{ any } E\}$. The generators of $\hat{\mathcal{G}}_0$ are represented by:

$$M=0\quad H=E,\quad \mathbf{P}=\mathbf{p},\quad \mathbf{K}=-i\mathbf{p}\frac{\partial}{\partial E},\quad \mathbf{J}=i\nabla_p\times\mathbf{p}+\mathbf{Z},\tag{50}$$

where $\mathbf{Z}=\pm s(p_1(p+p_3)^{-1}, p_2(p+p_3)^{-1}, \mathbb{1})$. Now, the elements $\mathbf{Q},\mathbf{S}$ of $D(U)$ do not admit a representation as operators in the Hilbert space. On the other hand, the helicity observable $\Lambda\equiv\mathbf{J}\cdot\mathbf{P}/p$ is represented by $\pm s\mathbb{1}$ and the operator $\tau\equiv-\mathbf{K}\cdot\mathbf{P}/p^2$ may be interpreted as the "age" of the system.

The QES described by $[0/p/\pm s]$ has zero mass and helicity $\pm s$. We note that it does not admit of a position observable.[22]

B. Relativistic elementary systems

Let $\mathcal{G}_0$ be the Poincare group; its projective group coincides with the universal covering group. Thus $\hat{\mathcal{G}}_0$ is the set of elements (a,A), $a\in\mathbb{R}^4$, $A\in\mathrm{SL}(2,\mathbb{C})$ with the group law

$$(a_1,A_1)(a_2,A_2)=(a_1+\Lambda(A_1)a_2, A_1A_2),\tag{51}$$

where $\Lambda(A)$ denotes the Lorentz transformation associated with $A\in\mathrm{SL}(2,\mathbb{C})$. The projective homomorphism $\hat{q}$ is given by $(a,A)\to(a,\Lambda(A))$ and Ker $\hat{q}$ is the central subgroup $\{(0,\pm\mathbb{1})\}$ of $\hat{\mathcal{G}}_0$. A basis $B=\{H,P,K,\mathcal{J}\}$ of $\hat{G}_0$ is determined by the equations:

$$(a,\mathbb{1})=\exp(-a^0H+\mathbf{a}\cdot P),\quad (0,A(\mathbf{n},\theta))=\exp(\theta\mathbf{n}\cdot\mathcal{J}),\tag{52}$$
$$(0,H(\mathbf{n},\psi))=\exp(\psi\mathbf{n}\cdot K),$$

where $A(\mathbf{n},\theta)=\cos\theta/2-i\mathbf{n}\sigma\sin\theta/2$, $H(\mathbf{n},\psi)=\cosh\psi/2-\mathbf{n}\sigma\sinh\psi/2$, with $|\mathbf{n}|=1$. We have

$$[\mathcal{J}^i,K^j]=\epsilon^{ijk}K^k,\quad [\mathcal{J}^i,P^j]=\epsilon^{ijk}P^k,\quad [\mathcal{J}^i,\mathcal{J}^j]=\epsilon^{ijk}\mathcal{J}^k,\tag{53}$$
$$[K^i,K^j]=-\epsilon^{ijk}\mathcal{J}^k,\quad [K^i,P^j]=-\delta^{ij}H,\quad [K^i,H]=-P^i,$$

all others commutators are zero. Sometimes, it will be convenient to use of the covariant generators $\{P^\mu, M^{\mu\nu}=-M^{\nu\mu}\}$ defined by $P^\mu=(H,P)$, $M^{0i}=K^i$, $M^{ij}=\epsilon^{ijk}\mathcal{J}^k$. The characteristic dimensions are $n(\hat{G}_0)=4$, $s(\hat{G}_0)=2$.

The coadjoint action of $\hat{\mathcal{G}}_0$ takes a simple form when expressed in the coordinates $\{p^\mu, m^{\mu\nu}=-m^{\nu\mu}\}$ associated with the covariant generators in $\hat{G}_0$. Indeed, we get

$$p'=\Lambda(A)p,$$
$$m'^{\mu\nu}=\Lambda(A)^\mu_\lambda\Lambda(A)^\nu_\rho m^{\lambda\rho}+a^\mu\Lambda(A)^\nu_\rho p^\rho-a^\nu\Lambda(A)^\mu_\lambda p^\lambda.\tag{54}$$

The invariant functions are generated by p^2 and w^2, where w is the Pauli–Lubansky 4-vector $w^\mu = \epsilon^{\mu\nu\lambda\rho}m_{\nu\lambda}p_\rho/2$. In the coordinates $\{h,\mathbf{p},\mathbf{k},\mathbf{j}\}$ the action of the subgroup $\{(a,A(\mathbf{n},\theta)\}$ is

$$h'=h,\quad \mathbf{k}'=R(\mathbf{n},\theta)(\mathbf{k}+a^0\mathbf{p})-h\mathbf{a},$$
$$\mathbf{p}'=R(\mathbf{n},\theta)\mathbf{p},\quad \mathbf{j}'=R(\mathbf{n},\theta)\mathbf{j}+\mathbf{a}\times R(\mathbf{n},\theta)\mathbf{p}.\tag{55}$$

The classical algebraic structures suggest the coordinates[1]:

$$\mathbf{q}=-\frac{\mathbf{k}}{m}+\frac{\mathbf{p}\times\mathbf{w}}{mh(m+h)},\quad \mathbf{s}=-\frac{\mathbf{w}}{m}+\frac{(\mathbf{wp})\mathbf{p}}{mh(m+h)},\tag{56}$$

where $m=(h^2-\mathbf{p}^2)^{1/2}$. Evidently they are defined only at points with $p^2=h^2-\mathbf{p}^2>0$. From (55) we obtain that their transformation law under the subgroup $\{(a,A(\mathbf{n},\theta))\}$ is of the form

$$\mathbf{q}'=R(\mathbf{n},\theta)(\mathbf{q}-a^0\mathbf{p}/h)+\mathbf{a},\quad \mathbf{s}'=R(\mathbf{n},\theta)\mathbf{s}.\tag{57}$$

One finds the following physical CES:

(I) $\{m/s\}$, $m>0$, $s\geq 0$. The state space is $\mathbb{R}^6(\mathbf{q},\mathbf{p})\times S_s^2(\mathbf{s})$ with the symplectic structure induced by the standard Poisson bracket relations (36). The kinematical action of $\mathcal{G}_0$ is generated by

$$H=(m^2+\mathbf{p}^2)^{1/2},\ P=\mathbf{p},\quad K=-\mathbf{q}h+\frac{\mathbf{s}\times\mathbf{p}}{m+h},\tag{58}$$
$$\mathcal{J}=\mathbf{q}\times\mathbf{p}+\mathbf{s}.$$

We have from (57) that the evolution law takes the form:

$$\mathbf{q}(t)=\mathbf{q}(0)+t\mathbf{p}(0)/h,\quad \mathbf{p}(t)=\mathbf{p}(0),\quad \mathbf{s}(t)=\mathbf{s}(0).\tag{59}$$

The observable $\mathbf{q}$ transforms as a position observable under translations, but its transformation properties under pure Lorentz transformations[5] correspond to this character only when $s=0$. Thus, $\{m/s\}$ describes a relativistic CES with mass m and spin s, and it admits a position observable with the limitations quoted above. Since the state space is simply connected, there are not other CES locally equivalent to $\{m/s\}$.

(II) $\{0\pm s\}$, $s\geq 0$. These orbits correspond to points in $\hat{G}_0^*$ with $p^2=w^2=0$ and $p^0>0$. The coordinates $\{\mathbf{q},\mathbf{s}\}$ are singular. A suitable coordinate system is $\{\mathbf{x},\mathbf{p},\lambda\}$ with

$$\mathbf{x}=-\frac{\mathbf{k}}{h},\quad \lambda=\frac{\mathbf{j}\cdot\mathbf{p}}{|\mathbf{p}|}.\tag{60}$$

The identity $pw=0$ implies the existence of $s\geq 0$ such that $w=\mp sp$. Then, if follows at once $\lambda=\pm s$.

We find the following Poisson bracket relations:

$$\{j^i,x^j\}=\epsilon^{ijk}x^k,\quad \{x^i,p^j\}=\delta^{ij},\quad \{x^i,x^j\}=-\lambda\epsilon^{ijk}p^k/h^3.\tag{61}$$

Therefore, $\mathbf{x}$ behaves as a position observable under space translations and rotations. Nevertheless $\{x^i,x^j\}$ vanish only when $s=0$. Hence, only in the case $s=0$ can we consider $\mathbf{x}$ as being a position observable. Moreover if $s\neq 0$, given $\mathbf{r}$ such that $\{j^i,r^j\}=\epsilon^{ijk}r^k$ and $\{r^i,p^j\}=\delta^{ij}$, we find that $\mathbf{r}-\mathbf{x}$ is of the form $f(\mathbf{p}^2)\mathbf{p}$, and therefore that $\{r^i,r^j\}=\{x^i,x^j\}\neq 0$.

One finds that the state space is $\mathbb{R}^3(\mathbf{x})\times(\mathbb{R}^3(\mathbf{p})-\{0\})$ and the generators of the kinematic action of $\mathcal{G}_0$ become

$$H=|\mathbf{p}|,\quad P=\mathbf{p},\quad K=-|\mathbf{p}|\mathbf{x},\quad \mathcal{J}=\mathbf{x}\times\mathbf{p}\pm s\mathbf{p}/|\mathbf{p}|.\tag{62}$$

The evolution law is $\mathbf{x}(t)=\mathbf{x}(0)+t\mathbf{p}(0)/|\mathbf{P}(0)|$, $\mathbf{p}(t)=\mathbf{p}(0)$. Hence $\{0/\pm s\}$ describes a relativistic CES with zero mass and helicity $\pm s$, and it admits a position observable with the correct transformation properties under translations and rotations only when $s=0$. Moreover it is not difficult to see that the isotropy subgroup at the point $\{\mathbf{x}=0,\mathbf{p}=(0,0,1)\}$ is the connected subgroup of $\mathcal{G}_0$ generated by the Lie subalgebra $\{H-P^3, \mathcal{J}^3, \mathcal{J}^1+K^2+\lambda P^1, \mathcal{J}^2-K^1+\lambda P^2\}$. Therefore, $\{0/\pm s\}$ determines a unique CES.

In the quantum algebraic context of the Poincare Lie algebra, the suitable elements to represent the position and spin observables are given by[1]:

$$\mathbf{Q} = -\tfrac{1}{2}(\mathbf{K}H^{-1} + H^{-1}\mathbf{K}) + \frac{\mathbf{P}\times\mathbf{W}}{MH(M+H)},$$

$$\mathbf{S} = -\frac{\mathbf{W}}{M} + \frac{(\mathbf{W}\cdot\mathbf{P})\mathbf{P}}{MH(M+H)}, \tag{63}$$

where $\mathbf{W}$ is the spatial part of the Pauli–Lubansky 4-vector $W^\mu = \epsilon^{\mu\nu\lambda\rho}M_{\nu\lambda}P_\rho/2$ and M is the square-root of $P^2 = H^2 - \mathbf{P}^2$. There are two independent Casimir invariants P^2 and W^2 which verify $P^2 = M^2$, $W^2 = -M^2\mathbf{S}^2$.

We have the following relativistic QES with physical meaning:

(I)′ $[m/s]$, $m > 0$, $2s$ a nonnegative integer. The group acts over the $(2s+1)$-component wavefunctions $\psi = \psi(\mathbf{p})$ with finite norm

$$\|\psi\|^2 = \int \psi^\dagger(\mathbf{p})\psi(\mathbf{p})\frac{d^3\mathbf{p}}{p^0}, \quad \psi^\dagger\psi = \sum_{-s\le j\le s}\psi_j^*\psi_j,$$

where $p^0 = (m^2 + \mathbf{p}^2)^{1/2}$. The generators of the Poincaré group are represented by the operators[23]

$$P^\mu = p^\mu, \quad \mathbf{K} = -ip^0\nabla_p - \frac{\mathbf{p}\times\mathbf{S}^{(s)}}{m+p^0}, \quad \mathbf{J} = i\nabla_p\times\mathbf{p} + \mathbf{S}^{(s)}, \tag{64}$$

where $\nabla_p = (\partial/\partial p^1, \partial/\partial p^2, \partial/\partial p^3)$, and $\mathbf{S}^{(s)}$ is the $(2s+1)$-dimensional spin operator. From (63) and (64) it is straightforward to obtain

$$\mathbf{Q} = i\nabla_p - i\frac{\mathbf{p}}{2(p^0)^2}, \quad \mathbf{S} = \mathbf{S}^{(s)}. \tag{65}$$

We see that $\mathbf{Q}$ is precisely the Newton–Wigner position operator.[24] It transforms in the usual way under the rotations and translations, but its behavior under pure Lorentz transformations has no simple interpretation.[25] The evolution law of the observables $\{\mathbf{Q}, P^\mu, \mathbf{S}\}$ is the same as the one verified by their classical analogs. Therefore, $[m/s]$ interpreted as a relativistic QES with mass m and spin s and it admits a position observable with the limitations already quoted.

(II)′ $[0/\pm s]$, $2s$ a nonnegative integer. They verify $M = W^2 = 0$, $p^0 > 0$. This implies a relation $W = \mp sP$. The Hilbert space is the set of complex-valued functions $\psi = \psi(p)$ with support in the future lightcone $C_+ = \{p: p^2 = 0, p^0 > 0\}$ and finite norm

$$\|\psi\|^2 = \int_{C_+} |\psi(p)|^2\frac{d^3\mathbf{p}}{p^0}.$$

The infinitesimal generators are represented by[23]:

$$P^\mu = p^\mu, \quad \mathbf{K} = -ip^0\nabla_p + \mathbf{Z}_1, \quad \mathbf{J} = i\nabla_p\times\mathbf{P} + \mathbf{Z}_2, \tag{66}$$

where $\mathbf{Z}_1 = \pm s(-p^2(p^0+p^3)^{-1}, p^1(p^0+p^3)^{-1}, 0)$ and $\mathbf{Z}_2 = \pm s(p^1(p^0+p^3)^{-1}, p^2(p^0+p^3)^{-1}, \mathbb{1})$. It is clear that $\mathbf{Q}$ and $\mathbf{S}$ are not defined in these representations. If we define

$$\mathbf{X} = -\tfrac{1}{2}(\mathbf{K}H^{-1} + H^{-1}\mathbf{K}), \quad \Lambda = \frac{\mathbf{J}\cdot\mathbf{P}}{|\mathbf{P}|}, \tag{67}$$

a simple computation shows that

$$\mathbf{X} = i\nabla_p - i\frac{\mathbf{P}}{2|\mathbf{P}|^2} - \frac{\mathbf{Z}_1}{|\mathbf{P}|}, \quad \Lambda = \pm s\mathbb{1}. \tag{68}$$

It is known[22] that $[0/\pm s]$ admits a position observable only when $s = 0$. In this case $\mathbf{X}$ reduces to the operator $i\nabla_p - i\mathbf{p}/2|\mathbf{p}|^2$ which is precisely the position observable of this QES.

The generators of the representation can be written in terms of $\mathbf{p}$ and $\mathbf{X}$ according to

$$H = |\mathbf{p}|, \quad \mathbf{P} = \mathbf{p}, \quad \mathbf{K} = -\tfrac{1}{2}(|\mathbf{p}|\mathbf{X} + \mathbf{X}|\mathbf{p}|), \quad \mathbf{J} = \mathbf{X}\times\mathbf{p} \pm s\frac{\mathbf{p}}{|\mathbf{p}|}. \tag{69}$$

The evolution laws of $\mathbf{P}$ and $\mathbf{X}$ are identical to their classical analogs. Thus, $[0, \pm s]$ describes a relativistic QES with zero mass and helicity $\pm s$, and only when $s = 0$ does it admit of a position observable.

C. Elementary systems of the Weyl group

The Weyl Lie group $\mathcal{G}_0$ is the group of Poincaré transformations and dilations acting on Minkowski space–time $x = (x^0, \mathbf{x})$ according to $x' = \lambda\Lambda x + a$. Since the Lie algebra G_0 has no nontrivial infinitesimal exponents,[26] the projective group $\hat{\mathcal{G}}_0$ coincides with the universal covering group. Thus $\hat{\mathcal{G}}_0$ consists of of the elements (a, A, λ) where $a \in \mathbb{R}^4$, $A \in \mathrm{SL}(2, C)$, $\lambda > 0$, with the group law,

$$(a_1, A_1, \lambda_1)(a_2, A_2, \lambda_2) = (a_1 + \lambda_1\Lambda(A_1)a_2, A_1A_2, \lambda_1\lambda_2). \tag{70}$$

The projective homomorphism $\hat{q}: \hat{\mathcal{G}}_0$ is given by $(a, A_1\lambda) \to (a, \Lambda(A), \lambda)$, and Ker $\hat{q}$ is the central subgroup $\{(0, \pm\mathbb{1}, 1)\}$ of $\hat{\mathcal{G}}_0$. We define the basis $\{H, \mathcal{P}, \mathcal{K}, \mathcal{J}, D\}$ of $\hat{G}_0$ by means of Eq. (52) for $\{H, \mathcal{P}, \mathcal{K}, \mathcal{J}\}$ and the following equation for D,

$$(0, \mathbb{1}, \lambda) = \exp(-\log\lambda\cdot D). \tag{71}$$

The nonnull commutation relations are given by (53) with the added one $[D, P^\mu] = -P^\mu$. One finds that the characteristic dimensions of $\hat{G}_0$ are $n(\hat{G}_0) = 5$ and $s(\hat{G}_0) = 1$.

In terms of the coordinates $\{p^\mu, m^{\mu\nu} = -m^{\nu\mu}, d\}$ of $\hat{G}_0^*$, the coadjoint action reads

$$\begin{gathered} p' = \lambda^{-1}\Lambda(A)p, \quad d' = d + \lambda^{-1}a\Lambda(A)p, \\ m'^{\mu\nu} = \Lambda(A)^\mu_\lambda\Lambda(A)^\nu_\rho m^{\lambda\rho} + \lambda^{-1}(a^\mu\Lambda(A)^\nu_\rho p^\rho - a^\nu\Lambda(A)^\mu_\lambda p^\lambda). \end{gathered} \tag{72}$$

There is only one invariant function given by w^2/p^2. The analysis[1] of the quotient field $D(S)$ suggests the use of the coordinates $\{r^\mu, p^\mu, w^\mu\}$, where $r^\mu = (dp^\mu + m^{\mu\nu}p_\nu/p^2$. Indeed, the coadjoint action takes the following simple form:

$$r' = \lambda\Lambda(A)r + a, \quad p' = \lambda^{-1}\Lambda(A)p, \quad w' = \lambda^{-1}\Lambda(A)w. \tag{73}$$

We shall consider only the CES corresponding to the orbits of points $\{r^\mu, p^\mu, w^\mu\}$ in $\hat{G}_0^*$ with $p^2 > 0$ and $p^0 > 0$. They are labeled by a nonnegative number $\{s\}$ determined by the constant value of the invariant function in the form $s^2 = -w^2/p^2$. The state space is the manifold $\mathbb{R}^4(r)\times\{(p, w)\in\mathbb{R}^8: p^2 > 0, p^0 > 0, pw = 0, w^2 = -s^2p^2\}$ with the sympletic structure induced by the Poisson bracket relations,[1]

$$\begin{gathered} \{p^\mu, p^\nu\} = \{p^\mu, w^\nu\} = 0, \quad \{p^\mu, r^\nu\} = g^{\mu\nu}, \\ \{r^\mu, r^\nu\} = \epsilon^{\mu\nu\lambda\rho}p_\lambda w_\rho/(p^2)^2, \quad \{r^\mu, w^\nu\} = (w^\mu p^\nu - w^\nu p^\mu)/p^2. \end{gathered} \tag{74}$$

One may also define the spin observable $\mathbf{s}$ as given by Eq. (56). Thus the orbit becomes the manifold $\mathbb{R}^4(r)\times\{p\in\mathbb{R}^4: p^2 > 0, p^0 > 0\}\times S^2_s(\mathbf{s})$.

From (73) we see that the evolution law is

$$r^0(t)=r^0(0)-t,\quad \mathbf{r}(t)=\mathbf{r}(0),\quad p^\mu(t)=p^\mu(0),$$
$$w^\mu(t)=w^\mu(0). \tag{75}$$

We notice that $r^\mu(t)$ does not correspond to the evolution law of the spacetime observable of a free particle. It is a consequence of the manifestly covariant kinematical action of the Weyl group as given by Eq. (73). We emphasize that this description in terms of r^μ is a generalization of Aaberge's analysis[27] of free relativistic particles of spin 0. Let us see that a convenient position observable is provided by

$$\mathbf{x}=\mathbf{r}-\frac{\mathbf{p}}{p^0}r^0. \tag{76}$$

In fact from (75) we easily find that its evolution law is $\mathbf{x}(t)=\mathbf{x}(0)+t\mathbf{p}/p^0$, which coincides with the expected one for a free-particle position observable. Under the dynamical action of one element $g=(a,\Lambda,\lambda)$ of $\hat{\mathcal{G}}_0$, every trajectory $x(t)\equiv(t,\mathbf{x}(t))$ is mapped into another one $x'(t)=(t,\mathbf{x}'(t))$. Now, we want to prove that $x(t)$ transforms like a space–time curve under the action of the Weyl group. Since $x(t)=r(0)-r^0(t)\,p/p^0$ we have

$$x'(t')=r'(0)-\frac{r'^0(t')}{p'^0}=\lambda\Lambda r(0)+a-\frac{r'^0(t')}{p'^0}\lambda^{-1}\Lambda p.$$

On the other hand, from (75) and (76) we have

$$r'^0(t')=r'^0(0)-t'=\lambda\left(\Lambda_0^0 r^0(t)+\Lambda_i^0\frac{p^i}{p^0}r^0(t)\right)=\lambda^2\frac{p'^0}{p^0}r^0(t).$$

Hence, we get

$$x'(t')=\lambda\Lambda r(0)+a-\lambda\frac{r^0(t)}{p^0}\Lambda p=\lambda\Lambda x(t)+a, \tag{77}$$

which proves the assertion. Therefore, $\mathbf{x}$ may be interpreted as a position observable which transforms under the Weyl group as we should expect. However, the components of $\mathbf{x}$ verify the standard relations $\{x^i,x^j\}=0$ only when $s=0$.

We see that $\{s\}$ describes a CES with spin s and arbitrary positive mass. It admits a position observable with the quoted properties. There are no other CES locally equivalent to $\{s\}$, since their associated orbit is simply connected.

The generic QES of the Weyl group are denoted by $[s]$ where $2s$ is a nonnegative integer. The Hilbert space is given by the $(2s+1)$-component functions $\psi=\psi\ (p)$ with support the future inner light cone $\Omega^+=\{p: p^2>0, p^0>0\}$ and finite norm

$$\|\psi\|^2=\int_{\Omega^+}\psi^\dagger(p)\,\psi(p)\,d^4p,\quad \psi^\dagger\psi=\sum_{-s\le j\le s}\psi_j^*\psi_j. \tag{78}$$

The action of the projective group is of the form[26]

$$[U(a,A,\lambda)\psi](p)=\lambda^2 e^{ipa}D^{(s)}[A(p)^{-1}AA(\Lambda(A)_p^{-1})]\psi[\lambda\Lambda(A)_p^{-1}], \tag{79}$$

where $A(p)\equiv(m+p^0+\boldsymbol{\sigma}\cdot\mathbf{p})[2m(p^0+m)]^{-1/2}$ and $D^{(s)}$ the $(2s+1)$-dimensional irreducible representation of SU(2). The Lie algebra is represented by the operators[26]:

$$P^\mu=p^\mu,\quad \mathbf{K}=-i\mathbf{p}\frac{\partial}{\partial p^0}-ip^0\nabla_p-\frac{\mathbf{p}\times\mathbf{S}^{(s)}}{m+p^0},$$
$$\mathbf{J}=i\nabla_p\times\mathbf{p}+\mathbf{S}^{(s)},\quad D=-i\ p^\mu\left(\frac{\partial}{\partial p^\mu}+2\right), \tag{80}$$

where $m=(p^2)^{1/2}$ is the mass variable with range $0<m<+\infty$.

From the analysis of the quotient field[1] $D(U)$ and Eq. (76) we have that the candidate to describe the position observable is

$$\mathbf{X}=\mathbf{R}-\frac{1}{2}\left[\frac{\mathbf{P}}{P^0},\ R^0\right]_+ \tag{81}$$

where $[\ ,\ ]_+$ denotes the anticommutator and R is defined by

$$R^\mu=\frac{1}{2}\left[D,\frac{p^\mu}{p^2}\right]_+ +\frac{1}{2p^2}[M^{\mu\nu},p_\nu]_+. \tag{82}$$

From (80) we find

$$R^0=-i\frac{\partial}{\partial p^0},\quad \mathbf{R}=i\nabla_p+\frac{\mathbf{S}^{(s)}\times\mathbf{p}}{m(m+p^0)}. \tag{83}$$

Therefore the position observable is represented by the operator

$$\mathbf{X}=i\nabla_p+i\frac{\mathbf{p}}{p^0}\frac{\partial}{\partial p^0}-i\frac{\mathbf{p}}{2(p^0)^2}+\frac{\mathbf{S}^{(s)}\times\mathbf{p}}{m(m+p^0)}. \tag{84}$$

Let us note that in the mass representation $\psi(p^0,\mathbf{p})\to\phi(m,\mathbf{p})\equiv\psi((p^2)^{1/2},\mathbf{p})$, it becomes

$$\mathbf{X}=i\nabla_p-i\frac{\mathbf{p}}{2(p^0)}+\frac{\mathbf{S}^{(s)}\times\mathbf{p}}{m(m+p^0)}, \tag{85}$$

which reduces to the Newton–Wigner operator when the spin vanishes. Moreover, we see that only in this case the components of $\mathbf{X}$ are compatible observables.

The evolution law of the operators $\{R^\mu, P^\mu, W^\mu\}$ is given by

$$R^0(t)=R^0(0)-t,\quad \mathbf{R}(t)=\mathbf{R}(0),\quad P^\mu(t)=P^\mu(0),\quad W^\mu(t)=W^\mu(0) \tag{86}$$

Hence we get

$$\mathbf{X}(t)=\mathbf{X}(0)+t\mathbf{P}/P^0. \tag{87}$$

The transformation properties of $\mathbf{X}$ under the dynamical action of the Weyl group are the expected ones as far as dilatations, translations and rotations are concerned. But its behavior under pure Lorentz transformations has no simple interpretation. Therefore, $[s]$ describes a QES with spin s and arbitrary mass and it admits a position observable with the above limitations.

If we define the observables $\tau=-r^0$ and $\tau=-R^0$. From (75) and (86) we see that their time evolutions are

$$\tau(t)=\tau(0)+t,\quad \tau(t)=\tau(0)+t \tag{88}$$

They may be interpreted as the "age" observables of the elementary system $\{s\}$ and $[s]$, respectively. We note that since $0<p^0<\infty$, the operator $\tau=i\,\partial/\partial p^0$ is symmetric but is not a self-adjoint operator.

[1] L. Martínez Alonso, J. Math. Phys. 18, 1577 (1977).
[2] E. P. Wigner, Ann. Math. (N.Y.) 40, 149 (1939).
[3] J. M. Lévy–Léblond, "Galilei Group and Galilean Invariance" in *Group Theory and its Applications*, edited by E. M. Loebel (Academic, New York, 1971), Vol. II.

[4]J. M. Souriau, *Structure des Systemes Dynamiques* (Dunod, Paris, 1970).
[5]E. C. G. Sudarshan and N. Mukunda, *Classical Dynamics: A modern perspective* (Wiley, New York, 1974).
[6]M. Paurí and G. M. Prosperi, J. Math. Phys. **7**, 366 (1966); **8**, 2256 (1967); **9**, 1146 (1968), **16**, 1503 (1975).
[7]This group has been named the "projective covering group" of $\mathcal{G}_0$ by J. Cariñena and M. Santander in J. Math. Phys. **16**, 1416 (1975). This denomination is not appropriated since in general $\hat{\mathcal{G}}_0$ is not a covering manifold of $\mathcal{G}_0$.
[8]V. Bargmann, Ann. Math. **59**, 1 (1954); see also V. S. Varadarajan, *Geometry of Quantum Theory* (Van Nostrand, New York, 1970), Vol. II.
[9]D. J. Simms and N. M. J. Woodhouse, Lecture Notes in Physics **53** (Springer, Berlin, 1976).
[10]L. Martínez Alonso, J. Math. Phys. **17**, 1177 (1976).
[11]A. Kirillov, *Eléments de la Théorie des Représentations* (Editions Mir, Moscou, 1974).
[12]B. Kostant, Lecture Notes in Mathematics **170** (Springer, Berlin, 1970).
[13]R. Abraham, *Foundations of Mechanics* (Benjamin, New York, 1967).
[14]G. Hochschild, Ann. Math. (N.Y.) **54**, 96 (1951).
[15]G. Warner, *Harmonic Analysis on Semi-Simple Lie Groups* (Springer, Berlin, 1972), Vol. I.
[16]M. Verge, Bull. Soc. Math. France **100**, 301 (1972).
[17]F. Brickell and R. S. Clark, *Differentiable Manifolds* (Van Nostrand, London, 1970); see also P. Tondeur, Lecture Notes in Mathematics **7** (Springer, Berlin, 1969).
[18]L. Abellanas and L. Martínez Alonso, J. Math. Phys. **16**, 1580 (1975).
[19]C. Chevalley, *Theory of Lie Groups* (Princeton, U.P., Princeton, New Jersey, 1946), Vol. 1.
[20]S. Kobayashi and K. Nomizu, *Foundations of Differential Geometry* (Interscience, New York, 1963), Vol. I.
[21]This is an important concept in the study of Lie group action; see for example L. Michel, "Nonlinear Group Action," in *Statistical Mechanics and Field Theory*, edited by R. N. Sen and C.Weyl (Israel Universities Press, 1972).
[22]A. S. Whigtman, Rev. Mod. Phys. **34**, 845 (1962).
[23]J. C. Guillot and J. L. Petit, Helv. Phys. Acta **39**, 300 (1966).
[24]T. D. Newton and E. P. Wigner, Rev. Mod. Phys. **21**, 400 (1949).
[25]D. G. Currie, T. F. Jordan, and E. C. G. Sudarshan, Rev. Mod. Phys. **35**, 350 (1963).
[26]D. J. Almond, Ann. Inst. Henri Poincaré XIX, 105 (1973).
[27]T. Aaberge, Helv. Phys. Acta **50**, 331 (1977).

Exotic spinors in superconductivity

Herbert Rainer Petry

Institut für Theoretische Kernphysik, Universität Bonn, Nussallee 14-16, D-5300 Bonn, West Germany
(Received 7 June 1977; revised manuscript received 30 December 1977)

The existence of inequivalent types of spinors in spaces which are not simply connected is mathematically investigated. The mathematical results provide a purely geometrical explanation of the charge dependence of quantized flux and Josephson current in superconductivity.

INTRODUCTION

When space–time is not simply connected, it is impossible to give a unique definition of spinors. There exist several inequivalent possibilities which are in one-to-one correspondence with a certain cohomology group of space–time (which we describe later on). If this cohomology group is nontrivial, we call the spinors *exotic*. In the first seven sections the mathematical aspects of exotic spinors are discussed. We concentrate mainly on Dirac spinors which represent physically the most interesting object. Beginning with Sec. 8, we give an application of the theory to superconducting rings. The geometrical aspects are particularly simple in this case: Only two types of spinors are possible. We then construct a model in which both types of spinors are used in the quantum theory of the superconducting ring. The functional form of the charge dependence of quantized magnetic flux and Josephson current is derived in a new way and it agrees with experiment. We test our assumptions by substituting electron pairs for exotic spinors and rederive the aforementioned quantities. The results are identical for both cases.

The first sections of this article are purely mathematical. They require some basic knowledge of modern differential geometry and cohomology (in particular the Čech definition). The reader who is not familiar with these subjects but interested in the application to super conductivity, may take Propositions III and IV of Sec. 7 for granted. The rest of the paper is then self-contained.

1. DEFINITION OF SPIN STRUCTURES

Suppose that we are given a four-dimensional, oriented manifold M with a Lorentz metric g (signature $+---$). We shall assume that M is time-oriented, i.e., there is a timelike vector field t on M. If y_x is a timelike vector at $x \in M$, then y_x is said to be positive if $g(y_x, t_x)$ is positive. In addition, we make the technical assumption that M has a *simple* covering by open sets U_α, which means that every intersection of the U_α is contractible. Let ξ denote the bundle of oriented and time-oriented orthonormal frames in M. ξ is a principal bundle with the special Lorentz group L_+ as structure group. L_+ has a universal covering called spin. The covering projection ρ: spin $\to L_+$ has a kernel K_0 isomorphic to $\mathbb{Z}_2$. K_0 lies in the center of spin. Now we want to define spinors. By definition, a *spin structure* consists of a principal bundle $\tilde{\xi}$ with group spin together with a bundle map η such that the diagram

$$\begin{tikzcd} \tilde{\xi} \arrow[rr, "\eta"] \arrow[dr, "\tilde{\pi}"'] & & \xi \arrow[dl, "\pi"] \\ & M & \end{tikzcd}$$

commutes. ($\pi, \tilde{\pi}$ are the canonical projections of the bundles ξ and $\tilde{\xi}$). Furthermore, η has to satisfy

$$\eta(zg) = \eta(z)\rho(g),$$

for all $z \in \tilde{\xi}$ and $g \in$ spin. Assume that $\tilde{\xi}$ exists. Let d be a representation of spin in $\mathbb{C}^m$.

A spinor field of type d is by definition a section in the associated vector bundle $\tilde{\xi} x_d \mathbb{C}^m$ (see Sec. 3). Thus the problem of defining spinors is reduced to the construction of $\tilde{\xi}$.

Remark: The definition of a spin structure has already been used in mathematics[1] and general relativity.[2,3] The definition of a spinor field of type d contains the usual tensor fields which are obtained for special representations d (see Sec. 3). Similar definitions are also used in modern gauge theories[4] which deal, of course, with different structure groups, e.g., SU(2).

2. MORE ABOUT SPIN

We need a more detailed definition of spin, K_0, and ρ. Following Atiyah *et al.*,[1] let E denote the ordinary Minkowski vector space with constant metric of signature $+---$. Consider $C(E)$, the corresponding Clifford algebra, with identity e. E is canonically included in $C(E)$. Let ω denote the canonical antiautomorphism of $C(E)$ which leaves every $x \in E$ fixed. By definition, spin is the connected component of the group of invertible elements $g \in C(E)$, which satisfy

(a) $gxg^{-1} \in E$ for $x \in E$,

(b) $g \cdot \omega(g) = e$.

[The group spin is isomorphic to $SL(\mathbb{C}, 2)$.]

The map ρ: spin $\to \mathrm{Gl}(E)$, defined by $\rho(g)(x) = gxg^{-1}$, for all $x \in E$, is shown to be a homomorphism of spin onto L_+ with kernel K_0 consisting of the elements e and $-e$. Denote by Lie $(L_+) \subset \mathrm{Gl}(E)$ and Lie (spin) $\subset C(E)$ the Lie algebras of L_+ and spin, respectively. ρ establishes an isomorphism ρ_0,

$$\rho_0 : \text{Lie (spin)} \to \text{Lie } (L_+),$$

such that

$$\rho_0(\alpha)x = \alpha x - x\alpha$$

for all $x \in E$ and $\alpha \in$ Lie (spin).

$C(E)$ has a faithful irreducible representation γ in $\mathbb{C}^4$.

 0022-2488/79/020231-10$01.00

Restricting γ to spin yields a representation of spin which we call the *Dirac representation.* Without proof we state[2]

Proposition I: There is a Hermitian metric $\tilde{\lambda}$ of signature $++--$ in $\mathbb{C}^4$ such that $\gamma(x)$ is self-adjoint for all $x \in E$. $\tilde{\lambda}(a, \gamma(x)a)$ is positive for all positive timelike x and arbitrary $a \in \mathbb{C}^4$ $(a \neq 0)$. Spin leaves $\tilde{\lambda}$ invariant and $\gamma(\alpha)$ is skew-adjoint for $\alpha \in$ Lie (spin).

3. NONUNIQUENESS OF SPIN STRUCTURES

It has been shown[5] that a spin structure $\tilde{\xi}$ exists if and only if the second Stiefel–Whitney class of M vanishes. But $\tilde{\xi}$ is not uniquely determined (up to trivial bundle isomorphisms), if $H^1(M, K_0)$ is nontrivial[6]. Assume that this is the case. If $\tilde{\xi}$ is a spin structure, choose a simple covering U_α of M and a system of local sections $\tilde{\sigma}_\alpha$: $U_\alpha \to \tilde{\xi}$ (which exist because U_α is contractible[7]). In $U_{\alpha\beta} = U_\alpha \cap U_\beta$, $\tilde{\sigma}_\beta(x)$ must satisfy

$$\tilde{\sigma}_\beta(x) = \tilde{\sigma}_\alpha(x)\tilde{\phi}_{\alpha\beta}(x)$$

for some function $\tilde{\phi}_{\alpha\beta}$: $U_{\alpha\beta} \to$ spin, called the transition map. In $U_{\alpha\beta\gamma} = U_\alpha \cap U_\beta \cap U_\gamma$ we have the identity

$$\tilde{\phi}_{\alpha\beta}(x)\tilde{\phi}_{\beta\gamma}(x) = \tilde{\phi}_{\alpha\gamma}(x). \tag{3.1}$$

When $k \in H^1(M, K_0)$ is nontrivial and is represented by the cocycle $K_{\alpha\beta}$: $U_{\alpha\beta} \to K_0$, set (for $x \in U_{\alpha\beta}$)

$$\tilde{\phi}'_{\alpha\beta}(x) = \tilde{\phi}_{\alpha\beta}(x)K_{\alpha\beta}. \tag{3.2}$$

The functions $\tilde{\phi}'_{\alpha\beta}$ also satisfy (3.1) and it has been shown[6,8] that they determine a new inequivalent spin structure $\tilde{\xi}(K)$ with a commuting diagram

$$\begin{tikzcd} \tilde{\xi}(K) \arrow[rr, "\eta(K)"] \arrow[dr, "\tilde{\pi}(K)"'] & & \xi \arrow[dl, "\pi"] \\ & M & \end{tikzcd}$$

and a system of local sections $\tilde{\sigma}_\alpha(K)$ with

$$\eta \circ \tilde{\sigma}_\alpha = \eta(K) \circ \tilde{\sigma}_\alpha(K). \tag{3.3}$$

The transition maps are given by (3.2). It can be shown[6,8] that this construction establishes a one-to-one correspondence between the inequivalent spin structures and the elements of $H^1(M, K_0)$.

Let d be a representation of spin in $\mathbb{C}^m$. Recall the definition of $\tilde{\xi} x_d \mathbb{C}^m$.[9] An equivalence relation $\simeq$ holds in $\tilde{\xi} x \mathbb{C}^m$:

$$(z, a) \simeq (z', a') \longleftrightarrow z' = zg \text{ and } a' = d(g^{-1})a$$

for some $g \in$ spin. Dividing out yields $\tilde{\xi} x_d \mathbb{C}^m$ as the set of equivalence classes. Denote by d_m: $\tilde{\xi} x \mathbb{C}^m \to \tilde{\xi} x_d \mathbb{C}^m$ the map which assigns to each pair its equivalence class.

Suppose that d maps K_0 onto the identity. Then d determines a representation $\bar{d}$ of L_+ in $\mathbb{C}^m$ in the following obvious way: Set $\bar{d}(g') = d(g)$ for some g with $\rho(g) = g'$. This makes sense because the definition is independent of the particular choice of g'. One can then easily show that $\tilde{\xi} x_d \mathbb{C}^m$ is isomorphic to $\xi x_{\bar{d}} \mathbb{C}^m$. A trivial consequence is that any spin structure leads to identical d-type spinors, when d has the aforementioned property, which is known to be valid for bosons. The nonuniqueness of spin structures can, therefore, only affect the properties of fermions, a result which is, of course, expected.

4. CONNECTION FORM AND COVARIANT DERIVATIVE

In this section we follow the standard differential-geometrical constructions as they are described in the books of Greub *et al.*[9] or Kobayashi and Nomizu.[10]

Consider the original frame bundle ξ. A point $z \in \xi$ in the fibre over $x \in M$ can be visualized as an orientation preserving isometry of E to T_x, the tangent space over x. *The fundamental form* θ of ξ with values in E is defined by [π^*: $T_z \to T_x$ $(z \in \xi)$ is the mapping induced by π in the tangent spaces]

$$\theta_z(t_z) = z^{-1}(\pi^*(t_z)),$$

for all $z \in \xi$ and $t_z \in T_z$. *The Riemannian connection form* ω of ξ is uniquely fixed by requiring vanishing torsion,[10]

$$\delta\theta + \omega \wedge \theta = 0.$$

From now on we use δ to denote the exterior derivative. If $\tilde{\xi}$ is a spin structure, then

$$\tilde{\omega} = \rho_0^{-1}(\eta_* \omega) \tag{4.1}$$

is a connection form in $\tilde{\xi}$. (*Notation:* If ϕ is any mapping of manifolds, ϕ_* denotes the induced pullback of forms.)

Let U_α be a simple covering of M and $\tilde{\sigma}_\alpha$ a system of local sections in $\tilde{\xi}$ with transition maps $\tilde{\phi}_{\alpha\beta}$. Then $\sigma_\alpha = \eta \circ \tilde{\sigma}_\alpha$ is a system of local sections in ξ with transition maps $\phi_{\alpha\beta}(x) = \rho[\tilde{\phi}_{\alpha\beta}(x)]$. The *local connection forms* $\tilde{\omega}_\alpha$ and ω_α defined in U_α for $\tilde{\xi}$ and ξ are given by $\tilde{\omega}_\alpha = \tilde{\sigma}_{\alpha*}(\tilde{\omega})$ and $\omega_\alpha = \sigma_{\alpha*}(\omega)$, respectively. From (4.1) it follows that

$$\tilde{\omega}_\alpha = \rho_0^{-1}(\omega_\alpha). \tag{4.2}$$

In $U_{\alpha\beta}$ we have the well-known transition law,[10]

$$\begin{aligned} \tilde{\omega}_\beta(x) &= \tilde{\phi}_{\alpha\beta}(x)\tilde{\omega}_\alpha(x)\tilde{\phi}_{\alpha\beta}(x)^{-1} + \tilde{\phi}_{\alpha\beta}(x)\cdot\delta\tilde{\phi}^{-1}_{\alpha\beta}(x), \\ \omega_\beta(x) &= \phi_{\alpha\beta}(x)\omega_\alpha(x)\phi_{\alpha\beta}(x)^{-1} + \phi_{\alpha\beta}(x)\cdot\delta\phi^{-1}_{\alpha\beta}(x). \end{aligned} \tag{4.3}$$

Recall now that a d-type spinor is a section ψ in $\tilde{\xi} x_d \mathbb{C}^m$. Let ψ_α: $U_\alpha \to \mathbb{C}^m$ denote the unique function which satisfies

$$d_m(\tilde{\sigma}_\alpha(x), \psi_\alpha(x)) = \psi(x).$$

In $U_{\alpha\beta}$ we have the transition law

$$\psi_\beta(x) = d(\tilde{\phi}_{\alpha\beta}(x))\psi_\alpha(x). \tag{4.4}$$

Recall that, conversely, any system of functions satisfying (4.4) determines a section. The functions ψ_α are called the *local components* of ψ with respect to σ_α. Let Y be a vector field on M. By setting

$$(\nabla_Y \psi)_\alpha(x) := i(Y)(\delta\psi_\alpha + d(\omega_\alpha)\psi_\alpha)(x) \tag{4.5}$$

[$i(Y)$ applied to a form means evaluation of the form by Y], we obtain, by using (4.3), a new system of functions satisfying (4.4) and, therefore, a new section $\nabla_Y \psi$, called the *covariant derivative* of ψ with respect to Y.

We stop here for a physical comment. Assume that one wants to describe the quantum-mechanical motion of particles in curved space by a wave equation, such that the gravitational interaction is mediated by purely geometrical entities like the covariant derivative. The formalism described above provides us with a construction principle:

(a) Specify the spin of the particle, i.e., specify the representation d.

(b) Construct $\tilde{\xi} x_d \mathbb{C}^m$. The "wavefunction" is a section in this vector bundle.

(c) Then there is a unique and natural definition of the covariant derivative which depends only on the properties of the underlying Lorentz manifold.

An analogous construction principle is used in the differential-geometrical approach to gauge theories,[4] in particular in electrodynamics. (Later on we will couple particles to an electromagnetic vector potential A which can also be interpreted as a part of a covariant derivative.)

If (c) is not observed, then we are forced to define the covariant derivative for each representation d separately. In physical terms this would be equivalent to the introduction of new fields of spin-dependent forces for which we do not have experimental evidence.

5. DIRAC SPINORS

Let γ and $\tilde{\lambda}$ be given as in Sec. 2. A *Dirac spinor* is by definition a section in $\tilde{\xi} x_d \mathbb{C}^4$. Define the *local Dirac forms* γ_α by

$$\gamma_\alpha(x) = \gamma((\sigma_{\alpha *}\theta)(x)). \tag{5.1}$$

One finds the transition law (valid in $U_{\alpha\beta}$),

$$\gamma_\beta(x) = \tilde{\psi}_{\alpha\beta}(x)\gamma_\alpha(x)\tilde{\psi}^{-1}_{\alpha\beta}(x), \quad (\tilde{\psi}_{\alpha\beta} = \gamma \circ \tilde{\phi}_{\alpha\beta}). \tag{5.2}$$

Now, if Y is a vector field on M and ψ a section in $\tilde{\xi} x_d \mathbb{C}^4$ with local components ψ_α, then $\gamma(Y)\psi$ is defined as the section with local components given by $(i(Y)\gamma_\alpha)\cdot \psi_\alpha(x)$. (5.2) and (4.4) ensure that this definition makes sense. For $x \in M$, let e_i $(i = 1, \dots, 4)$ be a set of linearly independent vector fields defined in an open set containing x. Denote by $g^{ij}(x)$ the matrix inverse to $g(e_i, e_j)(x)$.

The *Dirac operator* is defined by

$$(D\psi)(x) = \sum_{i,j=1}^{4} g^{ij}(x)\,(\gamma(e_i)\nabla_{e_j}\psi)(x). \tag{5.3}$$

The right side is independent of the particular choice of e_i.

Next, let ψ and ψ' be two sections with local components ψ_α and ψ'_α, respectively. Using the Proposition I we find that (for $x \in U_{\alpha\beta}$)

$$\tilde{\lambda}(\psi_\alpha(x), \psi'_\alpha(x)) = \tilde{\lambda}(\psi_\beta(x), \psi'_\beta(x)). \tag{5.4}$$

Thus we can uniquely define a *spinor metric* $\langle\, ,\rangle$ by setting

$$\langle \psi, \psi' \rangle(x) = \tilde{\lambda}(\psi_\alpha(x), \psi'_\alpha(x)),$$

for $x \in U_\alpha$. Two sections ψ, ψ' define a complex-valued 1-form $j(\psi, \psi')$ by

$$i(Y)\, j(\psi, \psi') = \langle \psi, \gamma(Y)\psi' \rangle, \tag{5.5}$$

for all vector fields Y on M.

The wavefunction of a particle of spin 1/2 and mass m is by definition a section ψ which satisfies the *Dirac equation*

$$iD\psi = m\psi. \tag{5.6}$$

If the particle is, for example, an electron with charge e coupled to an electromagnetic field with vector potential A, then (5.6) is replaced by the equation

$$iD\psi + e\gamma(A)\psi = m\psi. \tag{5.7}$$

If ψ and ψ' are solutions of (5.6) or (5.7), then the divergence of $j(\psi, \psi')$ vanishes.

6. EXOTIC SPINORS

Let $\tilde{\xi}$ be a spin structure, U_α a simple covering of M, and $\tilde{\sigma}_\alpha$ a system of local sections defined on U_α. Let $k \in H^1(M, K_0)$ be nontrivial, represented by the cocycle $k_{\alpha\beta}$: $U_{\alpha\beta} \to K_0$. According to Sec. 3, there is a second spin structure $\tilde{\xi}(k)$ with a system of local sections $\tilde{\sigma}_\alpha(k)$ such that

$$\sigma_\alpha = \eta(k) \circ \tilde{\sigma}_\alpha(k). \tag{6.1}$$

Furthermore, the transition maps satisfy

$$\tilde{\phi}^k_{\alpha\beta} = \tilde{\phi}_{\alpha\beta} \cdot k_{\alpha\beta}. \tag{6.2}$$

Now $k_{\alpha\beta}(x)$ is equal to $\pm e$ and, therefore, $\gamma(k_{\alpha\beta}) = \pm 1$, because γ is a faithful representation of $C(E)$. In order to simplify the discussion, we assume that there is a set of functions λ_α: $U_\alpha \to \mathbb{C}$, for each α, such that $|\lambda_\alpha| = 1$ and

$$\lambda_\alpha / \lambda_\beta = \gamma(k_{\alpha\beta}) \tag{6.3}$$

in $U_{\alpha\beta}$. [Such functions always exist[11] if $H^1(M, Z)$ has no torsion.[12]] It follows that $\lambda^2_\alpha = \lambda^2_\beta$ in $U_{\alpha\beta}$. Therefore, the local functions λ_α define a unique unimodular function λ: $M \to \mathbb{C}$,

$$\lambda(x) = \lambda^2_\alpha(x), \tag{6.4}$$

for $x \in U_\alpha$. We shall say that $k_{\alpha\beta}$ is generated by a system of local square roots λ_α of λ.

We now consider $\tilde{\xi}(k) x_d \mathbb{C}^4$ and endow it with a covariant derivative ∇^k and a spinor metric $\langle\, ,\rangle_k$, according to the canonical construction described earlier. Sections in $\tilde{\xi}(k) x_d \mathbb{C}^4$ will be indexed by k.

Using (6.1) one easily finds that the local connection forms and the local Dirac forms coincide for $\tilde{\sigma}_\alpha$ and $\tilde{\sigma}_\alpha(k)$. Using (6.2) for the local components ψ^k_α of a section ψ^k with respect to $\tilde{\sigma}_\alpha(k)$, we find the transition law

$$\psi^k_\beta(x) = \gamma(\tilde{\phi}_{\alpha\beta}(x))\gamma(k_{\alpha\beta})\psi^k_\alpha(x) \; (x \in U_{\alpha\beta}). \tag{6.5}$$

Then it follows from (6.3) that $\lambda_\alpha \cdot \psi_\alpha$ transforms as the local component of a section in $\tilde{\xi} x_d \mathbb{C}^4$. Consequently, we have found a bundle map T: $\tilde{\xi}(k) x_d \mathbb{C}^4 \to \tilde{\xi} x_d \mathbb{C}^4$ such that $T\psi^k_\alpha = \lambda_\alpha \psi^k_\alpha$ holds for the local components of each section. Using the equality of the local Dirac forms and the invariance of $\tilde{\lambda}$, we find immediately

$$\langle \psi^k, \psi'^k \rangle_k = \langle T\psi^k, T\psi'^k \rangle, \tag{6.6}$$

$$j(\psi^k, \psi'^k) = j(T\psi^k, T\psi'^k). \tag{6.7}$$

The only significant change appears in the covariant derivative; using the equality of the local connection forms and (6.4), one finds the equation

$$\nabla_Y T\psi^k = T\nabla^k_Y \psi^k + \tfrac{1}{2}[i(Y)\lambda^{-1}\,\delta\lambda]T\psi^k, \qquad (6.8)$$

which is valid for all sections ψ^k and all vector fields Y on M.

The most general bundle map between our two bundles can always be written as $\tilde{T}\circ T$, where $\tilde{T}$ is a bundle isomorphism of $\tilde{\xi}x_d\mathbb{C}^4$ into itself. Assume that:

(a) (6.6) and (6.7) hold with T replaced by $\tilde{T}\circ T$,

$$\langle\psi^k, \nabla^k_Y\psi'^k\rangle_k = \langle\tilde{T}\circ T\psi^k, \nabla_Y\tilde{T}\circ T\psi'^k\rangle, \qquad (6.9)$$

i.e., in quantum-mechanical terms, we require the physically important matrix elements to be invariant under the action of $\tilde{T}\circ T$.

If such a bundle map would exist, we would obviously be forced to regard both types of bundles only as different mathematical descriptions of the same physical situation. If such a bundle map does not exist, we say that the two bundles are physically inequivalent.

Now (6.6) and (6.7) imply that $\tilde{T}\psi = \lambda'\psi$ where λ' is a unimodular complex function. (6.8) and (6.9) yield

$$\lambda'\delta\lambda'^{-1} = \tfrac{1}{2}\lambda\delta\lambda^{-1} \quad \text{or} \quad \delta(\lambda'^2/\lambda) = 0.$$

Since λ' and λ are unimodular, there is a constant unimodular complex number C_0 such that $\lambda = (C_0\lambda')^2$. Therefore, $\lambda_\alpha/C_0\lambda' = k_\alpha = \pm 1$, which implies that $\gamma(k_{\alpha\beta}) = k_\alpha/k_\beta$. But then $k_{\alpha\beta}$ is a coboundary and the class k vanishes, contrary to the original assumption. The two bundles are, therefore, inequivalent in the physical sense.

By the same kind of reasoning it can be shown that the 1-form $\lambda^{-1}\delta\lambda$, which is obviously purely imaginary and closed, cannot be exact: Assume that $\lambda^{-1}\delta\lambda = i\delta\alpha$ for some real function α. Then $\delta\lambda = i\lambda\delta\alpha$, which implies $\lambda = \lambda_0 \exp i\alpha$, with $\lambda_0 \in \mathbb{C}$ and $|\lambda_0| = 1$. Then $\lambda_\alpha/\lambda_0 \exp i\alpha = k_\alpha = \pm 1$, which leads to $\gamma(k_{\alpha\beta}) = k_\alpha/k_\beta$ and implies a contradiction as shown above.

It is now convenient to introduce the 1-form

$$B = \frac{1}{2\pi i}\,\lambda^{-1}\delta\lambda \qquad (6.10)$$

which is real, closed (but not exact) and, moreover, defines an integer cohomology class. We sketch the proof of the last property: Since U_α is contractible (for $x\in U_\alpha$) we can write $\lambda(x) = \exp(2\pi i\Phi_\alpha(x))$ for some functions $\Phi_\alpha: U_\alpha \to R$. Therefore, $B = \delta\Phi_\alpha$ in U_α. In $U_{\alpha\beta}$ we must have $\exp(2\pi i\Phi_\alpha) = \exp(2\pi i\Phi_\beta)$, which implies $\Phi_\beta = \Phi_\alpha + z_{\alpha\beta}$, where $z_{\alpha\beta}$ is an integer. Consequently, B defines an integer cohomology class in the Čech sense. Using the equivalence of Čech and de Rham cohomology,[9] we conclude that the integral of B taken along any closed curve yields an integer.

Let us summarize our results: Starting with a spin structure $\tilde{\xi}$ we found a first type of Dirac spinor ψ defined as a section in $\tilde{\xi}x_\gamma\mathbb{C}^4$. For a nontrivial cocycle $k_{\alpha\beta}$, generated by a system of local square roots of λ, we have found a second, inequivalent spin structure $\tilde{\xi}(k)$ and a second type of Dirac spinor ψ^k defined as a section in $\tilde{\xi}(k)x_\gamma\mathbb{C}^4$. There is a canonical definition of a covariant derivative for both bundles. Now we see from (6.7) and (6.8) that the second type of spinor can also be represented differently, namely by a section $T\psi^k = \psi'$ in the original bundle $\tilde{\xi}x_\gamma\mathbb{C}^4$. According to (6.8), the only change required is in the definition of the covariant derivative. We formulate this in the following.

Proposition II: Let $\tilde{\xi}$ be a spin structure and let the nontrivial cocycle $k_{\alpha\beta}$ be generated by a system of local square roots of λ. In addition to the Dirac spinors ψ, which are defined as sections in the bundle $\xi x_\gamma\mathbb{C}^4$ with covariant derivative ∇, we get a second type of Dirac spinors ψ' which can be described by sections in the same bundle, but with a different formula for the covariant derivative,

$$\nabla'_Y\psi' = \nabla_Y\psi' - \tfrac{1}{2}(i(Y)\lambda^{-1}\delta\lambda)\psi', \qquad (6.11)$$

valid for all sections ψ' and vector fields Y. The spinor-metric $\langle\,,\rangle$ and the map $(Y,\psi)\to\gamma(Y)\psi$ are defined identically for both kinds of spinors.

7. DISCUSSION OF A SPECIAL EXAMPLE

Let E be the Minkowski space with constant Lorentz metric g (see Sec. 2). Choose an orthonormal basic e_μ ($\mu = 1, \ldots, 4$), such that $g(e_4, e_4) = 1$, i.e., e_4 is the time axis. Fix a representation λ of $C(E)$ in $\mathbb{C}^4$ by setting $\gamma(e_\mu) = \gamma_\mu$, where γ_μ is one of the standard sets of Dirac matrices with $\gamma_4^* = \gamma_4$ and $\gamma_i^* = -\gamma_i$ ($i = 1, 2, 3$). For $a, b \in \mathbb{C}^4$, define $\tilde{\lambda}(a, b) = a^*\gamma_4 b$. We also write $\tilde{\lambda}(a, b) = \bar{a}b$. Then γ and $\tilde{\lambda}$ have the properties stated in Proposition I.

Let us describe points in E by Cartesian coordinates x^μ with respect to the basis vectors e_μ and let $E_3 \subset E$ denote the subspace orthogonal to e_4. If $x \in E$, let $\mathbf{x}$ denote the projection of $\mathbf{x}$ onto E_3, i.e., $\mathbf{x}$ is the space component of x. Let $V_3 \subset E_3$ be an open set with boundary δV_3. Consider the Lorentz manifold $V \subset E$, $V = (x \in E;\, \mathbf{x} \in V_3)$, with induced metric and orientation. The tangent space of V is obviously trivial, and so is the frame bundle ξ. We have a global section σ in ξ given by the constant vector fields e_μ. The Riemannian connection form and the fundamental form then satisfy the well-known formulas

$$\sigma_* \cdot \theta(e_\mu) = e_\mu, \quad \sigma_* \cdot \omega(e_\mu) = 0. \qquad (7.1)$$

We find a first spin structure $\tilde{\xi} = V \times$ spin with η given by the map which sends the pair (x, g) into the frame $\rho(g^{-1})e_\mu$ ($\mu = 1, \ldots, 4$) at x. Define $\tilde{\sigma}$: $V \to \xi$ by $\tilde{\sigma}(x) = (x, e)$. Obviously, we have $\eta\circ\tilde{\sigma} = \sigma$. The associated vector bundle $\tilde{\xi}\mathbf{x}_\gamma\mathbb{C}^4$ is trivial. A section is completely specified by its component ψ_σ with respect to $\tilde{\sigma}$. ψ_σ: $V \to \mathbb{C}^4$ is a globally defined function. Using (4.1), (4.5), (5.1), and (5.4) together with (7.1) we find immediately:

$$(\nabla e_\mu\psi)_\sigma = \frac{\partial}{\partial x^\mu}\,\psi_\sigma, \qquad (7.2)$$

$$(D\psi)_\sigma = \gamma_\mu\,\frac{\partial}{\partial x^\mu}\,\psi_\sigma, \qquad (7.3)$$

$$(\gamma(e_\mu)\phi)_\sigma = \gamma_\mu\psi_\sigma, \qquad (7.4)$$

$$\langle \psi, \psi' \rangle = \bar{\psi}_\sigma \psi'_\sigma. \tag{7.5}$$

Proposition III: Let $M \subset E$ be given by the set $V = (x; \mathbf{x} \in V_3)$ where V_3 denotes a bounded open set with boundary δV_3 in E_3. Assume that there is a diffeomorphism $f: V_3 \to \mathbb{R}^2 \times S^1$.

(a) There are only two inequivalent types of Dirac spinors.

(b) The first type can be described by functions $\psi: V \to \mathbb{C}^4$ such that (7.2)–(7.5) hold.

(c) The second type of Dirac spinors is described by functions $\psi': V \to \mathbb{C}^4$, which satisfy (7.4) and (7.5), while (7.2) and (7.3) hold with $\partial/\partial x^\mu$ replaced by $\partial/\partial x^\mu - \frac{1}{2}[\lambda^{-1}(\partial/\partial x^\mu)\lambda]$.

(d) The function $\lambda: V \to \mathbb{C}$ satisfies $|\lambda| = 1$. It can be chosen such that $(\partial/\partial x^4)\lambda = 0$ and $\operatorname{div} \lambda^{-1}\delta\lambda = 0$. (div denotes divergence of a form.)

(e) Let A be the vector potential of an electromagnetic field. Solutions of the following two types of Dirac equations describe particle states of spin $\frac{1}{2}$, mass m, and charge e,

$$\begin{aligned} &iD\psi + e\gamma^\mu A_\mu \psi = m\psi, \\ &iD\psi' + \gamma^\mu \left[eA_\mu - \frac{i}{2}\left(\lambda^{-1}\frac{\partial}{\partial x^\mu}\lambda\right)\right]\psi' = m\psi'. \end{aligned} \tag{7.6}$$

Proof: The existence of f implies that $H^1(V, K_0) = H^1(S^1, K_0)$. But $H^1(S^1, K_0) = K_0$, so that (a) and (b) follow from Sec. 3.

Denote points in S^1 by complex numbers with modulus one. Let $\alpha: (S^1 - 1) \to S^1$ be the square root function (which is well defined on $S^1 - 1$). Let $f': V^3 \to S^1$ be the map induced by f. Define the open sets U_1, $U_2 \subset M$ by

$$U_1 = (x \in M; f'(\mathbf{x}) \neq 1),$$

$$U_2 = (x \in M; f'(\mathbf{x}) \neq -1).$$

U_1 and U_2 form a covering of M which is not simple, but sufficient for our purpose, since we can always find a simple refinement. Let $\phi: V_3 \to \mathbb{R}$ be an arbitrary function. Define $\lambda_i: U_i \to S^1$ by

$$\begin{aligned} &\lambda_1(x) = \alpha(f'(\mathbf{x}))\exp i\Phi, \\ &\lambda_2(x) = i\alpha(-f'(\mathbf{x}))\exp i\Phi. \end{aligned} \tag{7.7}$$

Now $U_{12} = U_1 \cap U_2$ is disconnected. The imaginary part of f', $\operatorname{Im} f'$, is nonzero on each connected component of U_{12}. The function $K_{12}(x) = \lambda_1(x)/\lambda_2(x)$ is well defined in U_{12} and K_{12} is given by

$$K_{12}(x) = -\operatorname{Im} f'(x)/|\operatorname{Im} f'(x)| \tag{7.8}$$

The functions λ_i are local square roots of the globally defined function $\lambda = f' \cdot \exp 2i\Phi$. Observe that $(\partial/\partial x^4)\lambda = 0$ and $\lambda^{-1}\delta\lambda = f'^{-1}\delta f' + 2i\delta\Phi$.

By setting

$$\Phi(x) = \frac{1}{8\pi i}\int_{V_3} |\mathbf{x} - \mathbf{y}|^{-1} \operatorname{div}(f'^{-1}\delta f')(y)\, dy^1\, dy^2\, dy^3$$

we ensure that $\operatorname{div}(\lambda^{-1}\delta\lambda) = 0$. (c)–(e) then follow immediately from Proposition II.

Remark: The formulas (7.6) might suggest that ψ and ψ' are related by a simple phase factor λ' and an a appropriate gauge transformation of A such that

$$\lambda'\delta\lambda'^{-1} = \tfrac{1}{2}\lambda\delta\lambda^{-1}.$$

But it has been shown quite generally in Sec. 6 that such a function cannot exist globally. It exists *locally*, e.g., on U_1, where we can set λ' equal to λ_1. The equation for the product $\lambda_1 \cdot \Psi'$ is indeed identical with the equation for Ψ but the product function discontinuously changes sign at the surface where λ_1 is discontinuous; hence it is *not* a well-defined solution of the first differential equation.

Proposition IV: Fix $x \in \mathbb{R}^2$. Denote by $c: S^1 \to V_3$ the closed path with $c(z) = f^{-1}(x, z)$. Then

$$\int_c \lambda^{-1}\delta\lambda = 2\pi i. \tag{7.9}$$

This is also true for any closed path homotopic to c.

Proof: By construction, λ satisfies $\int_c \lambda^{-1}\delta\lambda = \int_c f'^{-1}\delta f'$. Now $f' \circ c$ is the identity of S^1. Therefore,

$$\int_c \lambda^{-1}\delta\lambda = \int_{S^1} z^{-1}\, dz = i\int_0^{2\pi} d\phi = 2\pi i.$$

The last statement of the proposition follows because $\lambda^{-1}\delta\lambda$ is closed.

8. TWO MODELS FOR SUPERCONDUCTIVITY

Definitions and notations are as in Sec. 7. Suppose that electrons and photons are confined to the manifold V by some unspecified mechanism and that V satisfies the assumption of Proposition III. We have shown that exactly two types of Dirac spinors exist. If nature is democratic and does not suppress one of them, these electrons obviously have a new degree of freedom. This means that quantum electrodynamics must use two spinor-field operators ψ and ψ' together with the photon-field operator A. The equations of motion are then supplied by (7.6):

$$\begin{aligned} &iD\psi + e\gamma^\mu A_\mu \psi = m\psi, \\ &iD\psi' + e\gamma^\mu (A - B)_\mu \psi' = m\psi', \\ &B_\mu = \frac{i}{2e}\lambda^{-1}\frac{\partial}{\partial x^\mu}\lambda. \end{aligned} \tag{8.1}$$

The current operator is given by

$$I_\mu = (\bar{\psi}\gamma_\mu\psi + \bar{\psi}'\gamma_\mu\psi') \cdot e.$$

This expression is fixed by (8.1) and the requirement that the divergence of I vanishes. A mixed contribution of ψ and ψ' to I is not compatible with both conditions. A contribution with different weights contradicts the assumption that ψ and ψ' are associated to the same charge.

A satisfies

$$\Box A = I, \quad \operatorname{div} A = 0. \tag{8.2}$$

We require that ψ, ψ' and $F = \delta A$ have support in V and that the canonical equal-time commutation relations hold.

By support we mean that suitable boundary conditions are imposed at δV_3. These are not easily formulated for Dirac spinors. In principle, (8.1) should, therefore, be reduced to the nonrelativistic Pauli equation, where the conventional boundary conditions can be applied. A nonrelativistic reduction would also be more in line with the conventional treatment of superconductivity but it is omitted here because it does not affect our main conclusions (see also Sec. 13).

Parallel to the discussion of this model, we consider quantum electrodynamics with a boson field Φ of charge $2e$ and mass m'. Φ obeys the well-known equation[13]

$$g^{\mu\nu}\left(i\frac{\partial}{\partial x^\mu}+2eA_\mu\right)\left(i\frac{\partial}{\partial x^\nu}+2eA_\nu\right)\Phi=m'^2\Phi. \quad (8.3)$$

Φ has again support in V, the equal-time commutation relations are canonical, and the current is given this time by

$$I_\mu=2e\,\Phi^+\left(i\frac{\partial}{\partial x^\mu}+2eA_\mu\right)\Phi + \text{Hermitian conjugate}.$$

The well-known difficulties with the subsidiary condition $\mathrm{div}A=0$ do not concern us here. We assume that they are solved either by the use of an indefinite Hilbert space metric or by working in the radiation gauge.

Let us define

$$\begin{aligned} &T(\psi)=\psi' \quad T(\psi')=\psi\lambda,\\ &T(\Phi)=\Phi\lambda,\ \ T(A)=A-B. \end{aligned} \quad (8.4)$$

The equations (8.1)–(8.3) are invariant under the transformation $\psi\to T(\psi)$ etc. Moreover, the equal-time commutation relations are invariant. We assume that there is an operator $\bar{T}$ in the Hilbert space H, in which the field operators act, such that $\bar{T}$ is unitary and commutes with the Hamiltonian h of the system. $\bar{T}$ satisfies $\bar{T}\psi\bar{T}^{-1}=T(\psi)$, etc.; i.e., the formal symmetry (8.4) is unitarily implemented by $\bar{T}$. It follows that

$$\bar{T}I_\mu\bar{T}^{-1}=I_\mu. \quad (8.5)$$

Let c be a path as in Proposition IV. Consider the operator

$$\phi(t,c)=2e\int_c A_i(x_4=t,\mathbf{x})\,dx^i.$$

$\phi(t,c)$ is just the operator of the flux, times $2e$, through a surface in E bounded by c.

Let q be a fixed eigenvalue of the charge operator Q. Consider the subspace P of H, which is spanned by simultaneous eigenstates of Q (with eigenvalue q) and of h, with *lowest possible* eigenvalue $E(q)$. Denote the projection operator on P by $\bar{P}$ and set

$$A^P(\mathbf{x})=\bar{P}A(x)\bar{P},\quad I^P(\mathbf{x})=\bar{P}I(x)\bar{P},\quad \phi^P(c)=\bar{P}\phi(t,c)\bar{P}.$$

$\phi^P(c)$, $A^P(\mathbf{x})$, and $I^P(\mathbf{x})$ are independent of time (i.e., of x_4), since states in P have the same energy. This justifies our notation.

We now make the assumption of

(a) *rigidity*: $\bar{P}A(1-\bar{P})=0$.

Assumption (a) implies that for every state $|\alpha\rangle\in P$

$$A_\mu(x)|\alpha\rangle=A^P_\mu(\mathbf{x})|\alpha\rangle\in P. \quad (8.6)$$

It follows from (8.2) that

$$I_\mu(x)\,|\alpha\rangle\ =I^P_\mu(\mathbf{x})|\alpha\rangle\in P. \quad (8.7)$$

Using (a) together with the equal-time commutation relations, one finds

$$[A^P_\mu(\mathbf{x}),A^P_\nu(\mathbf{x})]_-=[A^P_\mu(\mathbf{x}),I^P_\nu(\mathbf{x})]_-=0. \quad (8.8)$$

The equal-time commutator of A_μ and $(\partial/\partial x^4)A_\nu$ is nonzero. On the other hand, a short calculation shows that (8.6) implies that this commutator vanishes. Therefore, assumption (a) can only be regarded as a more or less correct approximation. It means that quantum fluctuations in the photon field are totally neglected. The field behaves, therefore, like a (state-dependent) classical function. (8.6) implies now, in particular, that

$$\phi(t,c)|\alpha\rangle\ =\phi^P(c)\ |\alpha\rangle\in P. \quad (8.9)$$

Next we make the assumption of

(b) *simplicity*: For each $\phi\in R$ there is one and only one eigenstate $|\phi\rangle\in P$ of $\phi^P(c)$ with eigenvalue ϕ. The states $|\phi\rangle$ are normalized according to $\langle\phi'|\phi\rangle=\delta(\phi'-\phi)$ and span P.

(c) We also require the *Meissner effect*, according to which

$$\frac{\partial A^P_i(\mathbf{x})}{\partial x_j}-\frac{\partial A^P_j(\mathbf{x})}{\partial x_i}=0\quad(i,j=1,2,3)$$

for $\mathbf{x}$ sufficiently far from the boundary of V_3. With the help of these assumptions we now prove

Proposition V: (1) $|\phi\rangle$ is an eigenstate of $\phi(t,c)$ with eigenvalue ϕ.

(2) If c, c' are sufficiently far from the boundary of V_3 and c' is homotopic to c, then $\phi(t,c')\,|\phi\rangle=\phi\,|\phi\rangle$.

(3) $|\phi\rangle$ is an eigenstate of $I_\mu(x)$ and $A_\mu(x)$, with time-independent eigenvalues $I_\mu(\mathbf{x},\phi)$ and $A_\mu(\mathbf{x},\phi)$, respectively.

(4) $I_\mu(\mathbf{x},\phi)=I_\mu(\mathbf{x},\phi-2\pi)$, $\quad A_\mu(\mathbf{x},\phi)=A_\mu(\mathbf{x},\phi-2\pi)+B_\mu$.

Proof: (1) follows directly from (8.9).

(2) follows from assumption (c) and Stokes' theorem. From (8.8) and (8.9) we conclude that

$$[\phi^P(c),I^P_\mu(\mathbf{x})]_-=[\phi^P(c),A^P_\mu(\mathbf{x})]_-=0. \quad (8.10)$$

Therefore, $\langle\phi'|I_\mu(\mathbf{x})|\phi\rangle=\langle\phi'|A_\mu(\mathbf{x})|\phi\rangle=0$ if $\phi'\neq\phi$. Consequently,

$$\begin{aligned} &\langle\phi'|I_\mu(\mathbf{x})\,|\,\phi\rangle=I_\mu(\mathbf{x},\phi)\delta(\phi'-\phi),\\ &\langle\phi'|A_\mu(\mathbf{x})|\phi\rangle=A_\mu(\mathbf{x},\phi)\delta(\phi'-\phi), \end{aligned} \quad (8.11)$$

for some real functions $A_\mu(\mathbf{x},\phi)$ and $I_\mu(\mathbf{x},\phi)$.

(3) follows immediately from (8.6) and assumption (b).

(4) is the consequence of the existence of the unitary operator $\bar{T}$. Because $\bar{T}$ commutes with the Hamiltonian and the charge operator, it must leave P invariant, i.e.,

$$\bar{T}^{-1}\bar{P}\bar{T}=\bar{P}.$$

(8.4) and (8.5) imply that

$$\bar{T}^{-1}I^P_\mu(\mathbf{x})\bar{T}=I^P_\mu(\mathbf{x}),\quad \bar{T}^{-1}A^P_\mu(\mathbf{x})\bar{T}=A^P_\mu(\mathbf{x})+B_\mu\bar{P}. \quad (8.12)$$

From Proposition IV it follows that

$$\bar{T}^{-1}\phi^P(c)\bar{T} = \phi^P(c) - 2\pi\bar{P}$$

and, consequently,

$$\phi^P(c)\bar{T}\,|\,\phi\rangle = \bar{T}(\phi^P(c) - 2\,\pi)\,|\,\phi\rangle = (\phi - 2\pi)\bar{T}\,|\,\phi\rangle.$$

Now $T\,|\,\phi\rangle cP$ and, therefore, we conclude, from our assumption of simplicity, that

$$\bar{T}\,|\,\phi\rangle = |\,\phi - 2\pi\rangle. \tag{8.13}$$

If we insert (8.12) in (8.11) and use (8.13), we immediately verify the last point of the proposition.

9. FURTHER DEVELOPMENTS

After this exposition of the mathematical details of our two models, we wish to establish their connection to the theory of superconductivity. We use Schrieffer's book as a reference for the work which has been done in this field.[14]

Consider a massive ring of superconducting metal. The topology of the ring is equal to the topology of V_3. The only objects which can move inside the ring are the photons and the conducting electrons. We conclude that the low-lying states in the ring can be well described by a quantum-field theory which uses only electron and photon field operators. The ions of the metal are heavy particles which remain to a high degree of accuracy at a fixed average position. Their presence serves to keep the electrons inside the metal by maintaining an average electric potential. We incorporate this in our models by requiring that the electron-field operators have support in V_3. (More sophisticated models will, of course, take lattice vibrations[14] into account.) The standard assumption of superconductivity is that electrons form spinless pairs with charge $2e$. This corresponds to our boson model. In addition, we consider a model in which electrons make use of the second type of Dirac spinors, which is provided by the nontrivial topology of the ring, as we have shown. Because $H^1(E, K_0)$ vanishes, this second type of Dirac spinor cannot exist on the whole of Minkowski space E. Strictly speaking, this second type of Dirac spinors is the result of the approximation that electrons are completely confined to the ring.

In addition, we made the assumption that photons are also confined to the ring. This is an approximation which has to be improved, because the electrons will always create an electromagnetic field outside of the ring. The assumptions of rigidity and simplicity are then purely technical means for providing the proof of Proposition V.

Rigidity can be somewhat justified by the experimental fact that low-frequency radiation is not absorbed by superconductors.[14] The assumption of the Meissner effect uses the experimental result that magnetic fields do not exist in superconductors.[14]

Now recall Proposition V: We have found that the state of lowest energy for a definite charge is degenerate to within our approximations. The eigenstates $|\,\phi\rangle$ of flux (times $2e$) provide a basis for the eigenspace of lowest energy. $|\,\phi\rangle$ is also an eigenstate of current and vector potential with eigenvalues $I_\mu(\mathbf{x}, \phi)$ and $A_\mu(\mathbf{x}, \phi)$. We have already mentioned that this description is incomplete, since we have neglected the electromagnetic field outside the ring. We now try to improve our models by extending $A_\mu(\mathbf{x}, \phi)$ to the exterior of the ring. More precisely, we postulate the existence of a field $\tilde{A}_\mu(\mathbf{x}, \phi)$ which coincides with $A_\mu(\mathbf{x}, \phi)$ inside of the ring. It is then natural to assume that at large distances from the ring: (1) $\tilde{A}_4$ vanishes, because the electrostatic potential is completely shielded by the charges of the ions, and (2) that $\tilde{A}_i(\mathbf{x}, \phi)$ $(i = 1, 2, 3)$ approximates the classical expression given by Biot–Savart's law,

$$\tilde{A}_i(\mathbf{x}, \phi) = \frac{1}{4\pi}\int_{V_3} |\mathbf{x} - \mathbf{y}|^{-} I_i(\mathbf{y}, \phi). \tag{9.1}$$

We cannot postulate such a formula in the intermediate region, but we know experimentally that the extension of this region is given by the London penetration depth, which is very small.[14]

The external field makes a contribution $\Delta E(\phi)$ to the total energy of the state $|\phi\rangle$, namely half of the square of the magnetic field integrated over the external space. If we neglect the contribution of the transition region this additional energy can be simply computed from (9.1). We know that $I_i(\mathbf{x}, \phi)$ is periodic in ϕ with period 2π. It follows that $\Delta E(\phi)$ is also periodic. The degeneracy in energy is, therefore, partially but not completely removed by $\Delta E(\phi)$. If ΔE has an absolute minimum at ϕ_0, then there are other such minima at $\phi_0 + k \cdot 2\pi$. For obvious physical reasons, we would expect among those lowest-lying states there is one which carriers no current, no magnetic field, and no flux. This means that $\Delta E(0)$ is among the minima and, consequently, the states of lowest energy are given by the states $|2\pi k\rangle$.

With the assumptions quoted above we now prove

Proposition VI: Let c be a path which lies sufficiently deep inside the ring and assume that c has the properties stated in Proposition IV. If the ring is in a state of lowest energy, then the magnetic flux through any 2-surface bounded by c is equal to $k\pi/e$ with integer k. If c' is another such path which is homotopic to c in V_3, then the flux has the same value as for c.

Remark: The last point makes it possible to speak simply of the flux through the ring.

Proof: Let the ring be in state $|\,\phi\rangle$. The magnetic flux can be directly expressed as the line integral of $\tilde{A}(\mathbf{x}, \phi)$ over c. $\tilde{A}_i(\mathbf{x}, \phi)$ equals $A_i(\mathbf{x}, \phi)$ in V_3 and the line integral is equal to $\phi/2e$ as a consequence of Proposition V, which also guarantees the same result for c' homotopic to c. As we have shown before, ϕ equals $2\pi k$ with integer k in the states of lowest energy. This proves the proposition.

10. JUNCTIONS

We now turn to the discussion of junctions. A good survey of the existing theoretical and experimental work in this field has been presented in Waldram's recent review article.[15] Experimentally, a junction is prepared by inserting a thin insulating layer in a superconductor.

We represent this mathematically by an electric potential with a discontinuity at a surface which divides the superconducting ring. The absolute value of the discontinuity can be varied with the help of an external voltage source in parallel with the junction. We first discuss the geometrical aspects of this arrangement. For the following we need some topological notions. Notation and definitions are as in Secs. 7 and 8.

A 2-surface F in V_3 is called simple if there is a diffeomorphism of $V_3 - F$ onto $\mathbb{R}^3$ and if $F \cap c$ consists of a single point. A path c' is called simply homotopic to c if it is homotopic to c in V_3 and if $c' \cap F$ consists of single point.

A Josephson junction consists of a simple surface F and a real function μ (defined on $V_3 - F$), which has a constant discontinuity μ_0 at F.

A double junction consists of two surfaces F and F' which are both simple and do not intersect, together with a real function μ (defined on $V_3 - F - F'$), which has a constant discontinuity μ_0 at F and $-\mu_0$ at F'.

Recall that V_3 can be visualized as a ring. We see that in the case of the Josephson junction, F just cuts the ring. In the case of the double junction, F and F' cut the ring into two pieces. μ_0 may be time-dependent in both cases. We call $E_i = -(\partial\mu/\partial x^i)$ $(i=1,2,3)$ the electric field of the junction. Let $B(t)$ be a magnetic field on E_3 which vanishes in V_3. (we use implicitly the Meissner effect again here.) The external flux $\Phi_{ex}(t,c)$ through V_3 is defined by

$$\Phi_{ex}(t,c) = \int_{\tilde{F}} B df,$$

where $\tilde{F}$ is any surface bounded by c.

Proposition VII: (a) $\Phi_{ex}(t,c)$ is independent of the particular choice of the surface $\tilde{F}$.

(b) $\Phi_{ex}(t,c) = \Phi_{ex}(t,c')$ if c' is homotopic to c inside V_3.

(c) For the Josephson junction

$$\int_c d\mu = \int_{c'} d\mu = \mu_0$$

if c' is simply homotopic to c with respect to F.

(d) For the double junction

$$\int_c d\mu = \int_{c'} d\mu = 0$$

if c' is simply homotopic to F and F'.

Proof: (a) follows from $\operatorname{div} B = 0$ and Stokes' theorem.

(b) follows from $B=0$ in V_3. (c) For any path c' with $c' \cap F$ equal to a point, we have $\int_{c'} d\mu = \pm\mu_0$. The sign depends on the orientation of the path. The sign is $+$, if c' is homotopic to c. This proves (c). (d) is true for every path which meets F and F' only once.

Next we consider the effect of a junction for the two models of quantum electrodynamics discussed in the last section. The total flux $\Phi_{to}(t,c)$ is the sum of the externally-applied flux and $(1/2e)\phi(t,c)$ (note that ϕ was defined as flux times $2e$). Assume that $\mu(t)$ and $B(t)$ vanish for $t<0$. Let P have the superconducting properties (a)–(c). Proposition V applies for negative time. For positive time, the system interacts with the external fields of the junction which transfer energy into it. We assume that this energy is largely removed by some unspecified cooling mechanism. Let $U(t)$ denote the time-evolution operator of the system. $U(t)$ is unitary and satisfies $U(0)=1$. If a state was in P at time less than zero, it will remain in P, since energy and charge are not transferred. In particular, $U(t)|\phi\rangle$ is contained in P for all time.

We now make the assumption of

junction dominance: The contribution of the internal electric field to the time derivative of the total flux is small compared to the contribution of the junction field.

Maxwell's equations in integral form yield immediately

$$\frac{d}{dt}\Phi_{to}(t,c) = -\int_c E_i(t)\,dx^i \tag{10.1}$$

which is easily integrated and yields

$$\phi(t,c) = \phi(0,c) - \lambda(t),$$
$$\lambda(t) = 2e\Big[\Phi_{ex}(t,c) - \Phi_{ex}(0,c) + \int_0^t \int_c E_i(\tau)\,dx^i\,d\tau\Big]. \tag{10.2}$$

With the help of the time-evolution operator we write

$$\phi(t,c) = U(t)\,\phi(0,c)\,U^+(t).$$

With the help of (10.2) we find

$$\phi(0,c)U(t)|\phi\rangle = U(t)\phi(0,c)|\phi\rangle - \lambda(t)U(t)|\phi\rangle. \tag{10.3}$$

Now $\phi(0,c)|\phi\rangle = \phi|\phi\rangle$ as a result of Proposition V, and $U(t)|\phi\rangle \subset P$. The assumption of simplicity yields

$$U(t)|\phi\rangle = |\phi + \lambda(t)\rangle. \tag{10.4}$$

Evidently $U^+(t)|\phi\rangle = |\phi - \lambda(t)\rangle$.

Using $I_\mu(t,x) = U(t)I_\mu(0,x)U^+(t)$ and Proposition V again, we find immediately

$$I_\mu(t,x)|\phi\rangle = U(t)I_\mu(0,x)|\phi - \lambda(t)\rangle$$
$$= I_\mu(\mathbf{x}, \phi - \lambda(t))|\phi\rangle. \tag{10.5}$$

We stop here for a short remark. Only the last formula is important for us. We can also derive it under a less stringent condition, namely that (10.1) is valid only when applied to states in P. The proof of (10.5) follows when we observe that $U(t)\bar{P}U(t)^+ = \bar{P}$, where $\bar{P}$ is the projection operator on P.

Proposition VIII: Let the system be in an eigenstate of charge with lowest possible energy. Let c be a path sufficiently far from the boundary of V_3 and let rigidity, simplicity, Meissner effect, and junction dominance be valid. The electric current is given by the function $I_\mu(\mathbf{x}, \phi - \lambda(t))$ (the Josephson current) where $0 < \phi < 2\pi$ and:

(a) I_μ is a periodic function of $\phi - \lambda(t)$ with period 2π.

(b) $\lambda(t) = 2e[-\mu_0(t) + \Phi_{ex}(t,c) - \Phi_{ex}(0,c)]$ for the Josephson junction and

(c) $\lambda(t) = 2e[\Phi_{ex}(t,c) - \Phi_{ex}(0,c)]$ for the double junction.

(d) These formulas are also valid if c' is homotopic to c.

Proof: (a) is just formula (10.5).

(b) and (c) follow from the definition of $\lambda(t)$ given by (10.2) and Proposition VII.

(d) follows from the Propositions V and VII.

Formula (c) needs some discussion. It does not contain the electric field of the junction. This could lead to the conclusion that even when the junction field vanishes, we still get a current which varies with the external flux, i.e., we get a change in our system which is confined to V_3 by a magnetic field which vanishes in V_3.

This conclusion is wrong, since the derivation of our result is only correct under the assumption of junction dominance, which is not valid when the electric field of the junction vanishes. Junction dominance and vanishing junction field are just two limiting cases. In the latter case we can safely assume that the dynamics of the interior of V_3 are completely independent of the external magnetic field.

11. EXOTIC SPINORS VERSUS ELECTRON PAIRS

We discuss now the results which we have obtained in the last section. With the assumptions of Meissner effect, rigidity, simplicity, and junction dominance, we have derived the correct expressions for flux quantization and Josephson current. They are given by Proposition IV and VIII. Note that we obtain the Josephson current as a general periodic function of λ. This is in complete agreement with the experimental facts.[15] Our derivation yields identical results for two types of models: electrodynamics with exotic spinors of charge e and electrodynamics with bosons of charge $2e$, i.e., electron pairs. For a single Dirac field we would only obtain the original result of London with a factor 2 missing.[14] Note that we take the Meissner effect for granted, i.e., we do not explain every special feature in superconductivity, but use some of them to obtain our result. We cannot insure *a priori* that our models indeed satisfy all the additional assumptions which were made. Perhaps one has to allow for other interactions (e.g., with the phononfield) which insure that these assumptions are actually valid. This will not effect our conclusions as long as the operator $\overline{T}$ defined in Sec. 8 still defines a unitary transformation which leaves the equations of motion invariant.

The fact that we used relativistic quantum field theory does not mean that Josephson current, and flux quantization are of relativistic origin. In fact, the theory can also be developed in a nonrelativistic setting with Pauli spinors, but these would require a different mathematical formalism, since the relevant principal bundles have a different group. We mention here that the classification of different spin structures in the nonrelativistic case yields the same results as the relativistic case which we have described. We have also considered a boson field of charge $2e$, i.e., electron pairs, in order to facilitate the comparison with the conventional theory of superconductivity.[10] Our derivation of flux quantization and Josephson current is quite new in itself, so we wanted to check that it also works for electron pairs. As we have mentioned, electron pairs and exotic spinors yield identical results.

Measurements of the Josephson current which contains the factor $2e$ in a characteristic manner, are so precise that they are even used as the best experimental determination of e. The pairing hypothesis is, therefore, regarded as one of the best established theories in physics. On the other hand, it is clearly an approximation with limited applicability, e.g., it predicts a wrong behavior of the Knight shift.[14] The question arises why the predicted charge dependence of magnetic flux and Josephson current is experimentally verified to such an astonishing degree of accuracy. By relating the charge dependence to an invariant topological property of the whole macroscopic superconductor, our model with exotic spinors yields a natural answer to this question. On the other hand, the physical origin of exotic spinors is obscure. At the moment I can only offer the speculation that when the theory of superconductivity is formulated with the help of the Bogolyubov—Valatin transformation, it might happen that the two types of spinors emerge as quasiparticles for two different vacuum solutions of the Hartree—Bogolyubov equations. The work of Byers and Yang,[16] Bohr and Mottelson,[17] and especially of Uhlenbrock and Zumino[18] shows that such solutions indeed exist. Moreover, it is known that the Bogolyubov transformation even yields an exact solution in certain models. These models[19] are characterized by the appearance of inequivalent representations of the fermion anticommutation relations, which do not have a well-defined number operator so that gauge invariance of the first kind cannot be unitarily implemented. I do not think that the last point is relevant for the present problem, but also these models show that the appearance of different types of quasiparticles is not all unfamiliar to the standard theory of superconductivity. Up to now I have not been able to combine these results with my own geometrical considerations, though I see the necessity of such a connection, since the standard theory of superconductivity not only explains the effects mentioned in this paper, but also a lot of other phenomena.

Instead I have made a step backwards to phenomenology and have presented here a model in which the consequences of the geometrical facts can be easily demonstrated.

We must still say some words about our derivation of the current in the Josephson junction. In contrast to the conventional derivation, we decisively used that the current flows in a closed circuit. The shape of the circuit may be quite arbitrary, as we have shown. It may contain a device for applying the discontinuous voltage step at the junction (which is produced by a thin insulating layer[15]) and even an instrument for measuring the current. All this will not invalidate our conclusions, as we have seen, because the topology is always that of the ring. We must, of course, assume that the instrument used for measurements does not affect the superconducting properties of the whole system too much. The question now arises what happens when the geometry of the system is more complicated. The answer is that in such a case even more exotic spinors are possible, but by a more careful investigation which is beyond the scope of this article one can show that the predictions for flux and current are unchanged. [The reader who wants to check this must note that $(1/2\pi i)\lambda^{-1}\delta\lambda$ always defines an integer cohomology class for the characteristic function λ (compare Sec. 6).]

CONCLUSION

Exotic spinors always appear when the underlying manifold is not simply connected. Several beautiful solutions of Einstein's equations yield manifolds with this property.[20] The most prominent one is perhaps the Kerr solution with a ring singularity. In practically all solutions of Einstein's equations, the underlying manifold turns out to be parallelizable.[3] This provides us immediately with a first trivial spin-structure. The situation is in this respect not very different from our simple example. One then is readily inclined to reject the other nontrivial spin structures as somewhat artificial objects. Hopefully, our application to superconductivity has convinced the reader that this attitude is premature. A manifold which is not simply connected also appears in the discussion of the Bohm–Aharonov effect.[21] The mathematical results of Secs. 1–6 have been derived in such a way that they can be applied to all these cases.

ACKNOWLEDGMENTS

I thank Professor W. Greub, Professor K. Bleuler, and Professor D. Schütte for enlightening discussions on mathematical and physical problems connected with this article, which could not have been written without their continuous help.

[1] M. F. Atiyah, R. Bott, and A. Shapiro, Topology **3**, Suppl. 1 3 (1964).
[2] A. Lichnerowicz, Bull. Soc. Math. France **92**, 11 (1964).
[3] R. Geroch, J. Math. Phys. **9**, 1739 (1968).
[4] M. E. Mayer, *Differential Geometrical Methods in Mathematical Physics*, Lecture Notes in Mathematics **570** (Springer, New York, 1977).
[5] A. Haefliger, C. R. Acad. Sci. Paris **243**, 558 (1956).
[6] R. Bott and J. Mather, *Battelle Rencontres*, 1967 *Lectures in Mathematics and Physics* (Benjamin, New York, 1968).
[7] N. Steenrod, *The Topology of Fibre Bundles* (Princeton U. P., Princeton, N. J., 1951).
[8] W. Greub and H. R. Petry, *On the Lifting of Structure Groups* Differential Geometrical Methods in Mathematical Physics, II, Lecture Notes in Mathematics (Springer, New York, 1978).
[9] W. Greub, S. Halperin, and R. Vanstone, *Connections, Curvature and Cohomology* (Academic, New York, 1973).
[10] K. Kobayashi and S. Nomizu, *Foundations of Differential Geometry* (Interscience, New York, 1963).
[11] W. Greub, private communication.
[12] E. H. Spanier, *Algebraic Topology* (McGraw-Hill, New York, 1966).
[13] J. D. Bjorken and S. D. Drell, *Relativistic Quantum Mechanics* (McGraw-Hill, New York, 1964).
[14] J. R. Schrieffer, *The Theory of Superconductivity* (Benjamin, New York, 1964.
[15] J. R. Waldram, Rep. Prog. Phys. **39**, 751 (1976).
[16] N. Byers and C. N. Yang, Phys. Rev. Lett. **7**, 46 (1961).
[17] A. Bohr and B. R. Mottelson, Phys. Rev. **125**, 495 (1962).
[18] D. A. Uhlenbrock and B. Zumino, Phys. Rev. **133**, 350 (1964).
[19] R. Haag, Nuovo Cimento **25**, 287 (1962).
[20] S. W. Hawking and G. F. R. Ellis, *The Large Scale Structure of Space–Time* (Cambridge U. P., Cambridge, 1973).
[21] Y. Aharonov and D. Bohm, Phys. Rev. **115**, 485 (1959).

On the expansion of the propagator in power series of the coupling constant[a)]

R. B. Perez

Engineering Physics Division, Oak Ridge National Laboratory, Oak Ridge, Tennessee 37830

(Received 20 June 1978)

We show by simple rearrangements of the Lippmann–Schwinger equation, that the propagator (Green's function) of a linear physical field can be expanded in power series valid for values of the coupling constant either close to unity or very large. The new series satisfy the generalized resolvent operator equation first derived by Mockel.

I. INTRODUCTION AND REVIEW OF SOME RESULTS

The propagator (or Green's function) plays a central role in the mathematical description of a physical field. In most instances, only approximate expressions for the propagator can be obtained for the case of small departures from an exactly solvable physical situation. This results in the representation of the propagator as a power series in terms of the strength (coupling constant) of the interaction. This series is usually referred to as the Neumann expansion, as it arises from the iteration of the integral equation defining the propagator. In what follows, we would like to show that from the approximate knowledge of the propagator for values of the strength constant either equal to unity, or very large, one can derive two new series for the propagator, which we shall call the second and third Neumann series. This procedure will be illustrated for quantum mechanical systems, although it applies as well to any linear physical field.

Consider a quantum mechanical system represented by the Schrödinger equation,

$$(EI-H)\Psi=0, \tag{I.1}$$

where E is the energy of the system, I is the unit operator, and the Hamiltonian operator H is given in the form

$$H=H_0+\epsilon H_1, \tag{I.2}$$

with ϵ representing the interaction coupling constant. The unperturbed Hamiltonian is H_0, and the perturbation operator H_1 is written as

$$H_1=H_D+\tau H_N, \tag{I.3}$$

where we have split the interaction into a "diagonal" operator H_D and a nondiagonal operator H_N. The perturbation parameter τ ($0\le\tau\le1$) turns the nondiagonal interaction H_N "on" and serves as a convenient device to keep track of the order of the perturbation.

The following restrictions are imposed on the otherwise arbitrary operators H and H_1:

(a) The operators H_0, H_D, and H_N are compact operators of the Hilbert–Schmidt class, i.e.,

$$\operatorname{trace}(AA^\dagger)<\infty, \tag{I.4}$$

where A is any of the above operators and $A^\dagger$ is its adjoint.

(b) The inverse of the diagonal operator H_D, i.e., H_D^{-1}, exists as a bounded operator.

We also state without proof the following theorem.[1]

Theorem 1: The product of a compact operator and a bounded operator is also compact.

The full propagator G is defined by the relation

$$(EI-H)G=I, \tag{I.5}$$

which for the unperturbed case becomes

$$(EI-H_0)G_0=I. \tag{I.6}$$

The full and the unperturbed propagators are related by the Lippmann–Schwinger equation[2]

$$G=G_0+\epsilon GH_1G_0. \tag{I.7}$$

Let $B[(\lambda),\ (\eta)]$ be an operator, which is an analytical function of a set of parameters (η_i), and which is associated with a set (λ) of eigenvalues. We define its resolvent $R_B\{(\lambda),\ B[(\lambda),\ (\eta)]\}$ by

$$R_BB=I \tag{I.8}$$

for any

$$\{\lambda\}\in\rho(B), \tag{I.9}$$

where $\rho(B)$ is the resolvent set of B. It can be shown that the operator R_B satisfies the generalized resolvent equation[3]

$$\frac{\delta}{\delta n_i}R_B(\eta_i)=-R_B(\eta_i)\left[\frac{\delta}{\delta\eta_i}B(\eta_i)\right]R_B(\eta_i) \tag{I.10}$$

together with the reciprocity relation

$$R_B(\mathbf{x},\mathbf{x}',\eta_i)=R_B^*(\mathbf{x}',\mathbf{x},\eta_i). \tag{I.11}$$

From the comparison of Eqs. (I.5) and (I.8), we conclude that the propagator itself satisfies the following generalized resolvent equation,

$$\frac{\delta}{\delta\eta_i}G(\eta_i)=-G(\eta_i)\frac{\delta}{\delta\eta_i}[E(\eta_i)I-H(\eta_i)]\,G(\eta_i). \tag{I.12}$$

In particular, identification of the parameter η_i with the coupling constant ϵ yields

$$\frac{\delta}{\delta\epsilon}G(\epsilon)=G(\epsilon)[\epsilon H_1]G(\epsilon), \tag{I.13}$$

which together with the initial condition

[a)]Sponsored by the Department of Energy Office of Energy Technology, Reactor Research and Technology Division, Reactor Design, under contract with the Union Carbide Corporation.

$$G(\epsilon=0)=G_0 \tag{I.14}$$

defines the propagator $G(\epsilon)$.

The Lippmann–Schwinger equation (I.7) and the generalized resolvent equation (I.13) are equivalent. Iteration of the former yields the first Neumann series,

$$G=\sum_{n=0}^{\infty} G_0(\epsilon H_1 G_0)^n. \tag{I.15}$$

Identical results are obtained from the Taylor series expansion,

$$G=\sum_{n=0}^{\infty}\frac{\epsilon^n}{n!}\left[\frac{\delta^n G}{\delta\epsilon^n}\right]_{\epsilon=0}, \tag{I.16}$$

where the derivatives, $|\delta^{(n)}G/\delta\epsilon^n|$, are computed from Eq. (I.13) and the initial condition Eq. (I.14). The series, Eq. (I.15), converges when

$$\|\epsilon(H_1G_0)\|<1, \tag{I.17}$$

which implies that the operator H_1G_0 must be a compact operator. The interaction operator H_1 is compact by definition; hence, according to Theorem 1, G_0 must be a bounded operator.

We conclude the introductory material by pointing out that partial summation[4] of this series, Eq. (I.15), yields the more convergent result

$$G=\sum_{n=0}^{\infty} G_D(\epsilon H_N G_D)^n \tag{I.18}$$

in terms of the "diagonal" propagator

$$G_D=G_0(I-\epsilon H_0G_0)^{-1}, \tag{I.19}$$

A result which can be also arrived at (see Appendix A) by trivial rearrangement of the Lippmann–Schwinger equation (I.7) or from the generalized resolvent equation (I.13). Various techniques[5–7] have been developed in the past to improve on the convergence of the Neumann series (I.15) by suitable rearrangements of this series. Wellner[5] accomplishes the rearrangement of the series (I.15) by expansion of the coupling constant in power series of an auxiliary parameter λ. Rotenberg[6] introduces two new operators in place of the identity operator I and the operator H_1G_0 in Eq. (A9). Finally, Weinberg[7] utilizes a conformal mapping of the coupling-constant plane to rearrange the Neumann series (I.15). The present method is based on the rearrangement of the Lippmann–Schwinger equation (I.7) itself, instead of its iterated form (I.15).

In essence, the method developed here exploits the well-known algebraic analogy between the operator $(\tau-\epsilon H_1G_0)^{-1}$, and its expansion (I.15), and the expansion of the function $(1-X)^{-1}$, where X is a number real or complex.

II. THE "SECOND" NEUMANN SERIES EXPANSION OF THE PROPAGATOR

The purpose of this section is to construct and discuss a series expansion for the propagator for values of the coupling constant, ϵ, close to unity. To this end we rearrange the Lippmann– Schwinger equation (I.7) in the form

$$I=G^{-1}G_0+\epsilon H_1G_0, \tag{II.1}$$

which after adding and subtracting the operator H_1G_0 becomes

$$I=G^{-1}G_0+(\epsilon-1)H_1G_0+H_1G_0. \tag{II.2}$$

Next, from Eq. (I.7) evaluated at $\epsilon=1$, we obtain

$$G(1)=G_0(I-H_1G_0)^{-1}, \tag{II.3}$$

which can be solved for H_1G_0, i.e.,

$$H_1G_0=I-G^{-1}(1)G_0. \tag{II.4}$$

Introduction of Eq. (II.4) into Eq. (II.2), yields after trivial rearrangements the result

$$G(\epsilon)=G(1)[I-(\epsilon-1)Q_1]^{-1} \tag{II.5}$$

with

$$Q_1=H_1G(1). \tag{II.6}$$

Iteration of Eq. (II.5) yields the "second" Neumann series

$$G(\epsilon)=\sum_{n=0}^{\infty} G(1)[(\epsilon-1)Q_1]^n. \tag{II.7}$$

The convergence of the above series depends on the condition

$$\|(\epsilon-1)H_1G(1)\|<1, \tag{II.8}$$

which in turn demands that the operator $H_1G(1)$ be a compact operator of the Hilbert–Schmidt class. The interaction operator H_1 is by definition a compact operator, hence the propagator $G(1)$ must be bounded (see Appendix A). The propagator $G(1)$ is obtained from Eq. (A7) evaluated at $\epsilon=1$. The result is

$$G(1)=\sum_{n=0}^{\infty} G_D(1)[\tau H_N G_D(1)]^n \tag{II.9}$$

with the condition for convergence

$$\|\tau H_N G_D(1)\|<1, \tag{II.10}$$

which implies that the "diagonal" propagator $G_D^{(1)}$ (evaluated at $\epsilon=1$) must be bounded.

The second Neumann series, i.e., Eq. (II.7), must also satisfy the generalized resolvent equation (I.13). To prove this point it suffices to introduce the series Eq. (II.7) into Eq. (I.13). One obtains

$$\sum_{n=1}^{\infty} n(\epsilon-1)^{n-1}G(1)Q_1^n = \sum_{n=0}^{\infty}\sum_{n'=0}^{\infty}(\epsilon-1)^{n+n'-1}G(1)Q_1^{(n+n'+1)}, \tag{II.11}$$

an expression which after equating the coefficients of equal powers in the quantity $\epsilon-1$ becomes a set of identities.

III. THE THIRD NEUMANN SERIES EXPANSION OF THE PROPAGATOR

In this section we discuss a series expansion of the propagator for large values of the coupling constant ($\epsilon\gg1$). To this end we introduce the "strong coupling operator," Q_2, defined by

$$Q_2=G_0^{-1}H_1^{-1}. \tag{III.1}$$

On the basis of this operator the Lippmann–Schwinger equation (I. 7) is rewritten in the form

$$G = G_0 + \epsilon G Q_2^{-1}. \quad (\text{III}.2)$$

Left multiplication of Eq. (III. 2) by the Q_2 operator yields, after some rearrangement, the result

$$G = -\epsilon^{-1} G_0 Q_2 (I - \epsilon^{-1} Q_2)^{-1}, \quad (\text{III}.3)$$

which leads to the following series expansion for the propagator,

$$G = -\sum_{n=1}^{\infty} \epsilon^{-n} Q_2^n. \quad (\text{III}.4)$$

The convergence of the series (III. 3) depends on the condition

$$\|\epsilon^{-1} Q_2\| < 1, \quad (\text{III}.5)$$

the fulfillment of which requires that the Q_2 operator must be a compact operator of the Hilbert–Schmidt class. We shall now examine the condition

$$\|Q\| < M, \quad (\text{III}.6)$$

where M is a finite limit. From the definition [Eq. (III. 1)], of the strong coupling operator Q_2, rewritten in the form

$$Q_2 (H_1 G_0) = I, \quad (\text{III}.7)$$

we immediately realize that this operator is the resolvent of the assumedly compact operator $H_1 G_0$. It should then satisfy the generalized resolvent equation (I. 10), i. e.,

$$\frac{\delta}{\delta \tau} Q_2 = -Q_2 \left[\frac{\delta}{\delta \tau} H_1 G_0 \right] Q_2, \quad (\text{III}.8)$$

where the arbitrary parameter η_i has been replaced by the perturbation parameter τ. Introduction of Eq. (I. 3) into Eq. (III. 8) yields

$$\frac{\delta}{\delta \tau} Q_2 = -Q_2 (H_N G_0) Q_2, \quad (\text{III}.9)$$

with the initial condition

$$Q_2(\tau = 0) = Q_0, \quad (\text{III}.10)$$

where the initial value of the strong coupling operator is obtained by setting $\tau = 0$ in Eq. (III. 1), i. e.,

$$Q_0 = H_0^{-1} G_0^{-1}. \quad (\text{III}.11)$$

We can now proceed on the evaluation of Q_2 by expansion in Taylor series around $\tau = 0$. We have

$$Q_2(\tau) = \sum_{\nu=0}^{\infty} \frac{\tau^\nu}{\nu!} \left[\frac{d^\nu}{d\tau^\nu} Q(\tau) \right]_{\tau=0}, \quad (\text{III}.12)$$

where the derivatives in Eq. (III. 12) are computed from the generalized resolvent equation (III. 9). The result is

$$Q(\tau) = Q_0 P(\tau), \quad (\text{III}.13)$$

where the operator $P(\tau)$ is given by

$$P(\tau) = \sum_{\nu=0}^{\infty} (-\tau)^\nu X^\nu \quad (\text{III}.14)$$

with

$$X = H_N H_D^{-1}. \quad (\text{III}.15)$$

Whenever the operator H_D^{-1} exists as a bounded operator the operator X will be compact in view of the compactness of H_N. From Theorem 1, P will be in turn compact. Also, if

$$\tau \|X\| < 1, \quad (\text{III}.16)$$

the series (III. 14) will converge with

$$\|P\| < M_P, \quad (\text{III}.17)$$

where M_P is the bound of the norm $\|P\|$. Calling M_{Q_0} the bound of $\|Q_0\|$, we have from Eq. (III. 13) and the Schwartz inequality the result

$$\|Q\| \le M_P M_{Q_0}, \quad (\text{III}.18)$$

from which we conclude that the strong coupling operator Q_2 is compact and bounded when $\|Q_0\|$ is bounded and the condition Eq. (III. 16) is satisfied.

The third Neumann series, Eq. (III. 4) must also satisfy the generalized resolvent equation (I. 13). This is easily proven by inserting the expansion Eq. (III. 4) into Eq. (I. 13). We obtain

$$\epsilon^{-1} G_0 \sum_{n=1}^{\infty} n (\epsilon^{-1} Q_2)^n$$

$$= G_0 \sum_{n+n'=1}^{\infty} \epsilon^{-(n+n')} Q_2^n (H_1 G_0) Q_2^{n'}, \quad (\text{III}.19)$$

which, after equating coefficients of equal powers in the inverse coupling constant ϵ^{-1} and on account of Eq. (III. 7), becomes a set of identities.

Finally because the strong coupling operator Q_2 satisfies the generalized resolvent equation, it must also satisfy a reciprocity relation similar to Eq. (I. 11).

IV. CONCLUSIONS AND DISCUSSION

We have constructed, by means of simple rearrangements of the Lippmann–Schwinger equation, two new series for the propagator. The so-called second Neumann series, Eq. (II. 7), is a power series in $(\epsilon - 1)$. The convergence of this series depends on the existence of the propagator, $G(\epsilon = 1) = G(1)$, given by Eq. (II. 9). The third Neumann series is an expansion in terms of the inverse powers of the coupling constant Eq. (III. 4), where the strong coupling operator Q_2 is given by Eq. (III. 13).

The leading term in this series is from Eqs. (III. 1) and (III. 4),

$$G(\epsilon) = -(\epsilon H_1)^{-1}, \quad (\text{IV}.1)$$

which coincides with the results of taking the limit of the expression (A. 9) for the propagator in the case of large values of the coupling constant.

The three Neumann series satisfy both the Lippmann–Schwinger equation and the generalized resolvent equation, and in this sense they might be considered as analytical continuations of each other.

We have also studied the conditions required for Q_1 and Q_2 to be operators of the Hilbert–Schmidt class. Both operators are expressed as a power series in the perturbation parameter τ, which switches "on" the nondiagonal part of the interaction of the Hamiltonian. In consequence, the second and third Neumann series are in fact double series expansions.

In a forthcoming paper we show that the two series obtained by rearrangement of the Lippmann–Schwinger equation can also be thought of as the analytical continuation of a functional hypergeometric function by composition, first studied by Volterra[8,9] in connection with the Neumann series solution of integral equations.

APPENDIX A. EXPANSION OF THE PROPAGATOR IN TERMS OF THE "OFF-DIAGONAL" MATRIX ELEMENTS OF THE INTERACTION OPERATOR

In this Appendix we illustrate the use of the generalized resolvent equation (I.10) to obtain a perturbation expansion of the propagator in terms of the "off-diagonal" matrix elements of the interaction operator. To this end we identify the arbitrary parameter η_i with the parameter τ, and write

$$B(\tau) = EI - H(\tau), \tag{A1}$$

$$R_B(\tau) = G(\tau). \tag{A2}$$

Introduction of Eqs. (A1), (I.2), and (I.3) into Eq. (I.10) yields

$$\frac{\delta}{\delta\tau} G(\tau) = \epsilon G(\tau) H_N G(\tau), \tag{A3}$$

which is in the coordinates representation an integro-differential equation for the propagator, subject to the initial condition

$$G(\tau = 0) = G_D, \tag{A4}$$

where the "diagonal" propagator G_D is obtained from the Lippmann–Schwinger equation (I.7) evaluated at $\tau = 0$. We then have

$$G_D = G_0 (I - \epsilon H_D G_0)^{-1}, \tag{A5}$$

so that we can express the propagator $G(\tau)$ in the form of the Taylor series expansion

$$G(\tau) = \sum_{n=0}^{\infty} \frac{\tau^n}{n!} \left(\frac{d^n G}{d\tau^n}\right)_{\tau=0}, \tag{A6}$$

where the successive derivatives of the Green's function are computed from Eq. (A3). We obtain in this manner

$$G(\tau) = \sum_{n=0}^{\infty} G_D (\tau\epsilon H_N G_0)^n, \tag{A7}$$

which is a convergent series expansion for the propagator provided that

$$\| \tau\epsilon H_N G_D \| < 1, \tag{A8}$$

which implies that the operator $(H_N G_D)$ must be a compact operator of the Hilbert-Schmidt class. Hence, G_D must be a bounded operator, in view of Theorem 1 and since by definition H_N is a compact operator.

The result (A7) is also obtained from the Lippmann–Schwinger equation (I.7), rewritten in the form

$$G = G_0 (I - \epsilon H_1 G_0)^{-1}, \tag{A9}$$

which in view of Eq. (I.3) can be rearranged as

$$G = G_0 [(I - \epsilon H_0 G_0) - \epsilon\tau H_N G_0]^{-1}, \tag{A10}$$

so that expansion in terms of the nondiagonal matrix operation, $H_N G_0$, yields again the result (A7).

[1]A. E. Taylor, *Introduction to Functional Analysis* (Wiley, New York, 1958), p. 285.
[2]B. A. Lippmann and J. Schwinger, Phys. Rev. **79**, 469 (1950).
[3]A. J. Mockel, J. Math. Phys. **8**, 2318 (1967).
[4]R. P. Feymann, Phys. Rev. **76**, 769 (1949).
[5]M. Wellner, Phys. Rev. **132**, 1848 (1963).
[6]M. Rotenberg, Ann. Phys. (N.Y.) **21**, 579 (1963).
[7]S. Weinberg, J. Math. Phys. **5**, 743 (1964).
[8]V. Volterra, *Theory of Functional and of Integral and Integro-Differential Equations* (Dover, New York, 1959), p. 99.
[9]V. Volterra and J. Pérès, *Leçons sur composition et les functions permutables* (Coll. Borel, Gauthier–Villars, Paris, 1924).

Conformal flatness and the Schwarzschild interior solution

A. K. Raychaudhuri and S. R. Maiti

Department of Physics, Presidency College, Calcutta 700073, India
(Received 5 April 1978)

The paper gives a rigorous proof of the theorem that the only static conformally flat metric for a perfect fluid distribution is the Schwarzschild interior metric.

1. INTRODUCTION

As is well known, the Schwarzschild interior solution is static and conformally flat. The converse theorem that the only static, conformally flat metric for a perfect fluid distribution (subject to the field equations of relativity) is the Schwarzschild interior solution, has been claimed to be proved by quite a number of investigators at different times. However, none of these proofs seems to be quite general and/or free from questionable assumptions. Thus Buchdahl[1] (1971) assumed that a fluid distribution in equilibrium must be spherically symmetric. This however is not generally true as is evident from some solutions given somewhat later by Barnes[2] (1972). However all the relevant Barnes' solutions have singularities and it may well be that the theorem assumed by Buchdahl is true when one introduces the additional condition of regularity, but even then no proof in the literature has come to the notice of the present authors.

Shortly after Buchdahl, Misra, and Tribedi[3] gave a proof assuming that a static conformally flat metric must be of the form

$$ds^2 = e^{2\sigma}[dx^2 + dy^2 + dz^2 - dt^2],$$

with σ independent of time t. This, however, is not correct. A contrary example is provided by the de Sitter metric

$$ds^2 = e^{2g}[dx^2 + dy^2 + dz^2 - dt^2],$$

with "g" a function of the time coordinate. This metric is static as it is transformable to

$$ds^2 = \frac{dr^2}{1 - r^2/R^2} + r^2(d\theta^2 + \sin^2\theta\, d\varphi^2) - \left(1 - \frac{r^2}{R^2}\right) dt^2$$

which admits a translation along the t axis.

More recently, apparently ignorant of the earlier works of Buchdahl and Misra and Tribedi, Gurses and Gürsey[4] proved the theorem again with the condition of spherical symmetry. A somewhat different but related theorem has been proved recently by Collinson.[5] Every conformally flat axisymmetric stationary space–time is necessarily static and if the source is a perfect fluid, then the space–time metric is the Schwarzschild interior metric. In the present discussion, we first point out that for a perfect fluid, conformal flatness leads to either a spatial constancy of the energy density ρ and vanishing of the vorticity ω and shear σ_{ik} or a vanishing of $(p+\rho)$, where p is the pressure of the fluid. In the latter case, the metric reduces to the de Sitter form and the velocity of the fluid is indeterminate. In the first case if the expansion θ is assumed to vanish, the Schwarzschild interior metric follows. (If $\theta \neq 0$, then we have either the isotropic homogeneous cosmological solution or a family of nonhomogeneous cosmological models[6,7]).

2. THEOREM AND ITS PROOF

Theorem: If the source of the gravitational field be a perfect fluid with vanishing expansion and nonnegative density and pressure, then the only conformally flat space–time consistent with Einstein's gravitational equations is the Schwarzschild interior metric.

We recall some equations deduced by Ehlers, Kundt, and Trümper (all of which are presented in a review by Ehlers). We rewrite the necessary equations in the form for a perfect fluid as presented by Ehlers.[8]

$$h^f_a h^g_b \mathring{\sigma}_{fg} - h^f_a h^g_b \mathring{U}_{(f;g)} - \mathring{U}_a \mathring{U}_b + \omega_a \omega_b + \sigma_{af}\sigma^f_b + \tfrac{2}{3}\theta\sigma_{ab} + h_{ab} \times\left(-\tfrac{1}{3}\omega^2 - \tfrac{2}{3}\sigma^2 + \tfrac{1}{3}\mathring{u}^c{}_{;c}\right) + E_{ab} = 0, \tag{1}$$

$$h_a{}^b E_{bc;d} h^{cd} + 3H_{ab}\omega^b - \eta_{abcd} u^b \sigma^c_e H^{de} = -\tfrac{1}{3} h_a{}^b \rho_{,b} \tag{2}$$

$$h_a{}^b H_{bc;d} h^{cd} - 3E_{ab}\omega^b - \eta_{abcd} u^b \sigma^c_e E^{de} = (\rho + 3\rho)\omega_e, \tag{3}$$

$$\perp' E'_{ab} + h^f_{(a}\eta_{b)cde} u^c H_f{}^{d;e} + E_{ab}\theta - E^c{}_{(a}\omega_{b)c} - E^c{}_{(a}\sigma_{b)c} - \eta_{acde}\eta_{bpqr} u^c u^p \sigma^{dq} E^{er} + 2H^d{}_{(a}\eta_{b)cde} u^c u'^e = -\tfrac{1}{2}(\rho + p)\sigma_{ab}. \tag{4}$$

In the above u^a is the velocity vector of the fluid, ω^e, σ^{ab}, $\mathring{u}^a$ and θ are vorticity, shear, acceleration, and expansion defined in usual way. p and ρ are pressure and density of the fluid. E_{ab} and H_{ab} are the so-called electric and magnetic type of components of the Weyle tensor C_{abcd}, where

$$E_{ab} = C_{acbd} u^c u^d \tag{5}$$

and

$$H_{ab} = \tfrac{1}{2}\eta_{ac}{}^{gh} C_{ghbd} u^c u^d \tag{6}$$

For conformally flat space–time $E_{ab} = H_{ab} = 0$. With expansion $\theta = 0$, we have from Eqs. (2)–(4)

$$\rho_{,a} h^{ab} = 0 \tag{7}$$

and

$$(\rho + p)\omega_a = (\rho + p)\sigma_{ab} = 0.$$

Hence either

$$\rho + p = 0 \quad \text{or} \quad \omega_a = \sigma_{ab} = 0. \tag{8}$$

Equation (8) shows that the 3-space elements orthogonal to the velocity vector mesh together, and from Eq. (7) it follows that ρ is constant in 3-space. Again from the conservation relation

$$(p + \rho)\theta + \dot{\rho} = 0 \tag{9}$$

 0022-2488/79/020245-02$01.00

$\dot{\rho}=0$ if $\theta=0$, i.e., ρ is constant in time as well. The vanishing of σ,ω, and θ allow us to take the line element in the static form.

$$ds^2=g_{00}dt^2+g_{ik}dx^i dx^k, \tag{10}$$

where g_{00} and g_{ik}'s are all independent of time and the velocity vector of the fluid is $u^a=(g_{00})^{-1/2}\delta^a_0$.

The Ricci tensor for the 3-space metric g_{ik} can now be written as (cf. Ehlers[8])

$$R^*_{ik}=\dot{u}_{(i;k)}+\dot{u}_i\dot{u}_k+\tfrac{1}{3}g_{ik}(2\rho-\dot{u}^c{}_{;c}).$$

From Eq. (1) we get

$$\dot{u}_{(i;k)}+\dot{u}_i\dot{u}_k-\tfrac{1}{3}g_{ik}\dot{u}^c{}_{;c}=0. \tag{12}$$

Substituting (12) in (11)

$$R^*_{ik}=\tfrac{2}{3}\rho g_{ik}. \tag{13}$$

Since ρ is constant, the 3-space is a space of constant curvature. Hence the line element can be written as (cf. Eisenhart[9])

$$ds^2=g_{00}dt^2-\frac{dx^2+dy^2+dz^2}{(1+Kr^2/4)^2}. \tag{14}$$

Using Dingle's formulae (as reproduced by Tolman[10]), g_{00} in the metric (14) is

$$g_{00}=\left(A_1-\frac{B_1}{1+Kr^2/4}\right)^2, \tag{15}$$

where A_1 and B_1 are two arbitrary constants.

Presenting $\rho=3/R^2$, and using a scale factor $r=2RL$ the final form of the metric has become

$$ds^2=\left(A-B\,\frac{1-L^2}{1+L^2}\right)^2 dt^2 - \frac{4R^2}{(1+L^2)^2}[dL^2+L^2d\theta^2+L^2\sin^2\theta\, d\varphi^2], \tag{16}$$

where A and B are new constants. This is Schwarzschild interior metric.

3. THE CASE $p+\rho=0$

If $p+\rho=0$, hence ρ is constant, p is also a constant. Then the energy—momentum tensor may be written as

$$T^a{}_b=-\rho\delta^a_b$$

and from Einstein field equation

$$R^a{}_b=-\rho\delta^a_b$$

Hence the space is an Einstein space. The line element may then be written[9] as

$$ds^2=\frac{dt^2-dx^2-dy^2-dz^2}{\{1+(K_0/4)(t^2-r^2)\}^2}, \tag{17}$$

with $K_0=-p=\rho$. A proper transformation leads to de Sitter line element; but u^a and along with it ω^a and σ^{ab} remain indeterminate.

ACKNOWLEDGMENTS

One of the authors SRM is grateful to the University Grants Commission for the award of a teacher fellowship and to the authorities of the Presidency College, Calcutta and Midnapur College, Midnapur for hospitality and grant of leave respectively.

[1] R.A. Buchdahl, Am. J. Phys. **39**, 158 (1971).
[2] A. Barnes, J. Phys. A **5**, 374 (1972).
[3] M. Misra and V.M. Trivedi, Curr. Sci. **41**, 805 (1972).
[4] M. Gurses and Y. Gursey, Nuovo Cimento B **28**, 786 (1975).
[5] C.D. Collinson, Gen. Rel. Grav. 7, 419 (1976).
[6] M. Trümper, Z. Astro. Phys. **66**, 215 (1967).
[7] L.C. Shepley and T. Taub, Commun. Math. Phys. **5**, 237 (1967).
[8] J. Ehlers, *Relativity and Cosmology* (Academic, New York, 1971).
[9] L.P. Eisenhart, *Riemanian Geometry* (Princeton U.P., Princeton, N.J., 1949).
[10] R.C. Tolman, *Relativity, Thermodynamics & Cosmology* (Oxford U.P., Oxford, 1934).

Invariant operators of IU(n) and IO(n) and their eigenvalues

M. K. F. Wong

Fairfield University, Fairfield, Connecticut 06430

Hsin-Yang Yeh

Moorhead State University, Moorehead, Minnesota 56560

A systematic explicit evaluation of the invariant operators of IU(n) and IO(n) has been carried out. It is found that the invariant operators of IU(n) and IO(n) can be obtained from those of U(n,1) and O(n,1) by simple substitution. Similarly the eigenvalues of the invariant operators of IU(n) and IO(n) can be obtained from those of U(n,1) and O(n,1) by simple substitution. Since the invariant operators and their eigenvalues of U(n,1) and O(n,1) are closely related to those of U($n+1$) and O($n+1$) our results can be expressed in explicitly closed and simple form.

1. INTRODUCTION

Recently there has been some renewed interest in the explicit evaluation of the eigenvalues of the invariant operators of the unitary, orthogonal and simplectic groups. This attempt was started by Perelomov and Popov,[1–3] continued by Wong and Yeh,[4] who obtained the eigenvalues in closed but complicated form, and further improved by Nwachuku and Rashid,[5,6] who obtained the eigenvalues in closed and simple form. Okubo[7] and Edwards[8] then showed that the last result can be obtained by another simple method. So far, however, the results are confined to compact groups. This leads us to ask whether there are noncompact groups whose invariant operators and eigenvalues can be explicitly evaluated in closed form. Of course, the results of U($n+1$) and O($n+1$) can be trivially extended to U(n, 1) and O(n, 1). But U(n, 1) and O(n, 1) can be obtained by Chakrabarti.[9] However, he did not discuss we expect that it may be possible to carry out a systematic evaluation of the invariant operators of IU(n) and IO(n). In this article we show that this is indeed the case.

The representations of IU(n) and IO(n) have been obtained by Chakrabarti.[9] However, he did not discuss systematically the invariant operators and their eigenvalues. Rosen and Roman[10] obtained a sixth order invariant operator for IU(n) and fourth order invariant operator for IO(n), but did not calculate their eigenvalues. Nor, to our knowledge, do we know of any other systematic study of the above problem. In this article we shall show that the invariant operators and their eigenvalues of IU(n) and IO(n) can be obtained from those of U(n, 1) and O(n, 1) by direct substitution.

In Sec. 2 we give a brief summary of the representations of IU(n) and IO(n) and their connection with U(n, 1) and O(n, 1). In Sec. 3 we present a detailed calculation for the sixth and ninth order invariant operators of IU(n) and their relation to C_2 and C_3 of U(n, 1). In Sec. 4 we present a similar calculation for the fourth and eighth order invariant operators of IO(n) and their relation to C_2 and C_4 of O(n, 1). Finally, we generalize the results to all orders and present them as four theorems in Sec. 5.

2. REPRESENTATIONS OF IU(n) AND IO(n) AND THEIR RELATION TO U(n, 1) AND O(n,1)

We follow basically the notations of Chakrabarti.[9] IU(n) has generators A^i_j (i, $j=1,2,\ldots,n$) and $I^{\,i}_{n+1}$, I^{n+1}_i ($i=1,2,\ldots,n$). The commutation relations are

$$[A^i_j, A^k_l]=\delta^i_l A^k_j-\delta^k_j A^i_l, \tag{2.1}$$

$$[A^i_j, I^{\,k}_{n+1}]=-\delta^k_j I^{\,i}_{n+1}, \tag{2.2}$$

$$[A^i_j, I^{n+1}_k]=\delta^i_k I^{n+1}_j, \tag{2.3}$$

$$[I^{\,i}_{n+1}, I^{\,j}_{n+1}]=[I^{n+1}_i, I^{n+1}_j]=[I^{n+1}_i, I^{\,j}_{n+1}]=0 \tag{2.4}$$

with

$$(A^i_j)^+=A^j_i, \tag{2.5}$$

$$(I^{\,i}_{n+1})^+=I^{n+1}_i, \tag{2.6}$$

where

$$i,j,k,l=1,2\ldots,n.$$

The matrix elements of $I^{\,n}_{n+1}$ are given by Eq. (2.9) of Ref. 9. The (infinite-dimensional) basis is given by (2.1) of Ref. 9. The deformation to U(n, 1) is obtained as follows. Define

$$A^i_{n+1}=\pm[\triangle, I^{\,i}_{n+1}]+i\epsilon I^{\,i}_{n+1} \tag{2.7}$$

$$A^{n+1}_i=\pm[\triangle, I^{n+1}_i]+i\epsilon I^{n+1}_i \tag{2.8}$$

with

$$\triangle\equiv\frac{1}{2\sqrt{\triangle_{(2)}}}\left[\sum_{i,j=1}^{n}A^i_jA^j_i+\left(\frac{\triangle''_{(3)}}{\triangle_{(2)}}+n\right)A^{n+1}_{n+1}\right. \tag{2.9}$$

$$\triangle_{(2)}=\sum_{i=1}^{n}I^{n+1}_iI^{\,i}_{n+1}, \tag{2.10}$$

$$\triangle''_{(3)}=\sum_{i,j=1}^{n}I^{n+1}_iA^i_jI^{\,j}_{n+1}+A^{n+1}_{n+1}\triangle_{(2)}$$

$$=\Lambda+A^{n+1}_{n+1}\triangle_{(2)}, \tag{2.11}$$

$$A^{n+1}_{n+1}|h\rangle=(\zeta+\sum_{i=2}h_{i\,n+1}-\sum_{i=1}h_{in})|h\rangle. \tag{2.12}$$

Then one finds that

 0022-2488/79/020247-04$01.00

$$A_{n+1}^{\ i} = (A^{n+1}_{\ \ i})^{+}, \qquad (2.13)$$

$$[A_{n+1}^{\ i}, A_{n+1}^{\ j}] = [A^{n+1}_{\ \ i}, A^{n+1}_{\ \ j}] = 0, \qquad (2.14)$$

and

$$[A^{n+1}_{\ \ i}, A_{n+1}^{\ j}] = A^j_i - \delta^j_i A^{n+1}_{n+1} \qquad (2.15)$$

It can be easily seen that we can also write A^{n+1}_{n+1} as

$$A^{n+1}_{n+1} = -\frac{\Lambda}{\Delta_{(2)}} - n + \zeta, \qquad (2.16)$$

where Λ has been defined in (2.11). Now it has been shown by Wong and Yeh[11] that ζ and ϵ are related to $h_{1\,n+1}$ and $h_{n+1\,n+1}$ as follows:

$$h_{1\,n+1} = -\frac{n}{2} + \frac{\zeta}{2} + i\epsilon\kappa \qquad (2.17)$$

$$h_{n+1\,n+1} = \frac{n}{2} + \frac{\zeta}{2} - i\epsilon\kappa, \qquad (2.18)$$

where $\kappa^2 = \Delta_{(2)}$ and κ need not be 1. For IU(n) Chakrabarti has obtained the following invariant operators:

$$\Delta_{(2)} = \sum_{i=1}^{n} I^{n+1}_{\ \ i} I_{n+1}^{\ i} = \kappa^2 \qquad (2.19)$$

$$\Delta'_{(3)} = \Lambda - \sum_{i=1}^{n} A^i_i \Delta_{(2)} = -\kappa^2\Big(\sum_{i=1}^{n} h_{i\,n+1} + n\Big) \qquad (2.20)$$

$$\Delta''_{(3)} = \Lambda + A^{n+1}_{n+1}\Delta_{(2)} = \kappa^2(\zeta - n). \qquad (2.21)$$

We shall show in Secs. 3 and 5 that all other higher order invariant operators can be obtained in closed and simple form.

For IO(n) the generators are $J_{ab} = -J_{ba}$ ($a, b, = 1, 2, \ldots, n$) and $I_{n+1,\ a}$ ($a = 1, 2, \ldots, n$). The commutation relations are

$$[J_{ab}, J_{cd}] = i(\delta_{ac}J_{bd} + \delta_{bd}J_{ac} - \delta_{ad}J_{bc} - \delta_{bc}J_{ad}), \qquad (2.22)$$

$$[J_{ab} I_{n+1c}] = i(\delta_{ac} I_{n+1b} - \delta_{bc} I_{n+1a}), \qquad (2.23)$$

$$[I_{n+1a}, I_{n+1b}] = 0, \qquad (2.24)$$

where

$$a, b, c, d = 1, 2, \ldots, n.$$

The basis of IO($2k$) and IO($2k-1$) and matrix elements of $I_{2k+1\,2k}$, $I_{2k-1\,2k}$ are given by (6.3), (6.7), (6.5), and (6.9) of Ref. 9 respectively. The deformation to O(n, 1) are obtained as follows: Define

$$J_{n\,n+1} = \frac{i}{\sqrt{\Delta_{(2)}}}\left[\frac{\Delta}{2}, I_{n\,n+1}\right] + \lambda I_{n\,n+1} \qquad (2.25)$$

or

$$J_{j\,n+1} = \sum_{i=1}^{n} \frac{1}{2}(I_{i\,n+1}J_{ij} + J_{ij}I_{i\,n+1}) + \lambda I_{j\,n+1} \quad (j = 1, 2, \ldots, n), \qquad (2.26)$$

where

$$\Delta_{(2)}|h\rangle = \sum_{k=1}^{n} I_{n+1\,k} I_{n+1\,k}|h\rangle = \kappa^2|h\rangle \qquad (2.27)$$

and

$$\Delta = \sum_{i<j=1}^{n} J^2_{ij}. \qquad (2.28)$$

Then one finds that

$$[J_{n+1\,a}, J_{n+1\,b}] = -iJ_{ab}. \qquad (2.29)$$

We have shown in Ref. 11 that for O($2k-1$, 1)

$$\lambda = il_{2k,1},$$

where

$$l_{2k,\alpha} = h_{2k\alpha} + k - \alpha, \qquad (2.30)$$

and for O($2k$, 1)

$$h_{2k+1\,1} = \tfrac{1}{2} - k + i\lambda. \qquad (2.31)$$

The only invariant operator that Chakrabarti has evaluated is $\Delta_{(2)} = \kappa^2$. We shall show in Secs. 4 and 5 that all higher order invariant operators of IO(n) can be obtained in closed and simple form.

3. EXPLICIT EVALUATION OF THE SIXTH AND NINTH ORDER INVARIANT OPERATORS OF IU(n)

In this section we attempt an explicit evaluation of the sixth and ninth order invariant operators of IU(n). Our method is as follows. First, we write down the second and third order invariants (C_2 and C_3) of U(n, 1). We then substract from each expression all terms which contain ζ and ϵ. The result is an invariant operator in IU(n). Since these expressions contain $\Delta^2_{(2)}$ and $\Delta^3_{(2)}$ respectively respectively on the denominators, we obtain the sixth and ninth order invariant operators by multiplying the resulting expressions by $\Delta^2_{(2)}$ and $\Delta^3_{(2)}$ respectively.

The eigenvalues are obtained in the following way. The eigenvalues of C_2 and C_3 for U(n, 1) are well known. We substract the terms containing ζ and ϵ according to (2.17) and (2.18). The result is the eigenvalues of the invariant operators of IU(n) which, as one would expect, contain no terms in $h_{1\,n+1}$ and $h_{n+1\,n+1}$.

Thus for the sixth order invariant operator of IU(n) we proceed in the following way. For notational convenience summation over repeated indices is assumed from now on except where it is indicated. Also Latin letters (except n) always go from 1 to n, while Greek letters go from 1 to $n+1$.

$$C_2 = A^\beta_\alpha A^\alpha_\beta = A^j_i A^i_j + A^{n+1}_{\ \ i} A^i_{n+1} + A^j_{n+1} A^{n+1}_{\ \ j} + (A^{n+1}_{n+1})^2. \qquad (3.1)$$

We find that (3.1) contains the following terms in ζ and ϵ [from (2.7), (2.8), and (2.12)]:

$$-2\Delta_{(2)}\epsilon^2 + \tfrac{1}{2}\zeta^2. \qquad (3.2)$$

Therefore, the sixth order invariant in IU(n) is

$$\begin{aligned} I_6 = {} & \Lambda^2 + 2n\Lambda\Delta_{(2)} + \Delta^2_{(2)}(A^i_j A^j_i + \Delta' I_i \Delta' I^i - 2\Delta'^2\Delta_{(2)} \\ & - I_i\Delta'^2 I^i + I_i\Delta' I^i\Delta' + \Delta' I^i \Delta' I_i - I^i\Delta'^2 I_i \\ & + I^i\Delta' I_i\Delta' + n^2\Delta^2_{(2)} \end{aligned} \qquad (3.3)$$

where

$$\Delta' = A^j_i A^i_j / 2\sqrt{\Delta_{(2)}} \qquad (3.4)$$

and

$$I^i = I^i_{n+1}, \quad I_i = I^{n+1}_i. \qquad (3.5)$$

It can be easily calculated that the eigenvalue of I_6 is

$$\kappa^4\left[\sum_{i=2}^{n} h_{i\,n+1}(h_{i\,n+1} + n + 2 - 2i) - \frac{n^2}{2}\right] \qquad (3.6)$$

Next we calculate the ninth order invariant operator I_9 of IU(n). From U(n, 1) we have

$$C_3 = A^{\beta}_{\alpha}A^{\gamma}_{\beta}A^{\alpha}_{\gamma} = A^{j}_{i}A^{k}_{j}A^{i}_{k} + 3A^{j}_{n+1}A^{k}_{j}A^{n+1}_{k} + 3A^{n+1}_{n+1}A^{k}_{n+1}A^{n+1}_{k} + (A^{n+1}_{n+1})^3 + (1-2n)A^{k}_{n+1}A^{n+1}_{k} + 2A^{k}_{i}A^{i}_{k} - n(A^{n+1}_{n+1})^2 + A^{i}_{i} - nA^{n+1}_{n+1} - A^{i}_{i}A^{n+1}_{n+1}. \tag{3.7}$$

From (3.7) we find that it contains the following terms in ϵ and ζ:

$$-\epsilon^2(\zeta 3\Delta_{(2)} + n\Delta_{(2)} + \Delta_{(2)}) + \tfrac{1}{4}\zeta^3 + \left(\frac{n-1}{4}\right)\zeta^2 + \frac{\Delta'_{(3)}}{\Delta_{(2)}}\zeta - \frac{n^2}{4}\zeta + \frac{n}{2}\zeta. \tag{3.8}$$

Therefore, I_9 is

$$I_9 = \Big\{A^{j}_{i}A^{k}_{j}A^{i}_{k} + 3(\Delta' I^{j} - I^{j}\Delta')A^{k}_{j}(\Delta' I_k - I_k\Delta') + 3\left(-\frac{\Lambda}{\Delta_{(2)}} - n\right)(\Delta' I^{k} - I^{k}\Delta')\times(\Delta' I_k - I_k\Delta') + \left(-\frac{\Lambda}{\Delta_{(2)}} - n\right)^3 + (1-2n)(\Delta' I^{k} - I^{k}\Delta')(\Delta' I_k - I_k\Delta') + 2A^{k}_{i}A^{i}_{k} - n\left(+\frac{\Lambda}{\Delta_{(2)}} + n\right)^2 + A^{i}_{i} + A^{i}_{i}\left(\frac{\Lambda}{\Delta_{(2)}} + n\right) + n\left(\frac{\Lambda}{\Delta_{(2)}} + n\right)\Big\}\Delta^3_{(2)}. \tag{3.9}$$

The eigenvalue of I_9 is

$$\kappa^6\Big[\sum_{i=2}^{n} h_{i\,n+1}(h^2_{i\,n+1} + 3nh_{i\,n+1} + 3h_{i\,n+1} - 3ih_{i\,n+1} + 3n^2 + 6n + 3 - 6i - 6ni + 3i^2) - (n - \tfrac{1}{2})\sum_{i=2}^{n} h_{i\,n+1}(h_{i\,n+1} + 2n + 2 - 2i) - \tfrac{1}{2} \times \Big(\sum_{i=2}^{n} h^2_{i\,n+1} + 2\sum_{i\neq j=2}^{n} h_{i\,n+1}h_{j\,n+1}\Big) - n\sum_{i=2}^{n} h_{i\,n+1} - \frac{n^3}{4} - \frac{3n^2}{4}\Big]. \tag{3.10}$$

4. EXPLICIT EVALUATION OF FOURTH AND EIGHTH ORDER INVARIANT OPERATORS OF IO(n)

Following the same procedures as in the previous section, we calculate the fourth order invariant operator of IO(n). We have, for O(n, 1),

$$C_2 = \sum_{i<j=1}^{n} J^2_{ij} - \sum_{i=1}^{n} J^2_{i\,n+1}. \tag{4.1}$$

Equation (4.1) contains the following term in λ:

$$-\lambda^2\Delta_{(2)}. \tag{4.2}$$

Therefore, the fourth order invariant operator I_4 of IO(n) is easily seen to be

$$I_4 = \sum_{i>j=1}^{n} J^2_{ij} - \tfrac{1}{4}\sum_{i=1}^{n}(I_iJ_{ij} + J_{ij}I_i)^2, \tag{4.3}$$

where

$$I_i = I_{i\,n+1}. \tag{4.4}$$

The eigenvalue of I_4 can be easily calculated: For IO($2k$)

$$I_4 = \sum_{i=2}^{k} h_i^2 + \sum_{i=2}^{k-1}(2k - 2i + 1)h_i - (\tfrac{1}{2} - k)^2; \tag{4.5}$$

for IO($2k-1$)

$$I_4 = \sum_{i=2}^{k} h_i^2 + \sum_{i=2}^{k-1}(2k - 2i)h_i - (1 - k)^2. \tag{4.6}$$

For the eighth order invariant operator of IO(n), we have, starting from C_4 of O(n, 1):

$$C_4 = J_{ij}J_{jk}J_{kl}J_{li} - 4J_{ij}J_{jk}J_{k\,n+1}J_{n+1\,i} - 4J^2_{ij} - 4J_{jk}J_{kj}J_{n+1\,j}J_{j\,n+1} + 2J_{n+1\,j}J_{j\,n+1}J_{n+1\,j}J_{j\,n+1} + 2J_{n+1\,j}J_{j\,n+1}J_{n+1\,k}J_{k\,n+1} - (n-1)(n-4)J^2_{n+1\,l}, \tag{4.7}$$

where $i\neq j\neq k\neq l$. We find that (4.7) contains the following terms involving λ:

$$2\kappa^4\lambda^4 + (n-2)^2\kappa^2\lambda^2. \tag{4.8}$$

Therefore, the eighth order invariant operator I_8 of IO(n) is

$$I_8 = \Big[J_{ij}J_{jk}J_{kl}J_{li} + \frac{1}{\Delta_{(2)}}J_{ij}J_{jk}(I_aJ_{ak} + J_{ak}I_a)(I_bJ_{bi} + J_{bi}I_b) - 4J^2_{ij} - \frac{(n-1)(n-4)}{4\Delta_{(2)}}(I_aJ_{al} - J_{al}I_a)(I_bJ_{bl} + J_{bl}I_b) + \frac{J_{jk}J_{kj}}{\Delta_{(2)}}(I_aJ_{aj} + J_{aj}I_a)(I_bJ_{bj} + J_{bj}I_b) + \frac{1}{8\Delta^2_{(2)}} \times (I_aJ_{aj} + J_{aj}I_a)(I_bJ_{bj} + J_{bj}I_b)(I_cJ_{cj} + J_{cj}I_c)(I_dJ_{dj} + J_{dj}I_d) + \frac{1}{8\Delta^2_{(2)}}(I_aJ_{aj} + J_{aj}I_a)(I_bJ_{bj} + J_{bj}I_b)(I_cJ_{ck} + J_{ck}I_c) \times (I_dJ_{dk} + J_{dk}I_d)\Big]\Delta^2_{(2)} \tag{4.9}$$

Eigenvalues of I_8: For IO($2k$)

$$I_8 = \kappa^4\Big\{2\sum_{i=2}^{k} h_i[(h_i + r_i)^3 + (h_i + r_i)^2 r_i + (h_i + r_i)r_i^2 + r_i^3] - (2k-1)\sum_{i=2}^{k}(h_i + 2r_i)h_i + 2(k - \tfrac{1}{2})^3(-k + \tfrac{3}{2})\Big\} \tag{4.10}$$

where $r_i = k + \frac{1}{2} - i$; for IO($2k-1$)

$$I_8 = \kappa^4\Big\{\sum_{i=2}^{k} h_i[(h_i + r_i)^3 + (h_i + r_i)^2 r_i + (h_i + r_i)r_i^2 + r_i^3] - (2k-1)\sum_{i=2}^{k} h_i(h_i + 2r_i) + (k-1)^2(-2k^2 - 6k - 3)\Big\}, \tag{4.11}$$

where

$$r_i = k - i. \tag{4.12}$$

5. EXPLICIT EVALUATION OF ALL INVARIANT OPERATORS OF IU(n) AND IO(n)

From the results of Secs. 3 and 4 we arrive at the following theorems which are applicable to all orders for the invariant operators of IU(n) and IO(n).

Theorem 1: The invariant operators of IU(n) are obtained from those of U(n, 1) by the following substitution:

$$A^{i}_{n+1} \to [\Delta', I^{i}_{n+1}], \quad \text{where } \Delta' = A^{j}_{i}A^{i}_{j}/2\kappa, \tag{5.1}$$

$$A^{n+1}_{i} \to [\Delta', I^{n+1}_{i}], \tag{5.2}$$

$$A^{n+1}_{n+1} \to -\Lambda/\Delta_{(2)} - n. \tag{5.3}$$

Theorem 2: The eigenvalues of the invariant operators of IU(n) are obtained from those of U(n, 1) by the following substitution:

$$h_{1\,n+1} \to -n/2, \tag{5.4}$$

$$h_{n+1\,n+1} \to n/2. \tag{5.5}$$

Note that in the formula for C_p of U(n, 1)[3]

$$C_p = \sum_{i=1}^{n+1} \lambda_i^p \prod_{i \neq j} \frac{\lambda_i - \lambda_j - 1}{\lambda_i - \lambda_j}, \tag{5.6}$$

where

$$\lambda_i = h_i + n + 1 - i, \tag{5.7}$$

the term containing $1/(\lambda_1 - \lambda_{n+1})$ cancels out, giving $(\lambda_1 \to n/2,\ \lambda_{n+1} \to n/2)$

$$\left(\frac{n}{p}\right)^p \prod_{j=2}^{n} \frac{n/2 - \lambda_j - 1}{n/2 - \lambda_j}. \tag{5.8}$$

We now illustrate our theorems by recalculating the eigenvalues of $\Delta'_{(3)}$ obtaining the same result as Chakrabarti without going through the complicated procedures used by him. We have

$$C_1 = \sum_{i=1}^{n} A_i^i + A_{n+1}^{n+1}. \tag{5.9}$$

From (5.9) using Theorem 1, we obtain

$$I_1 = \sum_{i=1}^{n} A_i^i - \frac{\Lambda}{\Delta_{(2)}} - n. \tag{5.10}$$

From (5.10) using Theorem 2, we obtain for the eigenvalue

$$I_1 = \sum_{i=2}^{n} h_i. \tag{5.11}$$

Therefore, from (5.10), (5.11) we obtain

$$\Delta'_{(3)} \equiv \Lambda - \left(\sum_{i=1}^{n} A_i^i\right)\Delta_{(2)} = -(I_1 + n)\Delta_{(2)} = -\kappa^2\left(\sum_{i=2}^{n} h_i + n\right) \tag{5.12}$$

Equation (5.12) agrees with Eq. (2.40) of Chakrabarti.

It is now a simple matter to obtain an invariant operator in $IU(n)$ which is a polynomial in the generators of $IU(n)$, i.e., containing no terms on the denominator. This can be easily achieved by multiplying a suitable power of $\Delta_{(2)}$ with the expression obtained from Theorem 1. The result is that the pth order invariant operator of $U(n, 1)$ corresponds to the $(3p)$th order invariant operator of $IU(n)$.

Theorem 3: The invariant operators of $IO(n)$ are obtained from those of $O(n, 1)$ by the following substitution:

$$J_{j\,n+1} \to \tfrac{1}{2}(I_{i\,n+1}J_{ij} + J_{ij}I_{i\,n+1}). \tag{5.13}$$

Theorem 4: The eigenvalues of the invariant operators of $IO(n)$ are obtained from those of $O(n, 1)$ by the following substitution: For $IO(2k-1)$:

$$h_{2k\,1} \to -k + 1; \tag{5.14}$$

for $IO(2k)$:

$$h_{2k+1\,1} \to \tfrac{1}{2} - k. \tag{5.15}$$

From these two theorems, we find that the $(2p)$th order invariant operator of $O(n, 1)$ corresponds to the $(4p)$th order invariant operator of $IO(n)$. Now the invariant operators of $O(n+1)$ are

$$C_{2p} = J_{i_1 i_2} J_{i_2 i_3} J_{i_3 i_4} \cdots J_{i_{2p} i_1}. \tag{5.16}$$

The invariant operators of $O(n, 1)$ are obtained from (5.16) by replacing $J_{i\,n+1}$ by $iJ_{i\,n+1}$. In the meantime the eigenvalues are the same for both $O(n+1)$ and $O(n, 1)$ and can be expressed as[5-8] follows: for $O(2h+1)$

$$C_k = \sum_{i=1}^{n} p^k \prod_{\substack{i,j=1 \\ j \neq i}}^{n} \frac{p_j - p_i + 1 + \epsilon_{ji}}{p_j - p_i}, \tag{5.17}$$

where

$$n = 2h + 1, \quad p_i = h_i + n - 1 - i,$$
$$\epsilon_{ji} = \delta_{j,n+1-i} - \delta_{i,(n+1)/2} \tag{5.18}$$

For $O(2h)$, (5.17) still holds with

$$n = 2h, \ \epsilon_{ji} = -\delta_{j,n+1-i}. \tag{5.19}$$

Thus all invariant operators and their eigenvalues of $IO(n)$ can be obtained in closed and simple form.

It remains for us to choose a sufficient number of algebraically independent invariant operators so that each irreducible representation of the group is completely specified by these operators. We start with $IU(n)$. It can be seen that an irreducible representation of $IU(n)$ is specified by n mutually independent invariant operators. Since all the eigenvalues are known, it is easy to check that the following invariant operators are mutually independent. Thus we choose:

For $IU(2)$: $\Delta_{(2)},\ \Delta'_{(3)}$;

$IU(3)$: $\Delta_{(2)},\ \Delta'_{(3)},\ I_6$;

$IU(4)$: $\Delta_{(2)},\ \Delta'_{(3)},\ I_6,\ I_9$;

...

$IU(n)$: $\Delta_{(2)},\ \Delta'_{(3)},\ I_6,\ I_9, \ldots, I_{3(n-1)}$.

For $IO(2k+1)$, we need $k+1$ algebraically independent invariant operators, and for $IO(2k)$, we need k algebraically independent invariant operators. Again, since the eigenvalues are known, it can be checked that the following invariant operators are algebraically independent:

For $IO(2k+1)$: $\Delta_{(2)},\ I_4,\ I_8,\ I_{12}, \ldots, I_{4k}$;

for $IO(2k)$: $\Delta_{(2)},\ I_4,\ I_8,\ I_{12}, \ldots, I_{4(k-1)}$.

ACKNOWLEDGMENTS

We wish to thank Moorhead State University, especially Dr. W.B. Treumann and Dr. B.W. McCashland, for financial support of this work.

[1]A. M. Perelomov and V.S. Popov, Yad. Fiz. 3, 924 (1966) [Sov. J. Nucl. Phys. 3, 676 (1966)].
[2]A. M. Perelomov and V.S. Popov, Yad. Fiz. 3, 1127 (1966) [Sov. J. Nucl. Phys. 3, 819 (1966)].
[3]V.S. Popov and A.M. Perelomov, Yad. Fiz. 5, 693 (1967) [Sov. J. Nucl. Phys. 5, 489 (1967)].
[4]M.K.F. Wong and H.Y. Yeh, J. Math. Phys. 16, 1239 (1975).
[5]C.O. Nwachuku and M.A. Rashid, J. Math. Phys. 17, 1611 (1976).
[6]C.O. Nwachuku and M.A. Rashid, J. Math. Phys. 18, 1387 (1977).
[7]S. Okubo, J. Math. Phys. 18, 2382 (1977).
[8]S.A. Edwards, J. Math. Phys. 19, 164 (1978).
[9]A. Chakrabarti, J. Math. Phys. 9, 2087 (1968).
[10]J. Rosen and P. Roman, J. Math. Phys. 7, 2072 (1966).
[11]M.K.F. Wong and H.Y. Yeh, J. Math. Phys. 16, 800 (1975).

Automorphisms of the Bianchi model Lie groups

Alex Harvey

Queens College of the City University of New York, Flushing, New York 11367
(Received 23 February 1977; revised manuscript received 16 February 1978)

A derivation of the group of automorphisms of the Lie group of isometries characteristic of each of the nine Bianchi types of cosmologies is presented.

Among the various categories of cosmological models based on the Einstein equations are the so-called Bianchi models. These models are characterized by homogeneous, nonisotropic spatial hypersurfaces parametrized by time. In a synchronous coordinate system, that is, one in which the time axis is always normal to the hypersurfaces of homogeneity, the metric is

$$ds^2 = -dt^2 + g_{ij}(t)dx^i dx^j, \tag{1}$$

where $i, j = +1, 2, 3$. Space—time based on the possible (three-dimensional) isometries of g_{ij}, determined originally by Bianchi[1] and later applied to general relativity by Taub,[2] comprise the Bianchi models. There are nine distinct isometry groups in all.[3]

The study of these models[3–5] has been very active and fruitful in recent years. Many techniques and tools for this exploration have been developed. One of these, suggested by Heckmann and Schücking,[6] is to use the various groups of automorphisms of the Lie groups descriptive of the Bianchi isometries. This suggestion, however, does not seem to have been taken up. The author, in a current investigation, determined several of these groups of automorphisms, found them to be of sufficient interest to warrant determination of the entire set. It is the purpose of this paper to present them.

Recall briefly a few immediately pertinent if elementary facts[7] concerning Lie groups. Given the infinitesimal operators, X_i, of such a group, the structure constants are given in terms of the commutator

$$[X_i, X_j] = C^k{}_{ij} X_k. \tag{2}$$

These depend in general on the coordinate system in which the basis operators X_i are set and will generally change under a change in coordinate system. The group of automorphisms of a Lie group of isometries is that linear group of coordinate transformations with respect to which the structure constants are invariant. We are then concerned with the set of transformations, $A^i{}_j$, such that

$$C^f{}_{rs} = A^g{}_r A^h{}_s A^{-1f}{}_j C^j{}_{gh}. \tag{3}$$

The group of isometries is a subgroup of and possibly coincident with the group of automorphisms.

Of particular interest is the fact that of the nine Bianchi symmetry groups, types VIII and IX are (semi-) simple, types I—VII are not. It follows from general theorems that for types VIII and IX, all automorphisms are inner automorphisms, and the groups of automorphisms are thus isomorphic to the original symmetry groups.

The equations are cubic in the unknowns, i.e., the elements of $A^i{}_j$, but this may be substantially ameliorated by rewriting, e.g., (3) as

$$A^j{}_f C^f{}_{rs} = A^g{}_r A^h{}_s C^j{}_{gh}. \tag{4}$$

There is no general algorithm for solving such sets of equations, but recognition that $A^i{}_j$ is necessarily nonsingular facilitates greatly the solution. However, the process is quite simple if heuristic. Consequently, only one solution, Au(IV), is presented in detail. For the others, only the results are given.

For the various groups, then, the set of nonvanishing structure constants (following Ref. 3), the resulting set of equations other than those that vanish identically, and the solutions are as follows.

Bianchi I: All structure constants vanish. Consequently, the general linear group on three dimensions constitutes the group of automorphisms.

Bianchi II: $C^1{}_{23} = -C^1{}_{32} = 1$.

$$0 = A^2{}_1 A^3{}_2 - A^3{}_1 A^2{}_2, \tag{5a}$$

$$A^1{}_1 = A^2{}_2 A^3{}_3 - A^3{}_2 A^2{}_3, \tag{5b}$$

$$A^2{}_1 = 0, \tag{5c}$$

$$A^3{}_1 = 0, \tag{5d}$$

$$0 = A^2{}_3 A^3{}_1 - A^3{}_3 A^2{}_1, \tag{5e}$$

$$A^i{}_j = \begin{bmatrix} a & d & e \\ 0 & b & f \\ 0 & g & c \end{bmatrix}. \tag{6}$$

Necessarily $a = bc - fg \neq 0$.

$$A^{-1i}{}_j = \frac{1}{a^2}\begin{bmatrix} a & eg-cd & be-df \\ 0 & ac & -af \\ 0 & -ag & ab \end{bmatrix} \tag{7}$$

It may be noted that $A^{-1i}{}_j$ has, as it ought to, precisely the same structure as $A^i{}_j$.

Bianchi III: $C^1{}_{13} = -C^1{}_{31} = 1$.

$$0 = A^1{}_1 A^3{}_2 - A^3{}_1 A^1{}_2, \tag{8a}$$

$$0 = A^1{}_2 A^3{}_3 - A^3{}_2 A^1{}_3, \tag{8b}$$

$$-A^1{}_1 = A^1{}_3 A^3{}_1 - A^3{}_3 A^1{}_1, \tag{8c}$$

$$A^2{}_1 = 0, \tag{8d}$$

$$A^3{}_1 = 0, \tag{8e}$$

 0022-2488/79/020251-03$01.00

$$A^i{}_j = \begin{bmatrix} a & 0 & e \\ 0 & b & f \\ 0 & 0 & 1 \end{bmatrix}, \tag{9}$$

$$A^{-1i}{}_j = \frac{1}{ab} \begin{bmatrix} b & 0 & -be \\ 0 & a & -af \\ 0 & 0 & ab \end{bmatrix}. \tag{10}$$

Bianchi IV: $C^1{}_{13} = -C^1{}_{31} = 1,\ C^1{}_{23} = -C^1{}_{32} = 1,\ C^2{}_{23} = -C^2{}_{32} = 1.$

$$0 = A^1{}_1A^3{}_2 - A^3{}_1A^1{}_2 + A^2{}_1A^3{}_2 - A^3{}_1A^2{}_2, \tag{11a}$$

$$0 = A^2{}_1A^3{}_2 - A^3{}_1A^2{}_2, \tag{11b}$$

$$A^1{}_1 = A^1{}_1A^3{}_3 - A^3{}_1A^1{}_3 + A^2{}_1A^3{}_3 - A^3{}_1A^2{}_3, \tag{11c}$$

$$A^2{}_1 = A^2{}_1A^3{}_3 - A^3{}_1A^2{}_3, \tag{11d}$$

$$A^3{}_1 = 0, \tag{11e}$$

$$A^1{}_1 + A^1{}_2 = A^1{}_2A^3{}_3 - A^3{}_2A^1{}_3 + A^2{}_2A^3{}_3 - A^3{}_2A^2{}_3, \tag{11f}$$

$$A^2{}_1 + A^2{}_2 = A^2{}_2A^3{}_3 - A^3{}_2A^2{}_3, \tag{11g}$$

$$A^3{}_1 - A^3{}_2 = 0. \tag{11h}$$

Equation (11e) applied to Eq. (11h) shows that $A^3{}_2$ is also zero. These imply that $A^3{}_3$ may not vanish and reduce Eqs. (11a) and (11b) to trivial identities $0 = 0$. The remaining equations are also reduced.

$$A^1{}_1 = A^1{}_1A^3{}_3 + A^2{}_1A^3{}_3, \tag{11c$'$}$$

$$A^2{}_1 = A^2{}_1A^3{}_3, \tag{11d$'$}$$

$$A^1{}_1 + A^1{}_2 = A^1{}_2A^3{}_3 + A^2{}_2A^3{}_3, \tag{11f$'$}$$

$$A^2{}_1 + A^2{}_2 = A^2{}_2A^3{}_3. \tag{11g$'$}$$

Equation (11d′) implies that either

$$A^2{}_1 \neq 0, \quad A^3{}_3 = 1 \tag{12}$$

or

$$A^2{}_1 = 0, \quad A^3{}_3 \text{ arbitrary}. \tag{13}$$

If the first case is applied to Eq. (13c′) the result is

$$A^1{}_1 = A^1{}_1 + A^2{}_1, \tag{14}$$

which is inconsistent. Therefore, the alternate equation survives with $A^2{}_1 = 0$ and $A^3{}_3$ arbitrary. The remaining equations are

$$A^1{}_1 = A^1{}_1A^3{}_3, \tag{13c$''$}$$

$$A^1{}_1 = A^2{}_2A^3{}_3, \tag{13f$''$}$$

$$A^2{}_2 = A^2{}_2A^3{}_3. \tag{13g$''$}$$

Clearly, $A^3{}_3 = 1$ and $A^1{}_1 = A^2{}_2$ and are otherwise arbitrary. Thus

$$A^i{}_j = \begin{bmatrix} a & d & e \\ 0 & a & f \\ 0 & 0 & 1 \end{bmatrix} \tag{15}$$

$$A^{-1i}{}_j = \frac{1}{a^2} \begin{bmatrix} a & -d & df - ae \\ 0 & a & -af \\ 0 & 0 & a^2 \end{bmatrix}. \tag{16}$$

Bianchi V: $C^1{}_{13} = -C^1{}_{31} = 1,\ C^2{}_{23} = -C^2{}_{32} = 1.$

$$0 = A^1{}_1A^3{}_2 - A^3{}_1A^1{}_2, \tag{17a}$$

$$0 = A^2{}_1A^3{}_2 - A^3{}_1A^2{}_2, \tag{17b}$$

$$A^1{}_2 = A^1{}_2A^3{}_3 - A^3{}_2A^1{}_3, \tag{17c}$$

$$A^2{}_2 = A^2{}_2A^3{}_3 - A^3{}_2A^2{}_3, \tag{17d}$$

$$-A^1{}_1 = A^1{}_3A^3{}_1 - A^3{}_3A^1{}_1, \tag{17e}$$

$$-A^2{}_1 = A^2{}_3A^3{}_1 - A^3{}_3A^2{}_1, \tag{17f}$$

$$A^i{}_j = \begin{bmatrix} a & d & e \\ k & b & f \\ 0 & 0 & 1 \end{bmatrix}, \tag{18}$$

$$A^{-1i}{}_j = \frac{1}{ab - kd} \begin{bmatrix} b & -d & df - be \\ -k & a & ek - af \\ 0 & 0 & ab - dk \end{bmatrix}. \tag{19}$$

Necessarily $\dfrac{1}{ab - kd} \neq 0$.

Bianchi VI: $C^1{}_{13} = -C^1{}_{31} = 1,\ C^2{}_{23} = -C^2{}_{32} = h,\ h \neq 0, 1.$

$$0 = A^1{}_1A^3{}_2 - A^3{}_1A^1{}_2, \tag{20a}$$

$$0 = A^2{}_1A^3{}_2 - A^3{}_1A^2{}_2, \tag{20b}$$

$$hA^1{}_2 = A^1{}_2A^3{}_3 - A^3{}_2A^1{}_3, \tag{20c}$$

$$A^2{}_2 = A^2{}_2A^3{}_3 - A^3{}_2A^2{}_3, \tag{20d}$$

$$A^1{}_1 = A^1{}_1A^3{}_3 - A^1{}_3A^3{}_1, \tag{20e}$$

$$A^2{}_1 = h(A^2{}_1A^3{}_3 - A^2{}_3A^3{}_1) \tag{20f}$$

Apart from the factors h these equations are identical with those for Bianchi V.

$$A^i{}_j = \begin{bmatrix} a & 0 & e \\ 0 & b & f \\ 0 & 0 & 1 \end{bmatrix}, \tag{21}$$

$$A^{-1i}{}_j = \frac{1}{ab} \begin{bmatrix} b & 0 & -be \\ 0 & a & -af \\ 0 & 0 & ab \end{bmatrix}. \tag{22}$$

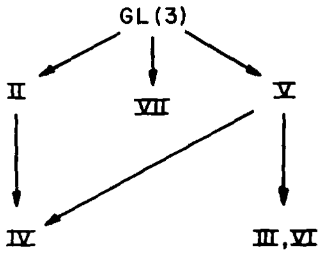
Bianchi VII: $C^1{}_{23} = -C^1{}_{32} = -1,\ C^2{}_{23} = -C^2{}_{32} = h,\ h^2 < 4,\ C^2{}_{13} = -C^2{}_{31} = 1.$

$$A^i{}_j = \begin{bmatrix} a & b & e \\ -b & a - hb & f \\ 0 & 0 & 1 \end{bmatrix}, \tag{23}$$

$$A^{-1i}{}_j = (a^2 - hab + b^2)^{-1} \begin{bmatrix} a - hb & -b & bf - e(a - hb) \\ b & a & -af - eb \\ 0 & 0 & a^2 - hab + b^2 \end{bmatrix}. \tag{24}$$

Bianchi VIII: $C^1{}_{23} = -C^1{}_{32} = -1,\ C^2{}_{31} = -C^2{}_{13} = 1,\ C^3{}_{12} = -C^3{}_{21} = 1.$

Bianchi IX: $C^1{}_{23} = -C^1{}_{32} = 1,\ C^2{}_{31} = -C^2{}_{13} = 1,\ C^3{}_{12} = -C^3{}_{21} = 1.$



SUBGROUP STRUCTURE

FIG. 1. Subgroup structures.

As noted earlier the groups of automorphisms for Bianchi types VIII and IX are isomorphic to the Bianchi type VIII and IX symmetry groups respectively. These are just the three-dimensional rotation group for type IX and the "2 + 1" Lorentz group for type VIII. For the choice of axes for this latter case, X^2 and X^3 are spacelike and X^1 is timelike.

It turns out that the automorphism groups for types III and VI are identical. This is not too surprising inasmuch as type III symmetry is obtained from type VI if, in the latter, $C^2{}_{23} = 0$. Also, the various groups of automorphisms have subgroup relations with each other and all are subgroups of GL(3), the general linear group on three dimensions. The specific relations are shown in Fig. 1 (arrows point from larger group to subgroup).

ACKNOWLEDGMENT

Interest in this problem was stimulated by conversations with Professor Englebert Schücking. Particular thanks are due to Professor Schücking for comments essential to the successful determination of these automorphisms.

[1]L. Bianchi, Mem. Soc. It. Sci. (Dei. XL) (3) **11**, 267 (1897).
[2]A. H. Taub, Ann. Math. 53, 472 (1951).
[3]M. P. Ryan, Jr. and L. C. Shepley, *Homogeneous Relativistic Cosmologies* (Princeton U. P., Princeton, N. J., 1975).
[4]G. F. R. Ellis and M. A. H. MacCallum, Commun. Math. Phys., 12, 108 (1969).
[5]M. A. H. MacCallum, *Cargese Lectures in Physics* (Gordon and Breach, New York, 1973), Vol. 6.
[6]O. Heckmann and E. Schücking, in *Gravitation*, edited by L. Witten (Wiley, New York, 1962).
[7]See, e. g., M. Hammermesh, *Group Theory* (Addison-Wesley, Reading, Mass. 1962).

Perturbative solution of the Percus–Yevick integral equation for a general class of intermolecular potential

M. Chen

Department of Mathematics, Vanier College, St. Laurent, Quebec
(Received 28 November 1977; revised manuscript received 27 February 1978)

A qualitative investigation of the Percus–Yevick integral equation by perturbation method is discussed for a general class of intermolecular potential. Under some general assumptions it is proved that the Percus–Yevick integral equation has a unique solution when the particle density ρ is in the region $0<\rho<0.33$, and a divergent solution when ρ is greater than 0.33. Moreover, the perturbation series is absolutely and uniformly convergent if the supremum norms of the nth order solutions are less than or equal to $n!$

I. INTRODUCTION

Consider a system of N molecules in a volume V and at temperature T. Suppose that the potential energy $\phi(\mathbf{r}_1,\mathbf{r}_2,\ldots,\mathbf{r}_N)$ of this system can be written as the sum of pairwise intermolecular potential $U(r_{ij})$,

$$\phi(\mathbf{r}_1,\mathbf{r}_2,\ldots,\mathbf{r}_N)=\sum_{1=i<j}^{N} U(r_{ij}),$$

where $\mathbf{r}_i$ is the position of the ith molecule and $r_{ij}=|\mathbf{r}_i-\mathbf{r}_j|$.

Define the configurational probability function by

$$p^{(N)}(\mathbf{r}_1,\mathbf{r}_2,\ldots,\mathbf{r}_N)=e^{-\beta\Phi}\left(\int e^{-\beta\Phi}d\mathbf{r}_1 d\mathbf{r}_2\cdots d\mathbf{r}_N\right)^{-1},$$

where $\beta=1/K_bT$, K_b is the Boltzmann constant.

The probability distribution functions of lower orders can then be obtained from $p^{(N)}(\mathbf{r}_1,\mathbf{r}_2,\ldots,\mathbf{r}_N)$. In particular, the probability of finding a molecule in a volume element $d\mathbf{r}_1$ at $\mathbf{r}_1$ and another molecule in $d\mathbf{r}_2$ at $\mathbf{r}_2$ is given by

$$n^{(2)}(\mathbf{r}_1,\mathbf{r}_2)d\mathbf{r}_1\,d\mathbf{r}_2=\frac{N!}{(N-2)!}\left(\int e^{-\beta\Phi}d\mathbf{r}_3\cdots d\mathbf{r}_N\right)\times d\mathbf{r}_1 d\mathbf{r}_2\left(\int e^{-\beta\Phi}d\mathbf{r}_1\cdots d\mathbf{r}_N\right)^{-1}.$$

For a simple fluid the intermolecular forces are central forces and consequently $n^{(2)}(\mathbf{r}_1,\mathbf{r}_2)$ depends only on the distance r_{12} between molecules 1 and 2. When r_{12} is large, we can obtain $n^{(2)}(r_{12})\sim N(N-1)/V^2\sim\rho^2$ in the thermodynamic limit $N\to\infty$, $V\to\infty$ but $\rho=\lim_{N\to\infty V\to\infty}(N/V)<\infty$.

The radial distribution function $g(r_{12})$ of a simple fluid is defined by

$$g(r_{12})=\frac{1}{\rho^2}n^{(2)}(r_{12}).$$

Since $g(r_{12})\to 1$ as $r_{12}\to\infty$, we define the total correlation function $h(r_{12})$ between molecules 1 and 2 by

$$h(r_{12})=g(r_{12})-1.$$

Following Ornstein and Zernike,[1] the total correlation function $h(r_{12})$ can be written as the sum of the direct correlation function $C(r_{12})$ and an indirect correlation function which accounts for the correlation of molecules 1 and 2 through a third molecule,

$$h(r_{12})=C(r_{12})+\rho\int h(r_{13})C(r_{23})\,d\mathbf{r}_3. \tag{1}$$

The convolution relation (1) is usually called the Orstein–Zernike (O.Z.) relation, which can be considered as the definition of $C(r_{12})$.

In order to obtain thermodynamic properties of a fluid it is essential to know the radial distribution function $g(r_{12})$.[2] Unfortunately, so far there is no exact theory for $g(r_{12})$. Several approximate theories have been proposed in the past. Based on numerical calculations the Percus–Yevick[3] (P.Y.) approximation seems to be the most successful theory.

Let

$$f(r)=e^{-\beta U(r)}-1,$$
$$y(r)=e^{\beta U(r)}g(r),\quad r=r_{12}.$$

The P.Y. approximation assumes that $C(r)$ vanishes outside the range of the intermolecular potential $U(r)$, specifically, $C(r)=f(r)\,y(r)$.

The P.Y. approximation together with the O.Z. relation (1) forms an integral equation for $g(r)$ in terms of $y(r)$,

$$y(r)=1+\rho\int y(r')f(r')\{\exp[-\beta U(r-r')]y(r-r')-1\}d\mathbf{r}',$$

which is called the P.Y. integral equation.

The P.Y. integral equation has been solved by Wertheim,[4] Thiel,[5] Baxter,[6] and recently by Chen[7] for the hard sphere potential. For a more realistic potential, Wertheim[4] had considered an attractive potential with a range less than the diameter of hard spheres. Unfortunately, his results were not very conclusive. On the other hand, Groeneveld[8] had studied the existence and analytic properties of solutions to the P.Y. integral equation by considering the series expansion of $C(r)$ and $y(r)$ in density ρ. Under the assumptions

(i) $A=\sup_{\mathbf{r}_1,\mathbf{r}_2}\{\exp[-\beta U(|\mathbf{r}_1-\mathbf{r}_2|)]\}<\infty$,

(ii) $B=\sup_{\mathbf{r}_1}\left(\int\{\exp[-\beta U(|\mathbf{r}_1-\mathbf{r}_2|)]-1\}d\mathbf{r}_2\right)<\infty$,

Groeneveld proved that there existed a unique solution of $C(r)$ and $y(r)$ in series of ρ which were analytic in the region $B|\rho|<(4A)^{-1}$.

Recently Watts[9] had numerically solved the P.Y. integral equation by truncating the Lennard–Jones potential at $r=3.5\sigma$, $r=5\sigma$, and $r=6\sigma$ (σ: hard sphere diameter). It was found that there existed a phase transition with critical density close to $\rho_c=0.27$ and a critical temtemperature dependent upon the truncation of the potential. Outside the critical region the P.Y. integral equation had two solutions. The solution in the higher density region was an unphysical one.

The purpose of this paper is to make a qualitative investigation of the P.Y. integral equation for a general class of intermolecular potential by perturbation method. We assume that the potential has a hard core and a weak attractive tail with infinite range. In contrast to Groeneveld's method, the attractive potential is considered as a perturbation on the repulsive potential.[10] The perturbation series is constructed by making use of Baxter's relations[6] (B.R.) together with the P.Y. assumption. A set of coupled integral-differential equations is then obtained. Under some general conditions we prove that the P.Y. integral equation has a unique solution when $0<\eta<0.175$ and $1<\eta<2.66$, and a divergent solution when $0.175<\eta<1$ and $\eta>2.66$, where $\eta=\frac{1}{6}\pi\rho$, ρ is the particle density. It is also shown that the perturbation series is absolutely and uniformly convergent if the supremum norms of the nth order solutions are less than or equal to $n!$. The method discussed in this paper is quite different from that of Groeneveld; However, the basic ideas are very similar.

II. PERTURBATION SERIES

Let $h(w)=\int e^{i\mathbf{w}\cdot\mathbf{r}}h(r)d\mathbf{r}$ and $S(r)=\int_r^\infty t(t)dt$. It has been proved[11] that the O.Z. relation (1) can be transformed into the following Baxter's relations (B.R.) if and only if $h(w)$ is bounded for real w:

$$rc(r)=-Q'(r)+12\,\eta\int_r^\infty Q'(t)Q(t-r)\,dt,\quad r\geq 0, \tag{2}$$

$$rh(r)=-Q'(r)+12\eta\int_0^\infty (r-t) \times h(|r-t|)\,Q(t)dt,\quad r\geq 0, \tag{3}$$

where $Q(r)$ is a continuous, bounded function on $[0,\infty)$, and that $Q(r)\to 0$, $|S(r)|\to ae^{-\delta r}$ as $r\to\infty$ (a,δ are real numbers).

In terms of the P.Y. assumption $c(r)=f(r)y(r)$ and Baxter's relations (2), (3), the P.Y. integral equation can be rewritten as the following coupled integral-differential equations for $r\geq 0$:

$$\begin{aligned} r\,c(r)&=-Q'(r)+12\eta\int_r^\infty Q'(t)Q(t-r)dt,\\ r\,h(r)&=-Q'(r)+12\eta\int_0^\infty (r-t)\,h(|r-t|)\,Q(t)\,dt,\\ c(r)&=f(r)\,y(r).\end{aligned} \tag{4}$$

Consider the intermolecular potential[12]

$$u(r)=u_0(r)-\lambda\xi v(r), \tag{5}$$

where

$$u_0(r)=\begin{cases}\infty, & r<1,\\ 0, & r\geq 1,\end{cases}$$

$$v(r)=\begin{cases}0, & r\leq 1,\\ \text{a positive smooth function for } r\geq 1 \text{ which monotonically decreases to 0 faster than } r^{-5} \text{ as } r\to\infty,\end{cases}$$

ξ denotes the maximum of the physical tail potential so that $\operatorname{Max}|v(r)|=1$, and $0\leq|\lambda|\leq 1$.

If the attractive potential $-\xi v(r)$ is considered as a perturbation on the hard sphere potential $u_0(r)$ we can obtain a series expansion in $\lambda\beta\xi$ for $f(r)$,

$$\begin{aligned} f(r)&=\exp[-\beta u_0(r)]\cdot\exp[\lambda\beta\xi v(r)]-1\\ &=f_0(r)+\sum_{n=1}^\infty \frac{\lambda^n}{n!}f_n(r),\end{aligned}$$

where

$$\begin{aligned} f_0(r)&=\exp[-\beta u_0(r)]-1,\\ f_n(r)&=\exp[-\beta u_0(r)]\cdot[\beta\xi v(r)]^n.\end{aligned}$$

Similarly we can write the following perturbation series expansion:

$$Q(r)=Q_0(r)+\sum_{n=1}^\infty\frac{1}{n!}(\lambda\beta\xi)^n Q_n(r), \tag{6}$$

$$y(r)=y_0(r)+\sum_{n=1}^\infty\frac{1}{n!}(\lambda\beta\xi)^n y_n(r), \tag{7}$$

$$h(r)=h_0(r)+\sum_{n=1}^\infty\frac{1}{n!}(\lambda\beta\xi)^n h_n(r), \tag{8}$$

$$c(r)=c_0(r)+\sum_{n=1}^\infty\frac{1}{n!}(\lambda\beta\xi)^n c_n(r), \tag{9}$$

where $(\beta\xi)^{-1}=K_bT/\xi$ is the reduced temperature, the subscript "0" in $Q_0(r)$, $y_0(r)$, $h_0(r)$, $c_0(r)$ denotes the unperturbed system with hard sphere potential $u_0(r)$, and

$$h_n(r)=\exp[-\beta u_0(r)]\cdot\sum_{i=0}^n\binom{n}{i}[v(r)]^i y_{n-i}(r),\quad n\geq 1,$$

$$c_n(r)=\exp[-\beta u_0(r)]\cdot\sum_{i=0}^n\binom{n}{i}[v(r)]^i y_{n-i}(r)-y_n(r),\quad n\geq 1.$$

From (4) and (6)–(9) we can obtain the following results:

$$\begin{aligned} r\,c_0(r)&=-Q_0'(r)+12\eta\int_r^1 Q_0'(t)\times Q_0(t-r)dt,\quad 0\leq r\leq 1,\\ r\,h_0(r)&=-Q_0'(r)+12\eta\int_0^1(r-t)\times h(|r-t|)\,Q_0(t),\quad r\geq 0,\\ c_0(r)&=f_0(r)y_0(r),\quad r\geq 0,\end{aligned} \tag{10}$$

and

$$r\,c_n(r)=-Q_n'(r)+12\eta\int_r^\infty\sum_{i=0}^n\binom{n}{i}Q_i'(t)\times Q_{n-i}(t-r)dt,\quad r\geq 1,\ n\geq 1, \tag{11}$$

$$\begin{aligned} r\,h_n(r)=-Q_n'(r)+12\eta\int(r-t)\Bigg[&\sum_{j=0}^{n-1}\binom{n}{j}Q_j(t)\times\{\exp[-\beta u_0(|r-t|)]\\ &\times\sum_{i=0}^{n-j}\binom{n-j}{i}[v(|r-t|)]^i y_{n-j-i}(|r-t|)\}\\ &+h_0(|r-t|)Q_n(t)\Bigg]dt,\quad r\geq 1,\ n\geq 1.\end{aligned} \tag{12}$$

Note that (10) is the P.Y. integral equation for a system of hard spheres. The solution of (10) is well known.[4-7]

Due to the nature of the intermolecular potential in (5), (11), and (12) can be further simplified. After some lengthly derivations we finally obtain the following results:

$$Q_n'(r)=A_n(r)-12\eta\int_0^\infty (r-t)Q_n(t)\,dt + 12\eta\int_{r+1}^\infty (r-t)y_0(t-r)Q_n(t)dt, \quad 0\le r<1, \tag{13}$$

$$Q_n'(r)=B_n(r)+12\eta\int_r^{r+1} Q_n'(t)Q_0(t-r)dt, \quad r\ge 1, \tag{14}$$

$$Y_n(r)=Q_n'(r)-12\eta\sum_{i=1}^{n-1}\binom{n}{i}\int_r^\infty Q_i'(t)Q_{n-i}(t-r)dt - 12\eta\int_r^1 Q_0'(t)Q_n(t-r)dt - 12\eta\int_r^{r+1}Q_n'(t) \times Q_0(t-r)dt, \quad 0\le r<1, \tag{15}$$

$$Y_n(r)=D_n(r)+12\eta\int_1^r Q_0(t-r)Y_n(t)dt, \quad 1\le r\le 2, \tag{16}$$

$$Y_n(r)=E_n(r)+12\eta\int_0^1 Q_0(t)Y_n(r-t)dt, \quad r\ge 2, \tag{17}$$

where $Y_n(r)=ry_n(r)$, $A_n(r)$, $B_n(r)$, $D_n(r)$ and $E_n(r)$ are functions of $Q_m(r)$ and $Y_m(r)$ for $m<n$, so that in the nth order perturbation, they can be considered as known functions.

Due to their complexity, the detailed expressions of $A_n(r)$, $B_n(r)$, $D_n(r)$, and $E_n(r)$ are omitted since we will not need them in subsequent discussions.

By considering the attractive potential as a perturbation on the hard sphere potential we have constructed a set of perturbation series in inverse temperature $\beta\xi$ expansion. From the P.Y. integral equation (4) we then obtain a set of coupled integral-differential equations (10) and (13)–(17). From (13)–(17), we note that the method of solving these equations for the nth order perturbation can be described by the following procedure:

$$B_n(r)\to Q_n'(r) \text{ for } r\ge 1 \to Q_n'(r)$$

$$\text{for } r<1 \to \begin{cases} D_n(r), & \text{for } 1\le r\le 2,\\ E_n(r), & \text{for } r\ge 2,\\ Y_n(r), & \text{for } 0\le r<1,\end{cases} \to \begin{cases} Y_n(r), & \text{for } 1\le r\le 2,\\ Y_n(r), & \text{for } r\ge 2.\end{cases}$$

III. SOLUTIONS OF PERTURBATION SERIES

In this section, for simplicity, we confine our discussion to solutions of the first order perturbation. The methods discussed can be applied to all orders of perturbation.

(i) Solution of $Q_1(r)$ for $r\ge 1$:

$$Q_1'(r)=B_1(r)+12\eta\int_r^{r+1} Q_0(t-r)Q_1'(t)dt, \quad r\ge 1. \tag{18}$$

We can rewrite (18) as

$$\phi(r)=\alpha(r)+\mu\int_r^{r+1}K(t-r)\phi(t)dt = \alpha(r)+\mu\int_0^1 K(t)\phi(t+r)dt, \quad r\ge 1, \tag{19}$$

where $\phi(r)=Q_1'(r)$, $\alpha(r)=B_1(r)=-rv(r)y_0(r)$, $K(r)=Q_0(r)$, and $\mu=12\eta$.

Let $I_2=[1,\infty)$. Consider the space $C_B(I_2)$ of all continuous and bounded functions defined on I_2 which approach to zero at least as fast as r^{-4} as $r\to\infty$, i.e., $C_B(I_2)=[f\,|\,f$ continuous and bounded on I_2 and that $\lim_{r\to\infty} r^4 f(r)=\text{const.}]$. Suppose that $C_B(I_2)$ is endowed with a sup norm and d is the metric function defined by

$$d(f_1,f_2)=\sup_{r\in I_2}|f_2(r)-f_1(r)|, \quad f_1, f_2\in C_B(I_2).$$

Then $C_B(I_2)$ together with the metric function d forms a complete metric space.[13]

Let ϕ be an arbitrary element of $C_B(I_2)$. We define an operator T on $C_B(I_2)$ by

$$(T\phi)(r)=\alpha(r)+\mu\int_0^1 K(t)\phi(t+r)dt, \quad \phi\in C_B(I_2). \tag{20}$$

Since $\alpha(r)$ is continuous and $\lim_{r\to\infty}r^4\alpha(r)=0$, $\alpha(r)$ is in $C_B(I_2)$. But $K(t)$ is continuous on $I_1=[0,1]$,[6] thus $(T\phi)(r)\in C_B(I_2)$, and T therefore transforms $C_B(I_2)$ into itself. In order to be able to apply fixed point theorem on $C_B(I_2)$, we next consider the conditions so that T is contractive.

Suppose ϕ_1 and ϕ_2 are two arbitrary elements of $C_B(I_2)$. Then

$$d(T\phi_1,T\phi_2)\le\mu\int_0^1|K(t)|\,dt\cdot d(\phi_1,\phi_2),$$

and T is a contractive mapping if $\mu\int_0^1|K(t)|dt<1$.

Now that $C_B(I_2)$ is a complete metric space and T is a contractive mapping on $C_B(I_2)$, by fixed point theorem, there exists a unique ϕ in $C_B(I_2)$ such that $T\phi=\phi$.

Proposition 1: Equation (19) has a unique solution which is an element of $C_B(I_2)$ if $\mu\int_0^1|K(t)|dt<1$.

Suppose $\{\phi_0,\phi_1,\phi_2,\ldots,\phi_n,\ldots\}$ is a sequence of functions in $C_B(I_2)$ defined by the following relations:

$$\phi_0(r)=\alpha(r),$$

$$\phi_n(r)=T\phi_{n-1}(r) = \alpha(r)+\mu\int_0^1 K(t)\phi_{n-1}(t+r)dt, \quad n\ge 1.$$

Let $\psi_0(r)=\phi_0(r)=\alpha(r)$ and $\phi_n(r)-\phi_{n-1}(r)=\mu^n\psi_n(r)$, $n\ge 1$. Then

$$\phi_n(r)=\sum_{m=0}^n \mu^m\psi_m(r),$$

where

$$\psi_m(r)=\left(\int_0^1\cdots\int_0^1\prod_{i=1}^m K(t_i)\alpha(r+t_1+t_2+t_3+\cdots+t_m)dt_1\cdots dt_m\right), \quad m\ge 1. \tag{21}$$

Suppose $\gamma=\sup_{r\in I_2}|\alpha(r)|$. It then follows from (21) that

$$\sup_{r\in I_2}|\psi_n(r)|\le\gamma\left(\int_0^1|K(t)|\,dt\right)^n.$$

Hence, if $\mu\int_0^1|K(t)|dt<1$, the series $\phi(r)=\sum_{n=0}^n\mu^n\psi_n(r)$ is absolutely and uniformly convergent for $r\in I_2$.

Corollary: Suppose $\mu\int_0^1|K(t)|dt<1$. The solution of (19) can be expressed as an absolutely and uniformly convergent series $\phi(r)=\alpha(r)+\sum_{n=1}^\infty\mu^n\psi_n(r)$, where $\psi_n(r)$ is given by (21).

In view of the asymptotic condition $\lim_{r\to\infty}r^4Q_1'(r)=\text{const}$, we can obtain a unique continuous and bounded function $Q_1(r)$ defined by

$$Q_1(r) = -\int_r^\infty \phi(t)dt \tag{22}$$

which satisfies the asymptotic condition $\lim_{r\to\infty} r^3 Q_1(r) = \text{const}$.

(ii) Solution of $Q_1(r)$ for $0 \le r \le 1$:

$$Q_1'(r) = -12\eta \int_0^\infty (r-t)Q_1(t)dt + 12\eta \int_{r+1}^\infty (r-t) \times y_0(t-r)Q_1(t)\,dt, \quad 0 \le r < 1. \tag{23}$$

Let

$$l_1 = 12\eta \int_0^1 tQ_1(t)dt, \tag{24}$$

$$m_1 = -12\eta \int_0^1 Q_1(t)dt, \tag{25}$$

and

$$\beta_1(r) = 12\eta \int_{r+1}^\infty (r-t)y_0(t-r)Q_1(t)dt - [12\eta \int_1^\infty Q_1(t)dt]\cdot r + 12\eta \int_1^\infty tQ_1(t)dt.$$

We can rewrite (23) as

$$Q_1'(r) = \beta_1(r) + m_1 r + l_1.$$

Hence

$$Q_1(r) = \int \beta_1(r)dr + \frac{m_1}{2} r^2 + l_1 r + p_1, \quad 0 \le r < 1. \tag{26}$$

In order to determine l_1, m_1, and p_1 in (26) we now assume that Q_1 is continuous at $r=1$. The boundary condition of continuity at $r=1$ together with (24), (25) provides us with three linear equations in three unknowns. Thus we can obtain a unique continuous function $Q_1(r)$ which satisfies (23) for $0 \le r \le 1$ and (18) for $r \ge 1$. By Stone–Weierstrass theorem Q_1 can be uniformly approximated by functions of the form $e^{-\alpha r}p(r)$ where $\alpha > 0$ and $p(r)$ is a polynomial. We summarize our results in the following:

Proposition 2: Suppose $\mu \int_0^1 |K(t)|dt < 1$. Then there exists a unique continuous and bounded function Q_1 which satisfies (23) for $0 \le r \le 1$ and (18) for $r \ge 1$. Furthermore, it is square integrable on $[0,\infty)$, and can be uniformly approximated by Laguerre functions $e^{-\alpha r}p(r)$, where $\alpha > 0$ and $p(r)$ is a polynomial.

(iii) Solution of $Y_1(r)$ for $0 \le r < 1$:

$$Y_1(r) = Q_1'(r) - 12\eta \int_r^1 Q_0'(t)Q_1(t-r)dt - 12\eta \int_r^{r+1} Q_0(t-r)Q_1'(t)dt, \quad 0 \le r < 1. \tag{27}$$

Once $Q_1(r)$ is known, straightforward integration of (27) yields $Y_1(r)$ which is continuous on $[0,1)$.

(iv) Solution of $Y_1(r)$ for $1 \le r \le 2$:

$$Y_1(r) = D_1(r) + 12\eta \int_1^r Q_0(r-t)Y_1(t)dt, \quad 1 \le r \le 2. \tag{28}$$

Since Q_0 is a quadratic function we can transform (28) into a third order linear differential equation

$$Y_1'''(r) + \frac{6\eta}{1-\eta} Y_1''(r) + \frac{18\eta^2}{(1-\eta)^2} Y_1'(r) - \frac{12\eta(1+2\eta)}{(1-\eta)^2} Y_1(r) = D_1'''(r) \tag{29}$$

with the following boundary condition:

$$Y_1(1) = D_1(1),$$
$$Y_1'(1) = D_1'(1) - \frac{6\eta}{1-\eta} D_1(1), \tag{30}$$
$$Y_1''(1) = D_1''(1) - \frac{6\eta}{1-\eta} D_1'(1) + \frac{18\eta^2}{(1-\eta)^2} D_1(1).$$

The general solution of the homogeneous equation of (29) is

$$Y_1(r) = \sum_{i=1}^3 \gamma_i \exp(t_i r), \tag{31}$$

where γ_i's are constants to be determined from (30), and the t_i's are roots of

$$R(t) = t^3 + \frac{6\eta}{1-\eta} t^2 + \frac{18\eta^2}{(1-\eta)^2} t - \frac{12\eta(1+2\eta)}{(1-\eta)^2} = 0.$$

Let $Y_1(r) = \sum_{i=1}^3 \gamma_i(r)\exp(t_i r)$. $Y_1(r)$ will be a solution of (29) if and only if

$$\sum_{i=1}^3 \gamma_i'(r)\exp(t_i r) = 0,$$
$$\sum_{i=1}^3 \gamma_i'(r)t_i \exp(t_i r) = 0, \tag{32}$$
$$\sum_{i=1}^3 \gamma_i'(r)t_i^2 \exp(t_i r) = 0.$$

These equations can be solved for $\gamma_i'(r)$ by Cramer's rule, and the results integrated to give $\gamma_i(r)$. From (30) a unique solution for $Y_1(r)$ can be obtained. The solution is of class C^2 on $[1,2]$.

(v) Solution of $Y_1(r)$ for $r \ge 2$:

$$Y_1(r) = E_1(r) + 12\eta \int_0^1 Y_1(r-t)Q_0(t)dt = E_1(r) + 12\eta \int_{r-1}^r Q_0(r-t)Y_1(t)dt, \quad r \ge 2. \tag{33}$$

It can be examined that $D_1(2) = E_1(2)$, $D_1'(2) = E_1'(2)$, and $D_1''(2) = E_1''(2)$. We have to find a Y_1 which satisfies (33) for $r \ge 2$, and for $1 \le r \le 2$, it is given by the solution of (29).

We can transform (33) into the following third order retarded linear nonhomogeneous differential-difference equations,

$$Y_1'''(r) + \frac{6\eta}{1-\eta} Y_1''(r) + \frac{18\eta^2}{(1-\eta)^2} Y_1'(r) - \frac{12\eta(1+2\eta)}{(1-\eta)^2} Y_1(r) + \frac{12\eta}{(1-\eta)^2}[(1+2\eta)Y_1(r-1) + (1+\tfrac{1}{2}\eta)Y_1'(r-1)] = E_1'''(r), \quad r \ge 2, \tag{34}$$

with initial condition given by the solution of (29) for $1 \le r \le 2$.

Following the standard continuation method of solving differential-difference equation,[14] the following result can be proved:

Proposition 3: There exists a unique $Y_1(r)$ of class C^2 on $[2,\infty)$ which satisfies (34) for $r \ge 2$ and the initial condition given by the solution of (29) for $1 \le r \le 2$.

By generalizing Theorem 3.5 of Bellman–Cooke,[15] we can obtain the exponential bound[7] for $Y_1(r)$, $|Y_1(r)| \leq K_1 \exp(K_2 r)$, where K_1, K_2 are positive constants. Since $\lim_{r\to\infty} E_1(r) = 0$, we can take the Laplace transform of (33).

Let

$$\bar{Y}_1(s) = \int_2^\infty Y_1(r) e^{-sr} dr,$$

$$\bar{E}_1(s) = \int_2^\infty E_1(r) e^{-sr} dr,$$

$$\bar{F}_1(s) = \bar{E}_1(s) + 12\eta \int_0^1 \left[\int_{2-v}^2 Y_1(u) e^{-su} du\right] \times Q_0(v) e^{-sv} dv,$$

$$H(s) = 1 - 12\eta \int_0^1 Q_0(r) e^{-sr} dr = s^{-3}[R(s) + L(s) e^{-s}],$$

where

$$R(s) = s^3 + \frac{6\eta}{1-\eta} s^2 + \frac{18\eta^2}{(1-\eta)^2} s - \frac{12\eta(1+2\eta)}{(1-\eta)^2},$$

$$L(s) = 12\eta(1-\eta)^{-2}[(1+2\eta) + (1+\tfrac{1}{2}\eta)s].$$

The Laplace transform of (33) yields

$$\bar{Y}_1(s) = \bar{F}_1(s)\,[H(s)]^{-1}. \tag{35}$$

Since all roots of $H(s) = 0$ lie in the left-hand plane except for the triple roots at the origin,[7] we can arrange the roots in order of nondecreasing absolute value with roots of equal absolute value put in any prescribed order. Let $\{S_n\}$ be a sequence of roots so arranged. The inverse Laplace transform of (35) then yields

$$Y_1(r) = \sum_{n=1}^\infty P_{n-1}(r) \exp(s_n r), \tag{36}$$

where $P_{n-1}(r)\exp(s_n r)$ denotes the residue of $e^{sr}\bar{F}_1(s)$ $[H(s)]^{-1}$ at a zero s_n of $H(s)$ and $P_{n-1}(r)$ is a polynomial at most of degree $n-1$ if s_n is an n multiple root. By Theorem 6.5 and Theorem 6.6 of Bellman–Cooke,[16] the series expansion in (36) is convergent for $r \geq 2$ and uniformly convergent over any finite interval for $r \geq 2$. Furthermore, from (33) we can obtain $\lim_{r\to\infty} Y_1(r) = 0$.

Proposition 4: The solution of (33) can be expressed as a convergent generalized Fourier-series type expansion in (36) for $r \geq 2$ which is uniformly convergent over any finite interval for $r \geq 2$. Moreover, $Y_1(r) \to 0$ as $r \to \infty$.

This completes our discussions of the first order perturbative solution. Same conclusions as described in Propositions 1–4 can be obtained for each order of the perturbation series.

IV. PHYSICAL SIGNIFICANCE OF THE CONDITION $\mu \int_0^1 |K(t)|\, dt < 1$

Notice $Q_0(r) = K(r) = \frac{1}{2}(1-\eta)^{-2}[(1+2\eta)r^2 - 3\eta r - (1-\eta)]$ for $0 \leq r \leq 1$ and $Q_0(r) = K(r) = 0$ for $r \geq 1$.[6] It can readily be seen that $K(r)$ has two roots $r_1 = (\eta - 1)/(\eta + 1)$, and $r_2 = 1$. But $K(r) \leq 0$ for $r \geq 0$ when $0 < \eta < 1$. Consequently the conditions $\mu\int_0^1 |K(t)|\, dt < 1$ implies $0 < \eta < (3 - \sqrt{7})/2$. On the other hand, when $\eta > 1$, we have $K(r) \geq 0$ for $0 < r < (\eta - 1)/(1 + 2\eta)$, and $K(r) \leq 0$ for $r \geq (\eta - 1)/(1 + 2\eta)$. Hence the condition $\mu \int_0^1 |K(t)|\, dt < 1$ implies $1 < \eta < 2.66$. So long as η is in the regions $(0, (3-\sqrt{7})/2)$ and $(1, 2.66)$, the solutions we have obtained for $Q_n(r)$ and $Y_n(r)$ are unique. $Q_n(r)$ is continuous and bounded on $[0, \infty)$, whereas $Y_n(r)$ is of class C^2 on $[1, \infty)$. (So far we have not been able to prove that Y_n is continuous at $r = 1$.) When $(3 - \sqrt{7})/2 < \eta < 1$ and $\eta > 2.66$, the operator T defined by (20) is no longer contractive and the series solution of $Q_1'(r) = \alpha(r) + \sum_{n=1}^\infty \mu^n \psi_n(r)$ diverges. The solution corresponding to the region $1 < \eta < 2.66$ should be considered as unphysical solutions because the density η is unreasonable high. However, $\eta = (3 - \sqrt{7})/2 \sim 0.175$ corresponds to $\rho = 0.33$, which is close to the critical density $\rho_c = 0.27$ of the Lennard–Jones fluids reported by Watts.

V. CONVERGENCE OF PERTURBATION SERIES AND PHASE TRANSITIONS

Since $Y_n(r)$ is of class C^2 on $[1, \infty)$ and that $Y_n(r) \to 0$ as $r \to \infty$ for $n \geq 1$, it follows that $Y_n(r)$ is uniformly continuous and bounded on $[1, \infty)$. So is $Q_n(r)$ on $[0, \infty)$ for $n \geq 1$. It is possible that $Y_n(r)$ is discontinuous at $r = 1$, however it can only be a finite discontinuity for $Y_n(r)$ is continuous in $[0, 1)$. For the discussion of convergence of pertrubation series, it suffices to consider $Q_n(r)$ and $Y_n(r)$ for $r \geq 1$.

By virture of Proposition 2 and Proposition 4, each $Q_n(r)$ is uniformly continuous, bounded, and square integrable on $[0, \infty)$ and can be uniformly approximated by Laguerre functions $e^{-\alpha r}p(r)$, where $\alpha > 0$ and $p(r)$ is a polynomial, whereas $Y_n(r) = \sum_{m=1}^\infty p_{m-1} \exp(s_m r)$, where $p_{m-1}(r)$ is a polynomial at most of degree $m - 1$ if S_m is an m multiple root, and $\{S_m\}$ is a sequence of roots in the left-hand plane (except for the triple root at the origin) arranged in nondecreasing order of absolute value. Morever, $Y_n(r)$ is uniformly continuous and bounded on $[1, \infty)$.

Although each $Q_n(r)$ and $Y_n(r)$ are bounded, $\sup|Q_n(r)|$ and $\sup|Y_n(r)|$ depend on n and may increase as n increases. So far, we have not been able to obtain the asymptotic behavior of $\sup|Q_n(r)|$ and $\sup|Y_n(r)|$ as $n \to \infty$. However, in view of the fact that

$$\left|\sum_{n=1}^\infty \frac{(\beta\xi)^n}{n!} Q_n(r)\right| \leq \sum_{n=1}^\infty \frac{(\beta\xi)^n}{n!} \sup|Q_n(r)|,$$

the series $\sum_{n=1}^\infty (\beta\xi)^n/n!\, Q_n(r)$ is absolutely and uniformly convergent for $0 < \beta < 1$ if and only if $\sup|Q_n(r)| \leq n!$ The same conclusion can be made for the series $\sum_{n=1}^\infty [(\beta\xi)^n/n!]Y_n(r)$.

Proposition 5: Suppose $\mu\int_0^1 |K(t)|\, dt < 1$. The perturbation series $\sum_{n=1}^\infty [(\beta\xi)^n/n!]Q_n(r)$ and $\sum_{n=1}^\infty [(\beta\xi)^n/n!]Y_n(r)$ are absolutely and uniformly convergent if and only if $\sup|Q_n(r)| \leq n!$ and $\sup|Y_n(r)| \leq n!$

Baxter[6] has shown that the inverse compressibility can be written as

$$\beta\left(\frac{\partial p}{\partial \rho}\right)_T = \tilde{Q}^2(0) = \left[1 - 12\eta \int_0^\infty Q(r)\, dr\right]^2.$$

By the absolute and uniform convergence of the perturbation series we have

$$\tilde{Q}(0)=1-12\eta\int_0^1 Q_0(r)dr-12\eta\sum_{n=1}^{\infty}\frac{(\beta\xi)^n}{n!}\int_0^{\infty}Q_n(r)dr. \tag{37}$$

For hard sphere potential it can easily be seen that $\tilde{Q}(0)=1-12\eta\int_0^1 Q_0(r)dr=(1+2\eta)/(1-\eta)^2$. But the phase transition is completely determined by the condition $(\partial p/\partial\rho)_T=\tilde{Q}(0)=0$. Consequently there exists no gas-liquid phase transition for hard sphere potential. For the intermolecular potential considered in this paper the possibility of a phase transition can not be completely ruled out because (37) can be approximated by a polynomial $P(\beta\xi)$ due to the convergence of the series [the convergence should be rapid since $Q(r)$ is convergent under the supremum norm.]. The real roots of $P(\beta\xi)$ in the interval $(0,1)$[17] give rise to phase transitions.

ACKNOWLEDGMENT

The author wishes to thank the referee for his valuable suggestions which have helped improve the presentation of the paper.

[1] L. S. Ornstein and F. Zernike, Proc. Acad. Sci. Amsterdam **17**, 793 (1914).
[2] T. L. Hill, *Statistical Mechanics* (McGraw—Hill, New York, 1956).
[3] J. K. Percus and G. J. Yevick, Phys. Rev. **110**, 1 (1958).
[4] M. S. Wertheim, Phys. Rev. Lett. **10**, 321 (1963); J. Math. Phys. **5**, 643 (1963).
[5] E. Thiel, J. Chem. Phys. **39**, 474 (1963).
[6] R. J. Baxter, Phys. Rev. **154**, 170 (1967); Aust. J. Phys. **21**, 563 (1968); J. Chem. Phys. **49**, 2770 (1968).
[7] M. Chen, "Differential-Difference Equations for the Radial Distribution Functions of Hard Rods and Hard Spheres Under Percus—Yevick Approximation," J. Math. Anal. Appl. **64**, 629 (1978).
[8] J. Groeneveld, J. Chem. Phys. **53**, 3193 (1970).
[9] R. O. Watts, J. Chem. Phys. **49**, 50 (1968); **50**, 984 (1968).
[10] R. W. Zwanzig, J. Chem. Phys. **22**, 1420 (1954); J. A. Barker and D. Henderson, J. Chem. Phys. **47**, 2856, 4714 (1967).
[11] Suppose $c(r)=0$ for $r>R$, R a real number. Under the assumption that $\tilde{h}(w)$ is bounded for real w, Baxter has derived the following relations:

$$rc(r)=-Q'(r)+12\eta\int_r^R Q'(t)Q(t-r)dt,\quad 0\le r\le R,$$

$$rh(r)=-Q'(r)+12\eta\int_0^R (r-t)h(r-t)Q(t)\,dt,\quad r\ge 0,$$

where $Q(r)$ is bounded for $0\le r\le R$ and $Q(r)=0$ for $r<0$, $r\ge R$. Recently Chen has proved that the boundedness of $\tilde{h}(w)$ is a necessary and sufficient condition for Baxter's relations [M. Chen, J. Math. Phys. **16**, 1150 (1975)]. The Baxter's relations (2) and (3) considered in this paper can be obtained by assuming R as a parameter and finally taking the limit $R\to\infty$. Alternatively following Baxter's method, Eqs. (2) and (3) can be proved under the assumption that $|S(r)|\to ae^{-6r}$ as $r\to\infty$ and $\tilde{h}(w)$ is bounded for real w.
[12] Alternatively we can introduce a cutoff into $v(r)$ by assuming $v(r)=0$ for $r\ge R$ and finally taking the limit $R\to\infty$.
[13] S. T. Hu, *Introduction to General Topology* (Holden—Day San Francisco, 1966), p. 148.
[14] R. Bellman and K. L. Cooke, *Differential-Difference Equation* (Academic, New York, 1963), Theorem 6.2, p. 171.
[15] See Ref. 14, p. 59.
[16] See Ref. 14, p. 186.
[17] It is possible that the interval of convergence for $\beta\xi$ in $Q(r)$ greater than (0,1) if $\sup|Q_n(r)|\le\alpha^n$ (α positive constant).

Classification of gauge fields[a]

J. Anandan[b]

Department of Physics and Astronomy, University of Pittsburgh, Pittsburgh, Pennsylvania 15260
(Received 3 March 1978; revised manuscript received 2 June 1978)

It is shown that the Lorentz invariants of an arbitrary gauge field are double valued functions of a Lorentz invariant matrix L^{ij} when rank $L=3$ and single valued functions of L^{ij} when rank $L\neq 3$. The question of how many Lorentz inequivalent realizations of the Lorentz invariants there are is answered. This leads naturally to a classification of an arbitrary gauge field at a space–time point, given previously by Anandan and Tod based on the rank of L. The answer to the analogous question for the gauge invariants of the SU(2) gauge field leads to a new classification of this field. Five more classifications of this field, including one which is symmetric with respect to space–time and isospin groups, are also presented.

I. INTRODUCTION

The classification of gauge fields has recently attracted considerable interest.[1–6] In this paper we shall study classifications that arise from a systematic study of Lorentz invariants and gauge invariants at a space–time point for a guage field $F^i_{\mu\nu}$ ($i=1,2,\dots,N,\mu,\nu=0,1,2,3$, the skew symmetric indices μ,ν transform under the Lorentz group, and i transforms under the adjoint representation of the gauge group). We study the Lorentz invariants of an arbitrary gauge field and the gauge invariants of the SU(2) gauge field, addressing specifically the following questions: (1) How many functionally independent Lorentz (gauge) invariants are there? (2) what is a polynomial basis for the invariants which are polynomials in $F^i_{\mu\nu}$?, and (3) how many Lorentz (gauge) inequivalent realizations of the invariants are there? (By a realization of a given set of invariants, is meant a field $F^i_{\mu\nu}$, such that the values of its corresponding invariants are the same as the given set of invariants.)

We answer the above questions for the Lorentz invariants of an arbitrary gauge field and the gauge invariants of the SU(2) gauge field. The answers for Lorentz invariants lead naturally to a classification of an *arbitrary gauge field* based on the rank of a Lorentz invariant matrix L^{ij}. This classification was first obtained by Anandan and Tod[4] using different considerations. Essentially the same classification was subsequently obtained for the special case of the $SU(2)$ gauge field by Wang and Yang[6] by considering the Lorentz-gauge inequivalent realizations of the matrix L for this field. In the present work, which was done independently of the work of Wang and Yang, we study the Lorentz inequivalent realizations of *all* Lorentz invariants for an *arbitrary* gauge field. In particular we study a set of cubic Lorentz invariants T^{ijk} which are not considered by Wang and Yang. In addition, we also study the gauge invariants for the SU(2) gauge field. Our answers to the above questions for gauge invariants lead to a new classification of this field in terms of the number of linearly independent $F^i_{\mu\nu}$, and two more classifications from an eigenvector problem of a matrix N which is defined to be $F^i_{\mu\nu}F^i_{\rho\sigma}$. We also present three new classifications of the anti-self-dual Yang-Mills field $F^i_{\mu\nu}+iF^{*i}_{\mu\nu}$ by considering a gauge invariant matrix M, including a classification which is symmetric with respect to space–time and isospin groups.

The proofs of most of the results are more easily given using the spinor formalism,[7] which is reviewed in Sec. II. But the final results are also stated in the usual tensor notation so that they can be understood without any knowledge of spinor formalism. The question of Lorentz invariants is considered in Sec. III. We first show that all Lorentz invariants of an arbitrary gauge field are functions of two gauge tensors L^{ij} and T^{ijk}. An explicit reduction algorithm is presented which would enable one to write any Lorentz invariant, which is a polynomial in $F^i_{\mu\nu}$, as a polynomial combination of L^{ij} and T^{ijk}.

These two tensors are, however, not independent, and it is shown that T^{ijk} is determined up to a sign by L^{ij}. Also since the complex symmetric $N\times N$ matrix L, whose rank does not exceed 3, always has realizations, it follows that the number of functionally independent Lorentz invariants of $F^i_{\mu\nu}$ ($i=1,\dots,N$) is the same as the number of independent real parameters which determine L, i.e., $6N-6$ for $N>1$ and 2 when $N=1$. The number of Lorentz inequivalent realizations are then shown to depend on the rank of L and the number of linearly independent anti-self-dual fields $F^i_{\mu\nu}+iF^{*i}_{\mu\nu}$, which leads to a classification of gauge fields.

The question of Lorentz-gauge inequivalent realizations of the Lorentz invariants is also analyzed for the SU(2) gauge field. This provides a refinement of the above mentioned classfication of an arbitrary gauge field, for this special case. If our results here appear to differ from those of Wang and Yang[6] this is because we take account of the role of the number of linearly independent anti-self-dual fields $F^i_{\mu\nu}+iF^{*i}_{\mu\nu}$ in studying the realizations of Lorentz invariants, and because the spinor method we use is different from their method, which leads to a slightly different classification. In Sec. IV we introduce a relation called "conjugation," which enables one to obtain from any given classification of the SU(2) gauge field, based on the algebraic properties of the field, another classification. In Sec. V it is shown that the SU(2) gauge field has 15 functionally independent gauge invariants, and a polynomial basis is given. Also, we show that

[a]This work forms part of the author's PhD thesis at the University of Pittsburgh, May, 1978.
[b]Present address: Department of Physics and Astronomy, University of Maryland, College Park, MD 20742.

 0022-2488/79/020260-09$01.00

if two SU(2) gauge fields have the same values for all their gauge invariants, then they must be related by a gauge transformation. This study leads to three new classifications of the SU(2) gauge field. Most of the results are summarized in Tables I, II, III, and Figs. 1 and 2.

II. SPINOR FORMALISM

It is well known that the group SL(2,C) is (2–1) isomorphic to the proper, orthochronous Lorentz group **L**. This enables the association with a tensor that transforms under the Lorentz group, a corresponding spinor which transforms under SL(2,C). Formally this correspondence can be made by using the matrices $\sigma^\mu = (1,\boldsymbol{\sigma})$, where **1** is the identity matrix and $\boldsymbol{\sigma}^i$ are the Pauli spin matrices. Then the vector v_μ is represented by the spinor $v_\mu \sigma^\mu_{AA'}$. It can be shown that[7]

$$F^i_{\mu\nu}\sigma^\mu_{AA'}\sigma^\nu_{BB'} = \phi^i_{AB}\epsilon_{A'B'} + \bar{\phi}^i_{A'B'}\epsilon_{AB}, \tag{2.1}$$

and if $F^{*i}_{\mu\nu} = \frac{1}{2}\epsilon_{\mu\nu}{}^{\rho\sigma}F^i_{\rho\sigma}$, then

$$(F^i_{\mu\nu} + F^{*i}_{\mu\nu})\sigma^\mu_{AA'}\sigma^\nu_{BB'} = 2\phi^i_{AB}\epsilon_{A'B'}, \tag{2.2}$$

where ϕ^i_{AB} are complex and symmetric in the spinor indices $A, B = 0, 1$, $i = 1,\ldots N$, the bar denotes complex conjugation, and ϵ_{AB} (or $\epsilon_{A'B'}$) is the antisymmetric tensor in two dimensions with $\epsilon_{01} = \epsilon_{0'1'} = 1$. The spinor indices are raised and lowered using the ϵ spinor, which in this respect plays a role analogous to the metric in tensor calculus.

Given an arbitrary symmetric rank-2 spinor ϕ_{AB}, there exist rank-1 spinors α_A and β_A such that

$$\phi_{AB} = \alpha_{(A}\beta_{B)}, \tag{2.3}$$

where $\alpha_{(A}\beta_{B)} = \frac{1}{2}(\alpha_A\beta_B + \alpha_B\beta_A)$. α_A and β_A are linearly independent if and only if $\phi_{AB}\phi^{AB} \equiv \lambda \neq 0$. So when $\lambda = 0$, $\phi_{AB} = \alpha_A\alpha_B$ for same spinor α_A. Given two linearly independent spinors α_A and β_B, a general symmetric spinor ϕ_{AB} can be written as

$$\phi_{AB} = a\alpha_A\alpha_B + b\beta_A\beta_B + c\alpha_{(A}\beta_{B)}, \tag{2.4}$$

where a, b, and c are complex numbers.

III. LORENTZ INVARIANTS

We shall define a Lorentz invariant of an arbitrary gauge field $F^i_{\mu\nu}$ ($i = 1,\ldots N$) to be an algebraic function of $F^i_{\mu\nu}$ which is invariant under the Lorentz group for all values of $F^i_{\mu\nu}$.[8] Since an infinite number of Lorentz invariants can be formed from any given $F^i_{\mu\nu}$, the study of these invariants and their realizations may at first sight appear to be formidable. The following theorem which is proven in Appendix A, however, leads to a remarkable simplification:

Theorem 1: Two gauge fields have the same values for all their Lorentz invariants if and only if they have the same values for L^{ij} and T^{ijk} defined by

$$L^{ij} = 2\phi^{iB}_{A}\phi^{jA}_{B} = K^{ij} + iJ^{ij} \tag{3.1}$$

and

$$T^{ijk} = 4\phi^{iB}_{A}\phi^{jC}_{B}\phi^{kA}_{C} = t^{ijk} + it'^{ijk}, \tag{3.2}$$

where

$$K^{ij} = \tfrac{1}{2}F^i_{\mu\nu}F^{j\nu\mu}, \quad J^{ij} = \tfrac{1}{2}F^{*i}_{\mu\nu}F^{j\nu\mu}, \tag{3.3}$$

and

$$t^{ijk} = F^{i\,\nu}_{\mu}F^{j\rho}_{\nu}F^{k\mu}_{\rho}, \quad t'^{ijk} = -F^{*i\,\nu}_{\mu}F^{*j\rho}_{\nu}F^{*k\mu}_{\rho}. \tag{3.4}$$

Moreover, any Lorentz invariant which is a polynomial in $F^i_{\mu\nu}$ can be written as a polynomial in K^{ij}, J^{ij}, t^{ijk}, and t'^{ijk}.

L^{ij} is a symmetric matrix, whereas T^{ijk} is completely antisymmetric. Hence, when $N = 3$, which is the case for the SU(2) gauge field,

$$T^{ijk} = \tau\epsilon^{ijk}, \tag{3.5}$$

where

$$\tau = \tfrac{2}{3}\epsilon_{ijk}\phi^{iB}_{A}\phi^{jC}_{B}\phi^{kA}_{C} = t + it', \tag{3.6}$$

and ϵ^{ijk}, ϵ_{ijk} are completely antisymmetric with $\epsilon^{123} = \epsilon_{123} = 1$. t and t' are defined by[2]

$$t = \tfrac{1}{6}\epsilon_{ijk}F^{i\nu}_{\mu}F^{j\rho}_{\nu}F^{k\mu}_{\rho},$$
$$t' = -\tfrac{1}{6}\epsilon_{ijk}F^{*i\nu}_{\mu}F^{*j\rho}_{\nu}F^{*k\mu}_{\rho}. \tag{3.7}$$

It follows that for $N = 3$, T^{ijk} in Theorem 1, can be replaced by τ. Also in this case it has been shown that[4]

$$\tau^2 = -\det L. \tag{3.8}$$

The generalization of (3.8) for an arbitrary gauge field is

$$T^{ijk}T^{lmn} = -6L^{[i|l|}L^{j|m|}L^{k]n}, \tag{3.9}$$

where [ ] represents antisymmetrization over the indices i, j, k (| | means that the index enclosed is left out in the antisymmetrization). (3.9) is proved in Appendix B.

To find the inequivalent realizations of the Lorentz invariants, our strategy will be to first find the inequivalent realizations of L. Theorem 1 then ensures that two realizations of L will have the same values for all the Lorentz invariants if and only if they have the same values for T^{ijk}. The number of Lorentz inequivalent realizations of L for an arbitrary gauge field is given by the following theorem:

Theorem 2: Let L^{ij} be an $N\times N$ complex symmetric matrix. Then there exists a gauge field satisfying (3.1) if and only if rank $L \leqslant 3$. The number of Lorentz inequivalent realizations, when they exist are given by Table I.

Proof: Suppose there exists ϕ^i_{AB} satisfying (3.1). Define the column vectors

$$\phi^i_{\alpha} = \begin{pmatrix} \phi^i_{00} \\ \phi^i_{11} \\ \sqrt{2}\phi^i_{01} \end{pmatrix}, \quad \phi^{j\alpha} = \begin{pmatrix} \phi^i_{00} \\ \phi^i_{11} \\ \sqrt{2}\phi^i_{01} \end{pmatrix}. \tag{3.10}$$

Since ϕ^i_{AB} is symmetric in A, B, (3.1) implies

$$L^{ij} = -2\phi^i_{\alpha}\phi^{j\alpha}. \tag{3.11}$$

Since ϕ^i_{α}, $\phi^{j\alpha}$ are 3×1 column vectors, it follows that rank $L \leqslant 3$.[4]

Conversely suppose that rank $L \leqslant 3$. We shall consider separately the cases rank $L = 0$, 1, 2, and 3.

TABLE I. Classification of gauge fields based on the number of Lorentz inequivalent realizations of the matrix L. The number in parenthesis gives the number of inequivalent realizations of *all Lorentz invariants* when this number differs from the number of inequivalent realizations of L. For the special case of the Maxwell field ($N=1$) one obtains the usual classification into radiative and nonradiative types. The labels I, II, etc., which denote the various types, were used in Ref. 4.

Rank L	No. of linearly independent ϕ^j_{AB} or $F^j_{\mu\nu}+iF^{j*}_{\mu\nu}$, $j=1,\cdots N$	Type of gauge field	No. of Lorentz inequivalent realizations of L for $N=1$	$N=2$	$N\geqslant 3$
3	3	I	–	–	2 (1)
2	2	II	–	1	1
1	2	A	–	2	$2\infty^{2N-4}$
	1	D	1	1	1
0	1	N	1	∞^2	∞^{2N-2}
	0	O	1	1	1

Rank $L=0$

Then $L=0$. This implies that either $\phi^i_{AB}=0$, $i=1,\ldots,N$, or without loss of generality,

$$\phi^1_{AB}=\alpha_A\alpha_B, \tag{3.12}$$

for some spinor $\alpha_A\neq 0$. Given any other $\phi^{1}_{AB}=\alpha'_A\alpha'_B$, $\alpha'_A\neq 0$, we can transform it to ϕ^1_{AB} by an SL(2,C) transformation that takes α'_A into α_A. Now write ϕ^i_{AB}, $i\geqslant 2$ in the general form (2.4) and using $L^{1i}=0=L^{ii}$ (no summation over i) we have

$$\phi^i_{AB}=\chi^i\alpha_A\alpha_B, \tag{3.13}$$

where $\chi^1=1$ and χ^i, $i\geqslant 2$, are arbitrary complex parameters. Clearly, given two fields ϕ^i_{AB} in the form (3.13), corresponding to two sets of χ^i (with $\chi^1=1$), it is not possible to Lorentz transform one to the other. Hence, the Lorentz inequivalent fields in this case depend on N-1 complex or 2N-2 real parameters.

Also it is clear that the two subcases above correspond to zero or one linearly independent ϕ^i_{AB} (with respect to complex coefficients).

Rank $L=1$

In this case there exists a nonsingular matrix P such that

$$\tilde{L}\equiv PLP^{T}=\begin{pmatrix} p & 0\ldots 0\\ 0 & \\ \cdot & \\ \cdot & 0\\ \cdot & \\ 0 & \end{pmatrix},\quad p\neq 0. \tag{3.14}$$

For each ϕ^i_{AB} satisfying (3.1), there exists a corresponding $\tilde{\phi}^i_{AB}$ satisfying

$$\tilde{L}^{ij}=2\tilde{\phi}^{iB}_{A}\tilde{\phi}^{jA}_{B}, \tag{3.15}$$

and conversely, where ϕ and $\tilde{\phi}$ are related by the nonsingular transformation

$$\tilde{\phi}^i_{AB}=P^{ij}\phi^j_{AB}. \tag{3.16}$$

Also note that *rank L and the number of linearly independent ϕ^i_{AB} are invariant under the above nonsingular transformation*. Now since $L^{11}=p$,

$$\tilde{\phi}^1_{AB}=p^{1/2}\alpha_{(A}\beta_{B)},\quad \alpha_A\beta^A=1. \tag{3.17}$$

for some spinors α_A, β_B. (p has two square roots. Pick one and denote it by $p^{1/2}$.) Given any other $\tilde{\phi}^{1'}_{AB}=p^{1/2}\alpha'_{(A}\beta'_{B)}$ with $\alpha'_A\beta'^A=1=\alpha_A\beta^A$, there exists an SL(2,$C$) transformation

$$\alpha'_A\to\alpha_A,\quad \beta'_A\to\beta_A,$$

which transforms $\tilde{\phi}^{1'}_{AB}\to\tilde{\phi}^1_{AB}$.

Now write $\tilde{\phi}^i_{AB}$, $i\geqslant 2$ in the general form (2.4) and using $\tilde{L}^{1i}=0$, $\tilde{L}^{ii}=0$, $i=2,\ldots,N$ (no summation over i) one obtains,

$$\tilde{\phi}^i_{AB}=\chi^i\alpha_A\alpha_B,\quad i\geqslant 2,\ \text{or}\ \tilde{\phi}^i_{AB}=\xi^i\beta_A\beta_B,\quad i\geqslant 2, \tag{3.18}$$

where χ^i or ξ^i ($i\geqslant 2$) are arbitrary complex numbers. Hence, in this case there can be one or two linearly independent $\tilde{\phi}^i_{AB}$. If only one $\tilde{\phi}^i_{AB}$ is linearly independent, then $\chi^i=0$, $\xi^i=0$ ($i\geqslant 2$) so that $\tilde{\phi}^i_{AB}$ is unique up to Lorentz transformations.

If there are two linearly independent $\tilde{\phi}^i_{AB}$, then $\chi^i\neq 0$ or $\xi^i\neq 0$ for some $i=r\geqslant 2$. Without loss of generality, $r=2$. Now the SL(2,C) transformation $\alpha_A\to(\chi^2)^{1/2}\alpha_A$, $\beta_A\to(\chi^2)^{-1/2}\beta_A$, leaves $\tilde{\phi}^1_{AB}$ invariant, but transforms $\chi^2\alpha_A\alpha_B$ to $\alpha_A\alpha_B$. Similarly $\xi^2\beta_A\beta_B$ can be transformed to $\beta_A\beta_B$ by a suitable SL(2,C) transformation. Also, it is not possible to transform $\alpha_A\alpha_B$ into $\beta_A\beta_B$ by an SL(2,C) transformation that leaves $\tilde{\phi}^1_{AB}$ invariant. Hence, when $N=2$, there are two Lorentz inequivalent fields. When $N\geqslant 3$, however, there are two sets of N-2 complex parameters χ^i or ξ^i ($i\geqslant 3$), i.e., 2N-4 real parameters, which describe the Lorentz inequivalent realizations.

Rank $L=2$

There exists a nonsingular matrix P such that

$$\tilde{L}\equiv PLP^{T}=\begin{pmatrix} p & r & \\ r & q & 0\\ & 0 & 0\end{pmatrix},\quad pq-r^2\neq 0. \tag{3.19}$$

Also, without loss of generality, $p\neq 0$. Then since $\tilde{L}^{11}=p$,

$$\tilde{\phi}^{1}_{AB}=p^{1/2}\alpha_{(A}\beta_{B)},\quad \alpha_A\beta^A=1 \tag{3.20}$$

for some spinors α_A, β_B. Write $\tilde{\phi}^{2}_{AB}$ in the general form (2.4). Then

$$\tilde{L}^{12}=r\Rightarrow cp^{1/2}=r, \tag{3.21}$$

$$\tilde{L}^{22}=q\Rightarrow -4ab+c^2=q, \tag{3.22}$$

which gives

$$4ab=\frac{r^2-pq}{p}=\mu \quad \text{(say)}. \tag{3.23}$$

Since $\mu\neq 0$ (rank $\tilde{L}$ = rank $L=2$), $a\neq 0$ and $b\neq 0$. The SL(2,C) transformation

$$\alpha_A\to\frac{\mu^{1/4}}{(2a)^{1/2}}\alpha_A,\quad \beta_A\to\frac{(2a)^{1/2}}{\mu^{1/4}}\beta_A$$

does not change (3.20) but changes $\tilde{\phi}^{2}_{AB}$, on using (3.23), to

$$\tilde{\phi}^{2}_{AB}=\frac{\mu^{1/2}}{2}(\alpha_A\alpha_B+\beta_A\beta_B)+rp^{-1/2}\alpha_{(A}\beta_{B)}. \tag{3.24}$$

Also $\tilde{L}^{1i}=0=\tilde{L}^{2i}=\tilde{L}^{ii}$, $i\geqslant 3$ imply

$$\tilde{\phi}^{i}_{AB}=0,\quad i\geqslant 2. \tag{3.25}$$

Since $\tilde{\phi}^{i}_{AB}$ have no arbitrary parameters, the field is unique up to Lorentz transformations. Also in this case there are exactly two linearly independent ϕ^{i}_{AB}.

Rank $L=3$

There exists a nonsingular matrix P such that

$$\tilde{L}\equiv PLP^{T}=\begin{pmatrix}\tilde{L}_3 & 0\\ 0 & 0\end{pmatrix}, \tag{3.26}$$

where the 3×3 matrix $\tilde{L}_3$ has one of the following canonical forms[9]:

$$\text{(a)}\begin{pmatrix}\lambda_1 & & 0\\ & \lambda_2 & \\ 0 & & \lambda_3\end{pmatrix},\quad \text{(b)}\begin{pmatrix}\lambda_1 & 0 & 0\\ 0 & \lambda_2+i & 1\\ 0 & 1 & \lambda_2-i\end{pmatrix},$$

or

$$\text{(c)}\begin{pmatrix}\lambda_1 & 1+i & 0\\ 1+i & \lambda_1 & 1-i\\ 0 & 1-i & \lambda_1\end{pmatrix}.$$

The above result is valid whenever rank $L\leqslant 3$. In the present case since rank $L=3$, we have $\lambda_i\neq 0$ for $i=1, 2, 3$.

Using now the same general techniques used for rank $L=0$, 1 and 2, we can write the general $\tilde{\phi}^{i}_{AB}$ up to Lorentz transformations, corresponding to case (a), in the following canonical way:

Case (a):

$$\tilde{\phi}^{1}_{AB}=\lambda_1^{1/2}\alpha_{(A}\beta_{B)},\quad \alpha_A\beta^A=1, \tag{3.27}$$

$$\tilde{\phi}^{2}_{AB}=i\frac{\lambda_2^{1/2}}{2}(\alpha_A\alpha_B+\beta_A\beta_B), \tag{3.28}$$

$$\tilde{\phi}^{3}_{AB}=\pm\frac{\lambda_3^{1/2}}{2}(\alpha_A\alpha_B-\beta_A\beta_B), \tag{3.29}$$

$$\tilde{\phi}^{i}_{AB}=0\quad \text{for } i>3. \tag{3.30}$$

Hence, there are exactly two Lorentz inequivalent fields corresponding to the $\pm$ sign in (3.29) in this case. These two fields can be related by multiplication of the whole field by -1 and the SL(2,C) transformation,

$$\alpha_A\to i\beta_A,\quad \beta_A\to i\alpha_A.$$

So in this case, L is realized by two Lorentz inequivalent fields related by multiplication by -1, which expresses the result in a gauge covariant manner.

For cases (b) and (c) we find that, similar to case (a), there are two Lorentz inequivalent fields $\tilde{\phi}^{i}_{AB}$ and $-\tilde{\phi}^{i}_{AB}$, where $\tilde{\phi}^{i}_{AB}$ can be written in the following canonical way:

Case (b):

$$\tilde{\phi}^{1}_{AB}=\lambda_1^{1/2}\alpha_{(A}\beta_{B)}, \tag{3.31}$$

$$\tilde{\phi}^{2}_{AB}=\frac{i}{2}(\lambda_2+i)^{1/2}(\alpha_A\alpha_B+\beta_A\beta_B), \tag{3.32}$$

$$\tilde{\phi}^{3}_{AB}=\tfrac{1}{2}(\lambda_2+i)^{1/2}\alpha_A\alpha_B+\frac{-\lambda_2+i}{2(\lambda_2+i)^{1/2}}\beta_A\beta_B, \tag{3.33}$$

$$\tilde{\phi}^{i}_{AB}=0,\quad i>3, \tag{3.34}$$

provided $\lambda_2+i\neq 0$. If $\lambda_2+i=0$, replace (3.32) and (3.33) by

$$\tilde{\phi}^{2}_{AB}=\alpha_A\alpha_B \tag{3.32$'$}$$

and

$$\tilde{\phi}^{3}_{AB}=-i\alpha_A\alpha_B-\tfrac{1}{2}\beta_A\beta_B. \tag{3.33$'$}$$

Case (c):

$$\tilde{\phi}^{1}_{AB}=\lambda_1^{1/2}\alpha_{(A}\beta_{B)},\quad \alpha_A\beta^A=1, \tag{3.35}$$

$$\tilde{\phi}^{2}_{AB}=\frac{1+i+\lambda_1}{2\lambda_1^{1/2}}\alpha_A\alpha_B+\frac{1+i-\lambda_1}{2\lambda_1^{1/2}}\beta_A\beta_B+\frac{1+i}{\lambda_1^{1/2}}\alpha_{(A}\beta_{B)}, \tag{3.36}$$

$$\tilde{\phi}^{3}_{AB}=\frac{i\lambda_1^{1/2}}{2}(\alpha_A\alpha_B+\beta_A\beta_B), \tag{3.37}$$

$$\tilde{\phi}^{i}_{AB}=0\quad \text{for } i>3. \tag{3.38}$$

This completes the proof of Theorem 2. We have now obtained the Lorentz inequivalent fields corresponding to any given L (of rank not exceeding 3). Now two fields having the same values for L will also have the same values for all their Lorentz invariants, except possibly when rank $L=3$. This follows from Theorem 1 and the following theorem:

Theorem 3: The tensor $T^{ijk}=0$ if and only if rank $L\neq 3$. When rank $L=3$, T^{ijk} is determined up to a sign by L^{ij}.

Proof: Suppose $T^{ijk}\neq 0$ for $(i,j,k)=(1,2,3)$ (say). Now from (3.9), $(T^{123})^2=-6L^{[1|1|}L^{2|2|}L^{3]3}=-\det L_3$, where L_3 is the 3×3 matrix formed from L^{ij}, $i,j,=1, 2, 3$. Therefore, $\det L_3\neq 0$, which implies that rank $L\geqslant 3$. But from Theorem 2 rank $L\leqslant 3$. Hence, rank $L=3$. Conversely, if L has rank 3, then, since L is symmetric, it must have a nonzero 3×3 principal minor N. It follows then from (3.9) that $(T^{i_1j_1k_1})^2=-N\neq 0$, where the indices i_1, j_1, k_1, label the rows (or columns) of the principal minor N.

The fact that T^{ijk} is determined up to a sign by L^{ij} follows immediately from (3.9).

The two Lorentz inequivalent realizations of L, ϕ^i_{AB} and $-\phi^i_{AB}$, which we found for rank $L=3$ (in the proof of Theorem 2), have different values for the Lorentz invariant T^{ijk}. This follows from (3.2) and the fact that $T^{ijk}\neq 0$ for some i, j, k when rank $L=3$, according to Theorem 3. (Indeed this provides an alternative proof of Lorentz inequivalence of ϕ^i_{AB} and $-\phi^i_{AB}$ when rank $L=3$: There cannot be a Lorentz transformation between them since they give different values for the same Lorentz invariant T^{ijk}.) So it follows from Theorems 1, 2, and 3 that the realization of a complete set of Lorentz invariants is unique up to Lorentz transformations when rank $L=3$. When rank $L\neq 3$, the number of inequivalent realizations of a complete set of Lorentz invariants is the same as the number of inequivalent realizations of L. The results are summarized in Table I.

It also follows from Theorems 1 and 3 that L^{ij} determine all the Lorentz invariants uniquely when rank $L\neq 3$. When rank $L=3$, however, corresponding to the given L^{ij} there will be two sets of Lorentz invariants depending on the sign of T^{ijk}. Hence, the Lorentz invariants may be regarded as double valued functions of L^{ij} when rank $L=3$ and single valued functions of L^{ij} when rank $L\neq 3$.

So far our analysis has been for an arbitrary gauge field. When the gauge group is specified, however, one can also ask the following different question: How many Lorentz-gauge inequivalent realizations of the Lorentz invariants are there?[6] This question can be easily answered from the above results, for the SU(2) Yang–Mills field. Consider first the realization of L, and notice that when rank $L=3$, ϕ^i_{AB} and $-\phi^i_{AB}$ are Lorentz-gauge inequivalent, since they have different values for the nonzero Lorentz-gauge invariant τ. When rank $L=2$ or when rank $L=1$ and one ϕ^i_{AB} is linearly independent, we have shown that the realization is unique up to Lorentz transformations and hence obviously also up to Lorentz-gauge transformations. When $L=0$, ϕ^i_{AB} is of the form (3.13) with χ^i being arbitrary complex numbers. Using the freedom of Lorentz transformations, χ^i can be normalized, so that either

$$\chi^k\chi^k=0 \tag{3.39}$$

or, without loss of generality,

$$\chi^k\chi^k=1. \tag{3.40}$$

It is easy to show that the set of χ^k satisfying (3.40) and are inequivalent under the gauge transformations [O(3)], are described by a single real parameter. If (3.39) is satisfied, however, $\phi^k_{AB}=\chi^k\alpha_A\alpha_B$ is unique up to Lorentz-gauge transformations.

Consider now the remaining case, namely rank $L=1$ and two ϕ^i_{AB} are linearly independent. Notice first from (3.17) and (3.18) that ϕ^i_{AB} determined by (3.16) must be of the form

$$\phi^i_{AB}=\alpha_{(A}\beta^i_{B)} \quad \text{or} \quad \alpha^i_{(A}\beta_{B)}. \tag{3.41}$$

So using the freedom of Lorentz transformations, we can write, without loss of generality,

$$\phi^1_{AB}=(L^{11})^{1/2}\alpha_{(A}\beta_{B)}, \quad \phi^2_{AB}=\pm(L^{22})^{1/2}\alpha_{(A}\beta_{B)}+\beta_A\beta_B,$$

$$\phi^3_{AB}=\pm(L^{33})^{1/2}\alpha_{(A}\beta_{B)}+d\beta_A\beta_B,$$

$$\alpha_A\beta^A=\pm 1, \quad L^{11}\neq 0. \tag{3.42}$$

The $\pm$ signs in front of $(L^{22})^{1/2}$ and $(L^{33})^{1/2}$ are determined by L^{12} and L^{13} (when L^{22} or $L^{33}\neq 0$). It is clear now that given another field ϕ'^i_{AB} in the form (3.42), with d' instead of d, then ϕ'^i_{AB} cannot be transformed to ϕ^i_{AB} by a gauge transformation, if $(d')^2\neq d^2$. Hence, in this case the Lorentz-gauge inequivalent fields are described by two real parameters corresponding to the different values of d^2.

The number of Lorentz-gauge inequivalent realizations of all the Lorentz invariants is now obtained immediately using theorems 1 and 3. We have above an additional type[5] $F=\chi^i\alpha_A\alpha_B$ with $\chi^k\chi^k=0$, so that we have a refinement of the original classification of an arbitrary gauge field, for the special case of an SU(2) gauge field. Type F can also be specified by the invariant conditions $L=0$ and $M=\widetilde{M}=0$ where L, M, and $\widetilde{M}$ are defined in (3.1), (4.2), and (4.6). The results are summarized in Table II.

TABLE II. Classification of SU(2) gauge field based on the number of Lorentz-gauge inequivalent realizations of the matrix L. A refinement of the classification in Table I is obtained for this special case, with type F as the additional case. The number in parenthesis refers to the number of Lorentz-gauge inequivalent realizations of all the Lorentz invariants when this number differs from the number of Lorentz-gauge inequivalent realizations of L. This classification can be more generally regarded as a classification of any gauge field with a compact gauge group.

Rank L	No. of Linearly independent ϕ^j_{AB} or $F^j_{\mu\nu}+iF^{*j}_{\mu\nu}$, $j=1,\cdots N$	Type of Gauge Field	Spinor Form	No. of Lorentz-gauge inequivalent realizations of L when G = SU(2)
3	3	I	–	2 (1)
2	2	II	–	1
1	2	A	$\alpha^k{}_{(A}\beta_{B)}$	$2\infty^2$
	1	D	$\chi^k\alpha_{(A}\beta_{B)}$	1
0	1	N	$\chi^k\alpha_A\alpha_B$	∞^1
	1	F	$\chi^k\alpha_A\alpha_B$, $\chi^k\chi^k=0$	1
	0	O	0	1

IV. CONJUGATION

In this section we shall specialize to the SU(2) gauge field. The index i in ϕ^{i}_{AB} can be replaced in this case by two SU(2) spinor indices in the following way: Let $\sigma^{i}{}_{P}{}^{Q}$, $i=1, 2, 3$, represent the Pauli spin matrices. Define

$$\phi_{AB,P}{}^{Q}=\phi^{i}{}_{AB}\sigma^{i}{}_{P}{}^{Q}. \tag{4.1}$$

We shall reserve the letters $P, Q,\cdots$ taken from the second half of the Latin alphabet for SU(2) indices while the indices $A, B\cdots$ taken from the first half of the Latin alphabet will continue to represent SL(2,C) indices.

Since the Pauli spin matrices $\sigma^{i}{}_{P}{}^{Q}$ are trace free, σ^{iPQ} are symmetric, where the index P has been raised using the alternating symbol ϵ^{PQ} which is invariant under SU(2). Therefore, ϕ^{PQ}_{AB} is symmetric in both pairs of indices (A,B) and (P,Q). Moreover, the algebraic properties for the indices A, B are the same as for the indices P, Q. This is basically due to the fact that the Lie algebra of SL(2,C) is the complexification of the Lie algebra of SU(2). This implies that for every algebraic statement on the indices A, B, there is a corresponding "*conjugate*" statement on the indices P, Q, and conversely.

Given a quantity X defined by a statement about the indices A, B and P, Q, we shall call the corresponding quantity defined by the conjugate statement its conjugate X^c. For instance, the conjugate of the type $A=\alpha_{(A}{}^{PQ}\beta_{B)}$ of the SU(2) gauge field is the type $A^{c}=\chi_{AB}{}^{(P}\delta^{Q)}$. The conjugate of type $N=\gamma^{(P}\delta^{Q)}\alpha_{(A}\alpha_{B)}$ is the type $N^{c}=\gamma^{(P}\gamma^{Q)}\alpha_{(A}\beta_{B)}$. Type $D=\gamma^{(P}\delta^{Q)}\alpha_{(A}\beta_{B)}$ is self-conjugate. It is clear now that for any given classification of the SU(2) gauge field based on the algebraic properties of the field, there is a corresponding *conjugate classification* consisting of the conjugates of the types in the former classification. More generally, if a statement that depends only on the algebraic properties of ϕ^{PQ}_{AB} is valid, then the conjugate statement is also valid.

We introduce now the matrix M, which is conjugate to L, defined by

$$M^{\beta}_{\alpha}\equiv-2\phi^{i}{}_{\alpha}\phi^{i\beta}=2\phi_{\alpha,P}{}^{Q}\phi^{\beta}{}_{Q}{}^{P},\quad \alpha,\beta=1,2,3, \tag{4.2}$$

where ϕ^{i}_{α} was defined in (3.10). Then

$$\mathrm{Tr}M=M^{\alpha}_{\alpha}=-2\phi^{i}{}_{AB}\phi^{iAB}=\mathrm{Tr}L. \tag{4.3}$$

Similarly,

$$\mathrm{Tr}M^{2}=\mathrm{Tr}L^{2},\quad \mathrm{Tr}M^{3}=\mathrm{Tr}L^{3}. \tag{4.4}$$

Therefore, by the Cayley–Hamilton theorem, $\det M=\det L$ and the eigenvalues of L and M are also the same. To express M in terms of more familiar fields, notice from (2.2) that there exists a nonsingular linear transformation that maps ϕ^{i}_{α} into $E^{i}_{l}-iH^{i}_{l}$, $i, l=1, 2, 3$, where the E^{i}_{l} and H^{i}_{l} are the "electric" and "magnetic" vectors of the gauge field $F^{i}_{\mu\nu}$. Hence, there exists a nonsingular 3×3 matrix P such that

$$PMP^{-1}=\widetilde{M}, \tag{4.5}$$

where $\widetilde{M}$ is defined by

$$\widetilde{M}_{lm}=(E^{j}_{l}-iH^{j}_{l})(E^{j}_{m}-iH^{j}_{m})$$

$$=K^{c}+iJ^{c}, \tag{4.6}$$

where

$$(K^{c})_{lm}=E^{j}_{l}E^{j}_{m}-H^{j}_{l}H^{j}_{m},\quad (J^{c})_{lm}=-E^{j}_{l}H^{j}_{m}-H^{j}_{l}E^{j}_{m}.$$

The Lorentz-gauge invariant defined in (3.6) can be written as

$$\tau=\frac{-2\sqrt{2}}{3}\phi_{A}{}^{B}{}_{P}{}^{Q}\phi_{B}{}^{C}{}_{Q}{}^{R}\phi_{C}{}^{A}{}_{R}{}^{P}, \tag{4.7}$$

where we have made use of

$$\epsilon_{ijk}\leftrightarrow\frac{1}{\sqrt{2}}(\epsilon_{PS}\epsilon_{TQ}\epsilon_{RU}-\epsilon_{PU}\epsilon_{TS}\epsilon_{RQ}), \tag{4.8}$$

the $\leftrightarrow$ denoting that the right-hand side is the spinor form of the left-hand side $[i\leftrightarrow(P,Q), j\leftrightarrow(R,S), k\leftrightarrow(T,U)]$. It is seen from (4.7) that τ is self-conjugate. The following theorem can now be proven using methods analogous to those used in proving theorem 2.

Theorem 4[10]: Let $M_{\alpha\beta}$ be a 3×3 complex symmetric matrix. Then there always exists a gauge field satisfying (4.2). The number of gauge and Lorentz-gauge inequivalent realizations is given in Table III.

TABLE III. Classification of SU(2) gauge field, based on the number of gauge inequivalent and Lorentz-gauge inequivalent realizations of the matrix M. The number of Lorentz-gauge inequivalent realizations is given for the general case when there are no relations on the eigenvalues of M other than what is implied by the rank of M.

Rank M	No. of linearly independent $\phi^{j}{}_{AB}$ or $F^{j}_{\mu\nu}+iF^{j}_{\mu\nu}{}^{*}$, $j=1,\cdots N$	Type of Field	Spinor form	No. of gauge inequivalent realizations of M	No. of Lorentz-gauge inequivalent realization of M
3	3	I	–	$2\infty^{3}$	$2\infty^{3}$
2	2	II^{c}	–	∞^{3}	∞^{3}
1	2	A^{c}	$\alpha_{AB}{}^{(P}\delta^{Q)}$	∞^{5}	∞^{1}
	1	D	$\gamma^{(P}\delta^{Q)}\alpha_{(A}\beta_{B)}$	∞^{1}	∞^{1}
0	1	N^{c}	$\gamma^{P}\gamma^{Q}\alpha_{(A}\beta_{B)}$	∞^{5}	∞^{1}
	1	F	$\gamma^{P}\gamma^{Q}\alpha_{A}\alpha_{B}$	∞^{3}	1
	0	O	0	1	1

For the SU(2) gauge field, using a proof similar to that of (3.9) (see Appendix B), one can also prove

$$V_{\mathbf{ABC}}V_{\mathbf{DEF}} = -6N_{[\mathbf{A}|\mathbf{D}|}N_{\mathbf{B}|\mathbf{E}|}N_{\mathbf{C}]\mathbf{F}}. \tag{5.4}$$

It follows from (5.4) that:

Theorem 7: The tensor $V_{\mathbf{ABC}} = 0$ if and only if rank $N \neq 3$. When rank $N = 3$, $V_{\mathbf{ABC}}$ is determined up to a sign by $N_{\mathbf{AB}}$.

The proof of Theorem 7 is analogous to that of Theorem 3.

Theorems 6 and 7 imply that except when rank $N = 3$, $N_{\mathbf{AB}}$ determines all the gauge invariants. When rank $N = 3$ however, corresponding to the given $N_{\mathbf{AB}}$ there will be two sets of gauge invariants depending on the sign of $V_{\mathbf{ABC}}$. Now since $F^i_{\mu\nu}$ may be regarded as six vectors in a three-dimensional Euclidean space, $N_{\mathbf{AB}}$ can be specified by giving the scalar products of these vectors with three among them. Hence $N_{\mathbf{AB}}$ has 15 and only 15 independent components and therefore exactly 15 functionally independent gauge invariants can be formed from the SU(2) gauge field $F^i_{\mu\nu}$. It also follows from Theorems 5, 6, 7, and the remark below Theorem 6, that *two SU(2) gauge fields have the same values for all their gauge invariants if and only if they are related by a gauge transformation.*

Classifications can also be obtained by considering the eigenvector problem

$$N^{\mathbf{A}}{}_{\mathbf{B}}V^{\mathbf{B}} = \lambda V^{\mathbf{A}}. \tag{5.5}$$

The matrix $N^{\mathbf{A}}{}_{\mathbf{B}}$ (unlike $N_{\mathbf{AB}}$) is not symmetric in general. However, since rank $(N^{\mathbf{A}}{}_{\mathbf{B}}) \leqslant 3$, $N^{\mathbf{A}}{}_{\mathbf{B}}$ can have three, four, five, or six linearly independent eigenvectors and three, four, five, or six zero eigenvalues. This provides more classification schemes.

In conclusion, we note that numerous classifications of gauge fields can be obtained. The classifications obtained in the present paper are intimately related to the study of invariants and may, therefore, be of importance.

ACKNOWLEDGMENTS

It is a pleasure to thank R. Roskies, P. Cvitanovic, G.A.J. Sparling, and K.P. Tod for useful and stimulating discussions.

APPENDIX A

We shall prove Theorem 1 by explcitly constructing a reduction algorithm that enables one to express all Lorentz invariants as well-defined functions of L^{ij} and T^{ijk}.

Consider first a Lorentz invariant which is a polynomial in $F^i_{\mu\nu}$. From the spinor form of $F^i_{\mu\nu}$ given in (2.1), it follows that such an invariant will consist of sums of products of tensors of the form

$$R^{ij\ldots l} = (\phi^i{}_A{}^B \phi^j{}_B{}^C \cdots \phi^l{}_D{}^A) \tag{A1}$$

and

$$\bar{R}^{pq\ldots s} = (\phi^p{}_{A'}{}^{B'} \phi^q{}_{B'}{}^{C'} \cdots \phi^s{}_{D'}{}^{A'}).$$

Clearly $R^{ij} = \frac{1}{2}L^{ij}$ and $R^{ijk} = \frac{1}{4}T^{ijk}$ where L^{ij} and T^{ijk} were defined in (3.1) and (3.2). A tensor $R^{ijk\cdots}$ with more than three indices can be written as the sum of the tensor $R^{[ijk]\cdots}$ and tensors for which at least two of the indices i,j,k are symmetric. The following identities will then enable us to reduce $R^{ijk\cdots}$ to products of L^{ij} and T^{lmn}.

$$\phi^{(i}{}_A{}^C \phi^{j)}_{CB} = \tfrac{1}{4}L^{ij}\epsilon_{AB}, \tag{A2}$$

and

$$\phi^{[i}{}_A{}^C \phi^j{}_C{}^D \phi^{k]}_{DB} = \tfrac{1}{8}T^{ijk}\epsilon_{AB}, \tag{A3}$$

where ( ) and [ ] denote respectively symmetrization and antisymmetrization. (A2) and (A3) can be proven by noticing first that the left-hand sides of (A2) and (A3) are antisymmetric in (A,B) and therefore must be proportional to ϵ_{AB}. The proportionality factor can then be determined by contraction.

Similarly $\bar{R}^{pqr\cdots}$ can be reduced to products of $\bar{L}^{pq}$ and $\bar{T}^{stu}$. So the polynomial invariant is a polynomial combination L, $\bar{L}$, T, and $\bar{T}$. Now a nonpolynomial invariant also has to be formed from polynomial invariants, e.g., as rational functions or square roots, etc. It follows that all Lorentz invariants which are well-defined functions of $F^i_{\mu\nu}$ can be expressed as well-defined functions of L^{ij} and T^{ijk}.

Note: A polynomial basis for the *Lorentz-gauge* invariants of an arbitrary gauge field can be constructed from the gauge invariants formed from L^{ij} and T^{ijk} alone. For Abelian gauge fields, however, we have above a reduction algorithm that enables one to express any Lorentz-gauge invariant as a polynomial in the fundamental set of invariants (since in this case, the action of G on L^{ij} and T^{ijk} is the identity). In particular for the Maxwell field, $L = F_{\mu\nu}F^{\nu\mu} + iF^*{}_{\mu\nu}F^{\nu\mu}$ and $T = 0$, and so we have shown that $F_{\mu\nu}F^{\nu\mu}$ and $F^*{}_{\mu\nu}F^{\nu\mu}$ form a polynomial basis for the invariants.[12]

APPENDIX B

The simplest way of proving (3.9) seems to be by replacing the SL(2,C) spinor indices in (3.1) and (3.2) by complex O(3) indices. Using

$$\delta_{pr} \leftrightarrow \epsilon_{(A|C|}\epsilon_{B)D}, \tag{B1}$$

where $p \leftrightarrow (A,B)$, $r \leftrightarrow (C,D)$, and (4.8), we can write

$$L^{ij} = -2\phi^i{}_p\phi^j{}_p, \tag{B2}$$

and

$$T^{ijk} = -2\sqrt{2}\epsilon_{pqr}\phi^i{}_p\phi^j{}_q\phi^k{}_r, \tag{B3}$$

where p, q, r are now complex O(3) indices.[13] It follows that

$$T^{ijk}T^{lmn} = 8\epsilon_{pqr}\phi^i{}_p\phi^j{}_q\phi^k{}_r\epsilon_{stu}\phi^l{}_s\phi^m{}_t\phi^n{}_u$$
$$= -6L^{[i|l|}L^{j|m|}L^{k]n},$$

where we have used (B2) and the identity

$$\epsilon_{pqr}\epsilon_{stu} = 6\delta_{[p|s|}\delta_{q|t|}\delta_{r]u} = \begin{vmatrix} \delta_{ps} & \delta_{pt} & \delta_{pu} \\ \delta_{qs} & \delta_{qt} & \delta_{qu} \\ \delta_{rs} & \delta_{rt} & \delta_{ru} \end{vmatrix}. \tag{B4}$$

[1]T. Eguchi, Phys. Rev. D **13**, 1561 (1976). There are some errors in this paper which are pointed out in J. Anandan, PITT-190. See also J. Anandan, Ph.D. thesis, University of Pittsburgh (1978).

[2]R. Roskies, Phys. Rev. D **15**, 1722 (1977).

[3]M. Carmeli, Phys. Rev. Lett. **39**, 523 (1977). There are serious errors in this letter which are pointed out in Ref. 5.

[4]J. Anandan and K.P. Tod, Phys. Rev. D **18**, 1144 (1978).

[5]J. Anandan and R. Roskies, Phys. Rev. D **18**, 1152 (1978).

[6]L.L. Wang and C.N. Yang, Phys. Rev. D **17**, 2687 (1978).

[7]For an introduction to spinors, see for instance F.A.E. Pirani in *Lectures on General Relativity, Brandeis Summer Institute in Theoretical Physics*, edited by S. Deser and K.W. Ford (Prentice Hall, New Jersey, 1965), I, Chap. 3.

[8]It is possible to have an algebraic function of $F^i_{\mu\nu}$ which is invariant under the Lorentz group only for special cases of the field $F^i_{\mu\nu}$. Such an "invariant" would be a tensor in which not all the space–time indices are contracted, or would be a function of such tensors. In the present paper, however, a Lorentz invariant means a gauge tensor in which all the space–time indices are contracted, or which is an algebraic function of such tensors. The limit of such a function, when some of the tensors tend to zero, is *not* considered to be an invariant in this paper.

[9]F.R. Gantmacher, *Applications of the Theory of Matrices* (Interscience, New York, 1959), Chap. I, Sec. 3. I would like to thank K.P. Tod for bringing this reference to my attention.

[10]In the proof of Theorem 4 it is convenient to use the canonical form for a complex 3×3 symmetric matrix that was given below (3.26). Notice that, under Lorentz transformations, the gauge invariant symmetric matrix $\widetilde{M}$ transforms as $\widetilde{M}\to O\widetilde{M}O^T$ where O is a complex orthogonal matrix. Hence, $\widetilde{M}$ can be made to have one of the canonical forms, (a), (b), or (c) [given below (3.26)] by a suitable Lorentz transformation. These three cases correspond to M (or $\widetilde{M}$) having respectively, three, two, or one linearly independent eigenvectors, which provides a classification of the SU(2) gauge field into three types. Another classification of this field in terms of the number of linearly independent eigenvectors of the matrix L is also possible, although in this case L cannot, in general, be made to have one of the above mentioned canonical forms by means of a gauge transformation.

[11]However M and the matrix P defined by $P_\alpha^{\beta'} = -2\phi^i_\alpha\bar{\phi}^{i\beta'}$ have information about all the gauge invariants, in some sense.

[12]A reduction algorithm for the Maxwell field has also been constructed by Predrag Cvitanović, using a different method which does not use spinors (privae communication).

[13]The remaining argument in the proof in Appendix B is due to P. Cvitanović (private communication).

[14]After this paper was submitted for publication, I learned that the gauge invariants for the self-dual Yang-Mills field has also been investigated by M.B. Halpern, Phys. Rev. D **16**, 3515 (1977).

Exceptional gauge groups and quantum theory[a]

L. P. Horwitz[b)]

Tel Aviv University, Ramat Aviv, Israel

L. C. Biedenharn[c)]

Duke University, Durham, North Carolina 27706
(Received 16 June 1978)

It is shown that a Hilbert space over the real Clifford algebra C_7 provides a mathematical framework, consistent with the structure of the usual quantum mechanical formalism, for models for the unification of weak, electromagnetic and strong interactions utilizing the exceptional Lie groups. In particular, in case no further structure is assumed beyond that of C_7, the group of automorphisms leaving invariant a minimal subspace acts, in the ideal generated by that subspace, as G_2, and the subgroup of this group leaving one generating element (e_7) fixed acts, in this ideal, as the color gauge group SU(3). A generalized phase algebra $\mathfrak{A}\, C_7$ is defined by the requirement that quantum mechanical states can be consistently constructed for a theory in which the smallest linear manifolds are closed over the subalgebra $\mathbb{C}(1,e_7)$ (isomorphic to the complex field) of C_7. Eight solutions are found for the generalized phase algebra, corresponding (up to an overall sign), in effect, to the use of $\pm e_7$ as imaginary unit in each of four superselection sectors. Operators linear over these alternative forms of imaginary unit provide distinct types of "lepton–quark" and "quark–quark" transitions. The subgroup in $\mathfrak{A}$ which leaves expectation values of operators linear over $\mathfrak{A}$ invariant is its unitary subgroup U(4), and is a realization (explicitly constructed) of the U(4) invariance of the complex scalar product. An embedding of the algebraic Hilbert space into the complex space defined over $\mathbb{C}(1,e_7)$ is shown to lead to a decomposition into "lepton" and "quark" superselection subspaces. The color SU(3) subgroup of G_2 coincides with the SU(3) subgroup of the generalized phase U(4) which leaves the "lepton" space invariant. The problem of constructing tensor products is studied, and some remarks are made on observability and the role of nonassociativity.

1. INTRODUCTION

The notion of non-Abelian gauge fields has become a useful theoretical tool in recent years. The fundamental idea of Yang and Mills[1] now plays an important role in models for the strong interactions,[2] where the global gauge group is called the "color group"[3] (as distinguished from the "flavor group" associated with hadron multiplets), and in renormalizable models for the weak and electromagnetic interactions.[4] In this framework, it appears possible to attempt a unification of the strong, electromagnetic and weak interactions by utilizing a gauge group which acts on the leptons and the color and flavor degrees of freedom of the quark fields, the spinor constituents of hadrons.

Gürsey and his co-workers have suggested[5] that the space of internal degrees of freedom of leptons and quarks, which form a basis for a gauge group of this type, may correspond to the space of exceptional quantum mechanical states discovered by Jordan, von Neumann, and Wigner[6] and that, in this framework, a spontaneously broken gauge field theory based on the exceptional Lie group E_7 is a reasonable candidate. The quantum mechanical spaces in which the transformation groups of algebraic automorphisms are F_4, E_6, E_7, and E_8 may be represented as matrices incorporating 3×3 submatrices with octonion (Cayley-number)-valued elements.[7] The generalized projective geometries associated with these spaces do not satisfy Desargues' theorem,[8] and Gürsey has suggested that this may lead to unusual consequences for the observability of the corresponding quark states. The interpretation of such a structure in terms of physical measurements, presumably associated with the question of "confinement," has not yet been worked out. Since the color singlet parts of given repesentations of the exceptional groups lie in a Desarguesian subset, however, they should be observable in the usual sense, and Gürsey associates them with the normal leptons and hadrons.

The possibility that internal degrees of freedom are associated with exceptional quantum mechanical states was considered several years ago. Stimulated by the work of Pais,[9] attempting to calssify hadrons with the help of the octonion algebra, and certain open questions raised by the work of Jordan, von Neumann, and Wigner,[6] Goldstine and Horwitz[10] defined the notion of a Hilbert space over Cayley numbers (octonions) and studied some of its properties. Immediate difficulties in the realization of such spaces in finite dimensions, due to the nonassociative property of Cayley numbers, were circumvented by the use of a real scalar product, and a spectral theorem was proved for a certain class of self-adjoint operators.[11] One must, however, consider the closure of linear manifolds under the action of multiplication by the elements of the nonassociative Cayley algebra (for example, in order to obtain Fourier series expansions); it was shown[10] that every vector generates a linear manifold over the reals of at most 128 dimensions, and that the basis for this manifold can provide a faithful representation of C_7,

[a)]A brief preliminary report of this work was given at the VI International Colloquium on Group Theoretical Methods in Physics, Tubingen, July, 18–22, 1977 (to appear in the Proceedings).
[b)]Research supported in part by the Binational Science Foundation (BSF) Jerusalem, Israel.
[c)]Work supported in part by the National Science Foundation.

 0022-2488/79/020269-30$01.00

the Clifford algebra of order seven. The appearance of a linear vector space over a finite associative algebra raised the question of the structure of Hilbert spaces over finite associative algebras. The general theory was worked out by Goldstine and Horwitz.[12]

In the special case of C_7, it was found that multiplication defined as an *equivalence relation* in a (minimal) one-sided ideal recovers the form of the nonassociative algebra. As pointed out above, the exceptional Lie groups proposed as general gauge groups unifying the weak, electromagnetic and strong interactions, containing both color and flavor degrees of freedom, arise as automorphisms of algebras constructed of matrices with octonion valued elements. This equivalence relation, however, permits us to reformulate the construction of these automorphisms in terms of matric algebras with elements in the associative algebra C_7. Since octonion multiplication rules are reproduced in a minimal ideal, *the automorphisms arise as the set of transformations which leave that ideal invariant.* For example, in the case that we shall study in detail here with matrices of dimension one (with no flavor degrees of freedom), the group of automorphisms which leaves invariant one minimal ideal of C_7 acts, in that ideal, as G_2. The properties desired of Hilbert spaces with octonion multipliers (structures which appear to be difficult to interpret at the present time) can therefore be studied in the framework of Hilbert spaces over associative algebras, for which a quantum mechanical interpretation of the ususal type is available.

Horwitz and Biedenharn[13] have shown that the propositional calculus associated with the algebraically closed linear manifolds of an algebraic Hilbert space of this type constitute a complete, weakly modular, orthocomplemented, atomic lattice, and therefore satisfies the axioms[14,15] of the calculus of propositions characterizing quantum mechanics. The fact that such a lattice can be embedded in a family of Hilbert spaces over a field Φ [15] was used[13] to show that the quantum theory described by a Hilbert space over an arbitrary finite algebra, in which the observables are linear with respect to the quantities of the algebra, is isomorphic to a quantum theory described by a Hilbert space over a field Φ in which there are superselection rules.

In this paper, we shall adapt a fundamental idea of Gürsey and Günaydin,[16] namely, the selection of a particular element of the Cayley algebra to represent the imaginary unit,[17] to the framework of the associative algebra C_7. The subalgebra $\mathbb{C}(1,e_7)$ of C_7 generated by unity and one of the generating elements, e_7, of C_7 (satisfying $e_7^2 = -1$), over the reals, is isomorphic to the complex field. The scalar product defined by the requirement of orthogonality between linear manifolds closed over $\mathbb{C}(1,e_7)$ is shown to be that given by Günaydin.[16] We then proceed to study the requirements for the construction of a physical state in a quantum mechanical framework in which the smallest linear manifold is spanned over the subalgebra $\mathbb{C}(1,e_7)$ (the field Φ) by a single vector, i.e., a structure isomorphic to the usual ray in complex Hilbert space. These requirements admit eight solutions for a generalized phase algebra $\mathfrak{A} \subset C_7$, analogous to the complex phase in the usual complex Hilbert space, which could be utilized to construct local gauge transformations. Each of these solutions is characterized by the fact that it commutes with one of a set of eight "imaginary" units closely related to e_7 (one of these is e_7 itself); they correspond, in fact, to the use of $\pm e_7$, with independent choice of signs, in each of four superselection sectors, up to an over-all sign. These distinct solutions provide alternative structures for operators which induce "lepton–quark" and "quark–quark"[18] transitions and their associated gauge algebras. Representations for the algebras are explicitly constructed. They provide a realization of the U(4) symmetry of the scalar product. The part of this U(4) of algebraic phases which represents transformations leaving invariant one minimal subspace (for which the equivalence relations discussed above define the nonassociative multiplication rules of the Cayley algebra) and the Clifford element (for example) e_7, is a U(3) which acts, in the ideal generated by this minimal subspace, as SU(3), coinciding with the color gauge subgroup of G_2.[16]

A Hilbert space over C_7, with linear manifolds closed over an algebra $\mathfrak{A} \subset C_7$ may be embedded in a family of Hilbert spaces over the complex field $\mathbb{C}(1,e_7)$, following the procedure used in our earlier work.[13] Defining "observables" as the self-adjoint operators linear over $\mathfrak{A}$, the embedding results in a quantum theory defined on a family of complex Hilbert spaces labelled by a superselection rule; these superselection spaces transform into each other under $\mathfrak{A}$, as a representation for the U(4) for which the complex scalar product is invariant. These spaces can be identified with the "leptonic" (observable) and the "quark" (unobservable) subspaces utilized by Gürsey and Günaydin.[16] We do not offer rigorous arguments on why states in the "quark" space are unobservable, but this demonstration shows the existence and interpretation of superselection rules which may provide a useful mathematical framework for the description of such phemomena.

We remark that the structure outlined above is predicated on the choice of a direction in the parameter space of G_2, i.e., the direction of e_7, which breaks the symmetry down to SU(3). The complex Hilbert space constructed over $\mathbb{C}(1,e_7)$ is therefore parameterically dependent on the choice of this direction. The full automorphism group of the minimal right ideal which defines G_2 is therefore represented by a family of such complex Hilbert spaces, with the direction of e_7 as the superselection parameter labelling the separate components that transform into each other under the action of operators which are linear only over the reals (including those of G_2).[19]

The use of complex-valued [in $\mathbb{C}(1,e_7)$] wavefunctions to express the quantum mechanical content of the algebraic Hilbert space facilitates the construction of tensor products (for the construction of many-body states). The situation is somewhat complicated, however, by the fact that we are working with a collection of complex Hilbert spaces that transform into each other under the action of operators that are for example, linear over $\mathbb{C}(1,e_7)$. In particular, transitions[20,21] from the "quark" or "unobservable" space to the "lepton" or "observable" (color singlet) space, and vice-versa, are accompanied by complex conjugation. Operators linear over $\mathbb{C}(1,e_7)$ on the whole space behave antilinearly, in the manner of Wigner's corepresentations,[22] on the component Hilbert spaces for transitions of this type. One of the conditions on a tensor product that one might require, namely, that there exist an operator on the tensor product space which is equivalent to the mapping induced by the action of reasonably well-behaved operators on each of the constituent spaces, is not, for example, satisfied for a tensor product of the type (antiassociator) discussed by Günaydin.[16] In this paper, we shall discuss the problem raised above in some detail, and provide a prescription for the construction of tensor products which is consistent for operators linear over $\mathbb{C}(1,e_7)$.

The plan of the paper is as follows. In Sec. II, we review briefly the special structure of an algebraic Hilbert space over the associative algebra C_7; by studying the orthogonality relations for the class of linear manifolds that is closed over the subalgebra $\mathbb{C}(1,e_7)$, we obtain the corresponding scalar product in the form given by Günaydin.[16] In sec. III, the properties of operators linear over $\mathbb{C}(1,e_7)$ and the superselection rules that they satisfy are discussed. In Sec. IV, the sesquilinear forms corresponding to quantum states are constructed in the most general form consistent with Gleason's theorem[23]; following a procedure established in our earlier work,[13] we characterize the generalized "phase" algebras $\mathfrak{A}\subset C_7$, and discuss the pure states. In Sec. V, the solutions for $\mathfrak{A}$ are made explicit, and superselection rules for operators linear over $\mathfrak{A}$ are displayed by embedding the Hilbert space[13] over C_7 into the complex Hilbert space over $\mathbb{C}(1,e_7)$. It is shown that the group contained in $\mathfrak{A}$ which leaves expectation values of operators linear over $\mathfrak{A}$ invariant is U(4), and is a realization of the U(4) invariance of the complex scalar product. The relation between the linear U(4) and the corepresentation quality of the transformations induced by it on the wavefunctions is explicitly given. It is also shown that the automorphisms of C_7 which leave invariant the subspace defining the Cayley multiplication rules act like G_2 in the ideal generated by that subspace, and the subset of these automorphisms which also leave e_7 invariant act like SU(3) in this ideal; this SU(3) is the intersection of the unitary subgroup U(4) in $\mathfrak{A}$ (which leaves the complex scalar product invariant) with G_2. A discussion of the construction of tensor products is given in Sec. VI, and, in a concluding section, we make some remarks on observability and the role of nonassociativity.

II. ALGEBRAIC HILBERT SPACE OVER C_7

We denote by $e_1,e_2,\ldots,e_7$ the generating elements of the associative real Clifford algebra C_7. These elements have the property

$$\{e_a,e_b\} = e_ae_b + e_be_a = -\delta_{ab}\cdot 2, \tag{2.1}$$

and an involution (called "conjugate")

$$e^*_a=-e_a,\quad (e_ae_b)^*=e^*_be^*_a. \tag{2.2}$$

If an element a of C_7 satisfies $a^*=a$, it is said to be symmetric.

We shall take this algebra, with unity quantity, as the set of "scalars" in a Hilbert space $\mathscr{H}$. The properties of such Hilbert spaces (for general finite associative algebras) were discussed by Goldstine and Horwitz[12]; definitions and some of the results will be restated here for convenience.

p. 1: $\mathscr{H}$ is a linear space in the sense that there exists an operation called multiplication by a scalar $f\cdot a$ for every a in C_7 (that is, 1, $e_1,\ldots,e_7$, $e_1e_2,\ldots,e_6e_7,e_1e_2e_3,\ldots,e_1e_2e_3e_4e_5e_6e_7$, and real linear combinations of these) and f in $\mathscr{H}$, with values in $\mathscr{H}$. There is also an operation called addition of vectors $f+g$, defined for every f, g, in $\mathscr{H}$ with values in $\mathscr{H}$. These properties are associative and distributive in the usual way.

It is well known[24] that the Clifford algebras are matrix algebras, i.e., with finite matrix representations over the reals. In any such representation, the usual concepts of the trace and positivity are well defined. We may therefore introduce an inner product process in $\mathscr{H}$.

p. 2: There exists an inner product (f,g) defined for all f,g in $\mathscr{H}$ with values in C_7 such that

$(f+g,h)=(f,h)+(g,h)$,

$(f,g)^*=(g,f)$,

$(f,f)\geqslant 0$ (symmetric), and it is zero if and only if $f=0$,

$(f,ga)=(f,g)a$ for $a\in C_7$

We define the modulus of a to be

$$|a|=\operatorname{tr}(aa^*)^{1/2}\geqslant 0, \tag{2.3}$$

where tr is the usual trace function (normalized so that $\operatorname{tr}1=1$); $|a|=0$ if and only if $a=0$. With the help of (2.3), we may define a norm for the space $\mathscr{H}$:

$$\|f\|^2=|(f,f)|=\operatorname{tr}(f,f). \tag{2.4}$$

The Schwarz inequality

$$|(f,g)|\leqslant\|f\|\cdot\|g\| \tag{2.5}$$

is valid. With the norm (2.4), we may construct Cauchy sequences. We state the completeness postulate.

p. 3: The space $\mathscr{H}$ is complete, i.e., every Cauchy sequence in $\mathscr{H}$ has its limit in $\mathscr{H}$.

We now state some geometrical notions which are central to our study,

Definition 2.1: The vectors f and g are said to be orthogonal in case $(f,g)=0$.

Definition 2.2: The set M is said to be a linear closed manifold in case it is closed and contains, along with f,g, $fa+gb$, where a, $b\in C_7$.

Let M be a linear closed manifold and f an arbitrary vector. Then[12] there is a unique decomposition of f into

$$f=g+h, \tag{2.6}$$

where g is in M and h is in $\mathscr{H}-M$. Given M and f, we say that P_Mf is the g in M existence is asserted in (2.6). The projection operator P_M is totally linear,[13] i.e.,

$$P_M(fa)=(P_Mf)a \tag{2.7}$$

for all $a\in C_7$, and it satisfies

$P_M(P_Mf)=P_Mf$,

$$(f, P_M g) = (P_M f, g) \;, \tag{2.8}$$

The projection operation P_M with properties (2.7), (2.8) have the same order relations that obtain among projection operators in a complex Hilbert space, i.e., for all f in $\mathscr{H}$,

$$\|P_M f\|^2 \leqslant \|P_N f\|^2 \quad \text{for} M \subset N. \tag{2.9}$$

In this case, $P_M P_N = P_M$, and we say that $P_M \leqslant P_N$.

The set of all closed linear manifolds on a Hilbert space over a finite associative algebra with unity quantity was shown by Horwitz and Biedenharn[13] to be a complete, orthocomplemented, weakly modular atomic lattice, and therefore to satisfy the axioms of the usual quantum theory[15] (the atoms are the one-dimensional linear manifolds generated by the minimal right ideals of C_7 on each f in $\mathscr{H}$). Such a lattice can be embedded in a Hilbert space over a field.[15] Embedding the closed linear manifolds of $\mathscr{H}$ in a Hilbert space $\mathscr{H}_R$ over the reals, one finds that the corresponding quantum theory, with totally linear observables, contains superselection sectors labelled by the minimal right ideals.[13]

We wish to make explicit the representation of C_7 in which we shall be working.[10] Since $e_1e_2e_3e_4e_5e_6e_7 \equiv E_7$ commutes with all of $e_1,\dots,e_7$ and $E_7^2 = I$, we may split the space $\mathscr{H}$ in an invariant way,

$$\mathscr{H} = \mathscr{H}_+ + \mathscr{H}_-,$$

where

$$f_+ = f_+P_+, \quad f_- = f_-P_-$$

and

$$P_\pm = \tfrac{1}{2}(I \mp E_7) \tag{2.10}$$

The decomposition (2.10) is invariant under the action of totally linear operators [with the property (2.7)] and under right multiplication by any element of C_7. In each subspace, the seven generating elements are not completely independent, but satisfy

$$e_7 P_\pm = \pm (e_1 \cdots e_6) P_\pm, \tag{2.11}$$

and an irreducible representation of the remaining C_6 algebra can be obtained in each of $\mathscr{H}_+$ and $\mathscr{H}_-$ by the usual quantum mechanical procedure of diagonalizing a complete set of symmetric commuting "operators." We take these to be

$$e_1e_2e_3, \quad e_5e_1e_6, \quad e_6e_2e_4. \tag{2.12}$$

The minimal projection[25] in the C_6 algebra, for which the three operators in (2.12) take on the value -1, is

$$P_0 = \tfrac{1}{8}(I - e_1e_2e_3)(I - e_5e_1e_6)(I - e_6e_2e_4). \tag{2.13}$$

The set of eight minimal projections spanning (the right space of) $\mathscr{H}_+$ (or of $\mathscr{H}_-$) is then

$$P_i = e_i P_0 e_i^*, \quad i = 0,1,\dots,7, \tag{2.14}$$

each choice of i corresponding to a distinct combination of signs replacing the negative signs in (2.13). Then,

$$I = \sum_i P_+ P_i + \sum_i P_- P_i. \tag{2.15}$$

We may now define a set of "multiplication laws" for elements of the Clifford algebra in each minimal subspace by means of the *equivalence relations*[12]

$$P_\pm P_i ab = P_\pm P_i (ab)_i^\pm, \tag{2.16}$$

where each of a, b, and $(ab)_i^\pm$ are of the form

$$a = \sum_{i=0}^{7} a_i e_i, \tag{2.17}$$

with a_i real. These relations may be used to reduce any element of C_7 to the form (2.17) by iterative use of (2.16). Of the sixteen equivalence relations (2.16), we write explicitly only the results for $(ab)_0^+$ here. Consider, for example,

$$\begin{aligned} P_+P_0e_1e_2 &= P_+P_0\tfrac{1}{2}(I - e_1e_2e_3)e_1e_2 \\ &= P_+P_0\tfrac{1}{2}(e_1e_2 + e_3\ ) \\ &= P_+P_0\tfrac{1}{2}(I - e_1e_2e_3)e_3, \end{aligned} \tag{2.18}$$

so that

$$(e_1e_2)_0^+ = e_3. \tag{2.19}$$

Together with the relation (2.11), (2.19) and similarly derived relations for the other pairs one finds the nonassociative multiplication laws for octonions in the form utilized by Gürsey and Günaydin[16]

$$\begin{aligned} &(e_1e_2)_0^+ = e_3, \quad (e_5e_1)_0^+ = e_6, \quad (e_6e_2)_0^+ = e_4, \\ &(e_4e_3)_0^+ = e_5, \\ &(e_7e_1)_0^+ = e_4, \quad (e_7e_2)_0^+ = e_5 \quad (e_7e_3)_0^+ = e_6, \\ &\qquad\qquad + \text{cyclic}. \end{aligned} \tag{2.20}$$

For values of i other than 0, there are similar rules to (2.20), with differing signs that can be generated by inner automorphisms of the algebra. The space $\mathscr{H}_-$ provides an inequivalent set obtained by the opposite convention for the sign of multiplication rules involving e_7.

It is easy to verify[10] that the basis set

$$\rho_{ij}^\sigma = e_i P_0 e_j^* P_\sigma \quad (\sigma = \pm 1), \tag{2.21}$$

where $\rho_{ii}^\sigma = P_i P_\sigma$, satisfies

$$\rho_{ij}^\sigma \rho_{kl}^{\sigma'} = \delta_{\sigma\sigma'} \delta_{jk} \rho_{il}^\sigma, \tag{2.22}$$

and that any element of C_7 can be represented as

$$a = \sum_{ij,\sigma} K_{ij}^\sigma(a) \rho_{ij}^\sigma,$$

where

$$K_{ij}^\sigma(a) = \sum_{k=0}^{7} \rho_{ki}^\sigma a \rho_{jk}^\sigma \tag{2.23}$$

We have, so far, discussed right multiplication of the vectors of $\mathscr{H}$ by the "scalars" $a \in C_7$. We shall, however, also admit left multiplication; since the scalar product is not defined to be linear under left multiplication by elements of the Clifford algebra, these elements act like nontrivial operators. We postulate that these operators are totally linear (verifiable in finite dimensions):

$p.\ 4$: $a(fb) = (af)b$,

where $a, b \in C_7$. Symmetric idempotents in C_7 in left multiplication satisfy all of the properties of the operator valued projections discussed in connection with (2.6) and, in fact, belong to this class. Since $P_\pm$ are invariant, we shall also postulate

$p.\ 5$: $P_\pm f = f\ P_\pm \in \mathscr{H}_\pm$,

and, in what follows, we shall work entirely in the subspace $\mathscr{H}_+$ and suppress the index $\sigma = \pm$, unless otherwise stated.

Let us define

$$f_{ij} = \sum_{k=0}^{7} \rho_{ki} f \rho_{jk}, \tag{2.24}$$

and therefore

$$f = \sum_{ij} f_{ij} \rho_{ij}. \tag{2.25}$$

It then follows that

$$(f,g) = \sum_{ijk} \rho_{ij} (f_{ki}, g_{kj})$$

$$= \sum_{ij} \rho_{ij} K_{ij}((f,g)), \tag{2.26}$$

since

$$K_{ij}(f,g)\ ) = \sum_{l} \rho_{li} (f,g)\ \rho_{jl}$$

$$= \sum_{l} (f\rho_{il}, g\rho_{jl})$$

$$= \sum_{l,l',k} (\rho_{lk} f \rho_{il}, \rho_{l'k}\ g \rho_{jl'}),$$

which verifies (2.26). Equation (2.26) provides a constructive realization of the algebra-valued scalar product (f,g).

An operator A in $\mathscr{H}$ may, among other things, multiply a vector f by a scalar $a \in C_7$ on the right, i.e., $Af = fa$. This is not possible for a totally linear operator unless a is in the center of C_7 (I or E_7) since otherwise $A\,(fb) = fba \neq fab$; $\forall b \in C_7$. A bounded totally linear Hermitian operator on $\mathscr{H}$ i.e., satisfying

$$A\,(fb) = (Af)b,$$

$$(f,Ag) = (Af,g), \tag{2.27}$$

has a spectral resolution of the form[12]

$$A = \int \lambda dP(\lambda), \tag{2.28}$$

where the $P(\lambda)$ are a totally linear spectral family. It is also possible to consider a theory in which the Hermitian operators are linear only over the reals, and satisfy the weaker condition

$$\mathrm{tr}(Af,g) = \mathrm{tr}(f,Ag). \tag{2.29}$$

It was shown in Ref. 12 that a Hilbert space $\mathscr{H}_R$ in which the scalar product is taken to be $\mathrm{tr}(f,g)$ (real), and for which the Hermitian operators satisfy (2.29), is isomorphic to a real Hilbert space. The linear manifolds relevant to a quantum theory in which Hermitian operators satisfying (2.29) are admitted as observables are closed over the reals.

A theory in which only the *totally linear* Hermitian operators are admitted as observables appears in the Hilbert space $\mathscr{H}_R$ as a quantum theory with superselection rules indexed by the minimal right ideals.[13]

We have so far considered two extreme algebraic requirements: (i) that the linear manifolds corresponding to the quantum lattice be closed only over the reals, resulting in a real Hilbert space and "observables" that are linear over the reals, and (ii) that the linear manifolds corresponding to the quantum lattice be closed over the full algebra of scalars C_7 to the right, resulting in an algebraic Hilbert space over C_7.

Gürsey and Günaydin[16] have argued, however, that operators linear over the complex field, such as occur in the unitary representations of Lie groups, and the extraction of the complex subalgebra itself, are important in the structure of gauge field theories of the type we are considering. The construction of tensor product spaces linear over the full algebra, or a non-Abelian subalgebra, furthermore, appears to be very difficult[17]; the use of a complex subalgebra is, in this sense, maximal. Gürsey and Günaydin[16] suggested, in the context of the algebra of octonions, that complex-valued wavefunctions, linear over a complex field defined as a subalgebra of the full algebra, be used. We shall adapt their suggestion to the framework of the Hilbert space over a Clifford algebra, selecting the subalgebra $\mathbb{C}(1,e_7)$ of C_7, generated by 1 and e_7, which is isomorphic to the algebraically closed field of complex numbers. The algebraic Hilbert space $\mathscr{H}$ can then be (somewhat more weakly than for $\mathscr{H}_R$) embedded in a Hilbert space $\mathscr{H}_c$, with scalar product linear over $\mathbb{C}(1,e_7)$, for which tensor products can also be consistently defined.

We wish now to construct a Hilbert space for which the scalar product is linear, and linear manifolds are closed, over $\mathbb{C}(1,e_7)$. To obtain a systematic procedure for constructing scalar products suitable for certain classes of linear manifolds, we consider the orthogonality definition 2.1 again. We assert that if

$$\mathrm{tr}((f,g)a) = 0 \tag{2.30}$$

for all $a \in C_7$, then f is orthogonal to g. In particular, (2.30) implies that

$$\mathrm{tr}((f,g)\rho_{ij}) = 0$$

or

$$0 = \sum_{kl} \mathrm{tr} K_{kl}(\ (f,g)\ ) \rho_{kl} \rho_{ij}$$

$$= K_{ji}(\ (f,g)\ ),$$

hence $(f,g)=0$. It is furthermore clear that if $(f,g)=0$, then $(fa,gb)=0$ for any $a,b \in C_7$, and therefore the algebraically closed linear manifolds generated by f,g are orthogonal. On the other hand, if we define orthogonality by (2.30) with a restricted to the real multiples of the identity, then we obtain the scalar product of the real Hilbert space $\mathscr{H}_R$, i.e., we say that f is real orthogonal to g if

$$\mathrm{tr}(f,g)=0. \tag{2.31}$$

It is not true that Eq. (2.31) implies $\mathrm{tr}((f,g)a)=0$ for every $a \in C_7$.

Let us now define complex orthogonality with

Definition 2.3: f and g are said to be complex orthogonal if $\mathrm{tr}((f,g)z)=0$ for all $z \in \mathbb{C}(1,e_7)$.

It is clear that the linear manifolds generated by f,g, and closed over $\mathbb{C}(1,e_7)$ are complex orthogonal if f and g are complex orthogonal, since $\mathrm{tr}(fz,gz')=\mathrm{tr}(z^*(f,g)z')=\mathrm{tr}((f,g)zz^*)=\mathrm{tr}((f,g)z'')=0$, when z, z', and $z''=zz^* \in \mathbb{C}(1,e_7)$. Hence the scalar product of Definition 2.3 is suitable for the construction of a Hilbert space with closed linear manifolds defined over $\mathbb{C}(1,e_7)$.

Since every $z \in \mathbb{C}(1,e_7)$ is of the form $\alpha+\beta e_7$ (α,β real), it suffices for the definition of complex orthogonality that

$$\mathrm{tr}((f,g))=0,$$
$$\mathrm{tr}((f,g)e_7)=0. \tag{2.32}$$

We shall now show that the requirement (2.32) is equivalent to the scalar product given by Günaydin.[16]

From (2.26)

$$\mathrm{tr}(f,g)=\sum_{k,i}(f_{ki},g_{ki}), \tag{2.33}$$

and

$$\mathrm{tr}((f,g)e_7)=\sum_{kij}(f_{ki},g_{kj})\,\mathrm{tr}(\rho_{ij}e_7). \tag{2.34}$$

According to (2.21)

$$\mathrm{tr}(\rho_{ij}e_7)=\mathrm{tr}(e_iP_0e_j^*e_7)$$

$$=\mathrm{tr}(P_0(e_j^*e_7e_i)); \tag{2.35}$$

to evaluate this expression, we may use the multiplication rules (2.20). In fact,

$$P_0e_j^*e_7e_i=P_0(e_j^*e_7)_0e_i$$

$$=P_0((e_j^*e_7)_0e_i)_0,$$

and therefore (2.35) vanishes unless $(j,7,i)$ are in the same quaternion subalgebra of (2.20). A simple calculation yields

$$\mathrm{tr}((f,g)e_7)=\sum_k[(f_{k0},g_{k7})-(f_{k7},g_{k0})$$
$$+(f_{k1},g_{k4})-(f_{k4},g_{k1})+(f_{k2},g_{k5})$$
$$-(f_{k5},g_{k2})+(f_{k3},g_{k6})-(f_{k6},g_{k3})]. \tag{2.36}$$

We remark that (for each value of the index k) (2.33) is invariant under SO(8), and (2.36) under Sp(8); the symmetry of the full scalar product (2.32) is therefore SO(8) $\cap$Sp(8) = U(4). To make explicit the identity of this scalar product with that of Gunaydin,[16] we write the vector representation (2.25), with the help of the multiplication rules (2.20), in the form

$$f=\sum f_{kj}e_kP_0e_j^*$$

$$=\sum_k e_kP_0[(f_{k0}-f_{k7}e_7)-(f_{k1}+f_{k4}e_7)e_1$$

$$-(f_{k2}+f_{k5}e_7)e_2-(f_{k3}+f_{k6}e_7)e_3]. \tag{2.37}$$

Let us call

$$\psi_0^k=f_{k0}-f_{k7}e_7,$$
$$\psi_\alpha^k=-f_{k\alpha}-f_{k\alpha+3}e_7, \tag{2.38}$$

so that

$$f=\sum_k e_kP_0\left(\psi_0^k+\sum_{\alpha=1}^3\psi_\alpha^ke_\alpha\right). \tag{2.39}$$

Note that, for $f\to fz$,

$$\psi_0^k\to\psi_0^kz,\quad \psi_\alpha^k\to\psi_\alpha^kz^*. \tag{2.40}$$

In terms of the definitions (2.38), the two parts of the scalar product (2.32) are (χ corresponds to g)

$$\mathrm{tr}((f,g))$$
$$=\sum_k\left[\mathrm{Re}(\psi_0^k,\chi_0^k)+\sum_{\alpha=1}^3\mathrm{Re}(\psi_\alpha^k,\chi_\alpha^k)\right] \tag{2.41}$$

and

$$\mathrm{tr}((f,g)e_7)$$
$$=\sum_k\left[-\mathrm{Im}(\psi_0^k,\chi_0^k)+\sum_{\alpha=1}^3\mathrm{Im}(\psi_\alpha^k,\chi_\alpha^k)\right] \tag{2.42}$$

where we have used the fact that scalar products of vectors of the form (2.38) belong to $\mathbb{C}(1,e_7)$ and have adopted the usual terminology for real and imaginary parts (complex conjugation is the star operation of C_7). The condition that (2.41) and (2.42) both vanish, i.e., that f and g are complex orthogonal, can be more concisely written if we define

$$(f,g)_c=\mathrm{tr}((f,g))-e_7\mathrm{tr}((f,g)e_7) \tag{2.43}$$

and require

$$(f,g)_c=0, \tag{2.44}$$

Combining (2.41) and (2.42) according to (2.43), we obtain

$$(f,g)_c$$
$$=\sum_k\left[(\psi_0^k,\chi_0^k)+\sum_{\alpha=1}^3(\chi_\alpha^k,\psi_\alpha^k)\right]. \tag{2.45}$$

For each k,[26] the scalar product defined by (2.45) is precisely that given by Günaydin.[16] We note that since $\mathrm{tr}[(f,f)e_7]=0$,

$$(f,f)_c = \mathrm{tr}(f,f) = \|f\|^2, \qquad (2.46)$$

and hence the topology of $\mathscr{H}_c$ is the same as that of $\mathscr{H}_+$.

III. COMPLEX LINEAR OPERATORS

It follows from Eqs. (2.40) and (2.45) that the scalar product $(f,g)_c$ is linear over $\mathbb{C}(1,e_7)$ in the sense

$$\begin{aligned}(f,gz)_c &= (f,g)_c z\\ (fz,g)_c &= z^*(f,g)_c,\end{aligned} \qquad (3.1)$$

for $z \in \mathbb{C}(1,e_7)$. It is furthermore clear from the discussion of Sec. II that projection operators defined on linear manifolds closed over $\mathbb{C}(1,e_7)$, and with orthogonality determined by the scalar product $(f,g)_c$, are linear operators with respect to $\mathbb{C}(1,e_7)$. In addition to operators linear over the reals, or over the entire algebra of C_7, we shall therefore study a third type of linearity, that of operators linear over $\mathbb{C}(1,e_7)$. Such operators, which we call complex linear, satisfy

$$A(fz) = (Af)z \qquad (3.2)$$

for $z \in \mathbb{C}(1,e_7)$.

As we have seen in Eq. (2.25), every $f \in \mathscr{H}_+$ has the representation

$$f = \sum_{ij} f_{ij}\rho_{ij}$$

and the most general action of an operator linear over the reals (see Appendix A) is (for A real-valued)

$$Af = \sum_{ij,kl} \mathfrak{A}_{ij,kl} f_{kl}\rho_{ij} \qquad (3.3)$$

Reorganizing the sum on the right-hand side according to the procedure used to obtain Eq. (2.39) from (2.37), we obtain (for $A^{kl}_{\alpha\beta}$ complex-valued)

$$\begin{aligned}Af = \sum_{kl} e_k P_0 \Big\{ & A^{kl}_{00}\psi^l_0 + A^{kl'}_{00}\psi^{l*}_0 + \sum_{\alpha=1}^{3} (A^{kl'}_{0\alpha}\psi^l_\alpha\\ &+ A^{kl}_{0\alpha}\psi^{l*}_\alpha) + \sum_{\alpha=1}^{3}\Big[A^{kl'}_{\alpha 0}\psi^l_0 + A^{kl}_{\alpha 0}\psi^{l*}_0\\ &+ \sum_{\beta=1}^{3} (A^{kl}_{\alpha\beta}\psi^l_\beta + A^{kl'}_{\alpha\beta}\psi^{l*}_\beta)\Big] e_\alpha \Big\},\end{aligned} \qquad (3.4)$$

where $A^{kl}_{\alpha\beta}, \cdots$ correspond to independent linear combinations of the (real) operators $\mathfrak{A}_{ij,kl}$ *over* $\mathbb{C}(1,e_7)$, and the ψ^l_0, ψ^l_α are the complex valued functions representing f as given in Eq. (2.38). With the help of Eq. (2.40), for $z \in \mathbb{C}(1,e_7)$, we obtain

$$\begin{aligned}(Af)z = \sum_{kl} e_k P_0 \Big\{ & A^{kl}_{00}\psi^l_0 z + A^{kl'}_{00}\psi^{l*}_0 z + \sum_{\alpha=1}^{3} (A^{kl'}_{0\alpha}\psi^l_\alpha z\\ &+ A^{kl}_{0\alpha}\psi^{l*}_\alpha z) + \sum_{\alpha}\Big[A^{kl'}_{\alpha 0}\psi^l_0 z^* + A^{kl}_{\alpha 0}\psi^{l*}_0 z^*\\ &+ \sum_{\beta=1}^{3} (A^{kl}_{\alpha\beta}\psi^l_\beta z^* + A^{kl'}_{\alpha\beta}\psi^{l*}_\beta z^*)\Big] e_\alpha \Big\},\end{aligned} \qquad (3.5)$$

and

$$\begin{aligned}(Af)z = \sum_{kl} e_k P_0 \Big\{ & A^{kl}_{00}\psi^l_0 z + A^{kl'}_{00}\psi^{l*}_0 z^* + \sum_{\alpha=1}^{3} (A^{kl'}_{0\alpha}\psi^l_\alpha z^*\\ &+ A^{kl}_{0\alpha}\psi^{l*}_\alpha z) + \sum_{\alpha=1}^{3}\Big[A^{kl'}_{\alpha 0}\psi^l_0 z + A^{kl}_{\alpha 0}\psi^{l*}_0 z^*\\ &+ \sum_{\beta=1}^{3} (A^{kl}_{\alpha\beta}\psi^l_\beta z^* + A^{kl'}_{\alpha\beta}\psi^{l*}_\beta z)\Big] e_\alpha \Big\},\end{aligned} \qquad (3.6)$$

Comparing Eqs. (3.5) and (3.6), the requirement (3.2) for complex linearity is satisfied for the operator A if and only if

$$\begin{aligned}A^{kl'}_{00} &= 0, \quad A^{kl'}_{0\alpha} = 0,\\ A^{kl'}_{\alpha 0} &= 0, \quad A^{kl'}_{\alpha\beta} = 0.\end{aligned} \qquad (3.7)$$

The vanishing of the transitions listed in Eq. (3.7) corresponds to superselection rules imposed on all operators linear over $\mathbb{C}(1,e_7)$ and, in particular, on self-adjoint operators of this type. We see, however, that there remain components of operators linear over $\mathbb{C}(1,e_7)$ on the entire space that appear to act *antilinearly*, in the manner of Wigner's corepresentations,[20] on the subspaces $\mathscr{H}_0$, $\mathscr{H}_\alpha$ over which we have decomposed $\mathscr{H}_+$ by means of the representation (2.39) (the antilinear structure of the theory is reformulated in a simple formal way in Appendix B). The general complex linear operator therefore has the action

$$Af = \sum_{kl} e_k P_0 \Big[A^{kl}_{00}\psi^l_0 + \sum_{\alpha=1}^{3} A^{kl}_{00}\psi^{l*}_\alpha + \sum_{\alpha=1}^{3} \Big(A^{kl}_{\alpha 0}\psi^{l*}_0 + \sum_{\beta=1}^{3} A^{kl}_{\alpha\beta}\psi^l_\beta\Big) e_\alpha \Big]. \qquad (3.8)$$

Following the interpretation of Gürsey and Günaydin[16] for the octonion Hilbert space, we take ψ^l_α to correspond to states in the "quark" or "unobservable" sector and ψ^l_0 to states in the "leptonic" or "observable" sector. Equation (3.8) indicates that a complex linear operator can act antilinearly on the unobservable space and admit a transition from that (conjugate) state to the observable space, and conversely. From the point of view of field theory, the operators $A^{kl}_{0\alpha}$ and $A^{kl}_{\alpha 0}$ behave like currents composed of lepton and quark fields.

Saclioglu[18] has pointed out that Nambu's suggestion,[27] based on the invariance of a massless fermion Lagrangian to Pauli–Gürsey transformations, of the existence of diquark currents (which then implies the existence of leptoquark currents), when applied to the generators of $SU(3)_{\text{color}} \times SU(n)_{\text{flavor}}$ charges of the usual type, results in an algebraic structure which is most economically accommodated in the framework of the exceptional groups. Equation (3.8) shows that the leptoquark transition operators occur naturally as pieces of operators linear over the complex subalgebra $\mathbb{C}(1,e_7)$. The self-adjoint operators of this class would be observables if complex linearity alone were sufficient to insure that no superselection rules are violated.[28] It appears, in any case, that operators of this type must be accommodated in theories which attempt to unify weak, electromagnetic and strong interactions.[18,20,21]

Furthermore, the linear operators corresponding to the U(4) symmetry of the scalar product, as noted in the remarks following Eq. (2.36), could not be realized in the absence of these terms [the maximum symmetry transformations available would otherwise be $U(1)\times U(3)$]. In Sec. V, we shall show that the generalized phase algebra $\mathfrak{A}\subset C_7$ which commutes with e_7 acts on the complex wavefunctions exactly as indicated in Eq. (3.8).

We conclude this section with some technical remarks on the properties of complex linear operators. From the definition of the scalar product (2.45), we define $A^\dagger$, the adjoint of A, according to

$$(f,Ag)_c=(A^\dagger f,g)_c \tag{3.9}$$

in the same way as for the usual theory of complex Hilbert spaces. It follows that (these operators are complex-valued)

$$(A^\dagger)^{kl}_{0\alpha}=A^{lk\dagger}_{00},\quad (A^\dagger)^{kl}_{\alpha 0}=A^{lk\dagger *}_{0\alpha},$$
$$(A^\dagger)^{kl}_{0\alpha}=A^{lk\dagger *}_{\alpha 0}\quad (A^\dagger)^{kl}_{\alpha\beta}=A^{lk\dagger}_{\beta\alpha}, \tag{3.10}$$

where the adjoints defined on the right of these relations are determined by complex scalar products among vectors from the subspaces $\mathscr{H}_0$, $\mathscr{H}_\alpha$, and the domain and range subspaces are indicated by the indices on the left-hand side. Symmetric operators therefore satisfy

$$\mathrm{A}^{kl}_{00}=A^{lk\dagger}_{00},\quad A^{kl}_{\alpha 0}=A^{lk\dagger *}_{0\alpha},$$
$$A^{kl}_{0\alpha}=A^{lk\dagger *}_{\alpha 0}\quad A^{kl}_{\alpha\beta}=A^{lk\dagger}_{\beta\alpha}. \tag{3.11}$$

We record here also the form of the products of two operators linear over $\mathbb{C}(1,e_7)$:

$$(BA)^{kl}_{00}=\sum_m\Big(B^{km}_{00}A^{ml}_{00}+\sum_\alpha B^{km}_{0\alpha}A^{ml*}_{\alpha 0}\Big),$$
$$(BA)^{kl}_{0\alpha}=\sum_m\Big(B^{km}_{00}A^{ml}_{0\alpha}+\sum_\beta B^{km}_{0\beta}A^{ml*}_{\beta\alpha}\Big),$$
$$(BA)^{kl}_{\alpha 0}=\sum_m\Big(B^{km}_{\alpha 0}A^{ml*}_{00}+\sum_\beta B^{km}_{\alpha\beta}A^{ml}_{\beta 0}\Big),$$
$$(BA)^{kl}_{\alpha\beta}=\sum_m\Big(B^{km}_{\alpha 0}A^{ml*}_{0\beta}+\sum_\gamma B^{km}_{\alpha\gamma}A^{ml}_{\gamma\beta}\Big), \tag{3.12}$$

IV. STATES AND THE ALGEBRA OF PHASES

As we have shown previously,[13] the algebraically closed linear manifolds of $\mathscr{H}_*$ form a complete, weakly modular, orthocomplemented, atomic lattice, and this structure can therefore be embedded in a Hilbert space over a field Φ.[15] In this earlier work, Φ was chosen to be the basic field over which the algebra was defined (the reals in the case of the real C_7 we are presently considering). Admitting as observables only operators linear over the entire algebra, we found that the embedding results in a Hilbert space over Φ with superselection rules that correspond to the primitive idempotents of the algebra. It was furthermore found that the pure states correspond to vectors multiplied (in the sense of a generalized phase) by the corresponding minimal right ideals over the full algebra (elements of unit modulus).

In the present work, we shall admit as the smallest linear manifold a vector multiplied by quantities from $\mathbb{C}(1,e_7)$; pure states therefore correspond to rays over $\mathbb{C}(1,e_7)$. The further resolution of these manifolds would imply the existence of observables which are sensitive to the global complex phase of a wavefunction. We shall therefore embed the Hilbert space over C_7 into the Hilbert space $\mathscr{H}_c$ with scalar product $(f,g)_c$, and linear manifolds closed over $\mathbb{C}(1,e_7)$.

We may then define "states" in terms of measures on the complex linear closed manifolds. There is, however, a larger algebra which may be applied to the vectors of this space, and we must therefore extend the notion of a state to measures defined on manifolds closed under an algebra $\mathfrak{A}\subset C_7$, larger than $\mathbb{C}(1,e_7)$ (as we shall show, $\mathfrak{A}$ cannot be the full algebra C_7 since there are elements of C_7 which can alter the structure of the minimal invariant subspaces). The pure states, evaluated on linear manifolds closed under this larger algebra, will be invariant under the action of the norm-preserving elements of $\mathfrak{A}$. It is precisely such an invariance which is characteristic of the fundamental idea of non-Abelian gauge groups,[1] and we therefore identify $\mathfrak{A}$ with the algebraic structure of a gauge group. It is then consistent to define an observable as a self-adjoint operator (with respect to a scalar product linear over $\mathfrak{A}$) which is linear over $\mathfrak{A}$; the expectation value of any such observable is independent of the choice of norm-preserving multiples from $\mathfrak{A}$ on the vectors involved.

The conditions defining the algebra $\mathfrak{A}\subset C_7$ follow from the requirement that the notion of a state can be extended from a function on the projection operators corresponding to linear manifolds closed over $\mathbb{C}(1,e_7)$ to a function on the projection operators corresponding to linear manifolds closed over $\mathfrak{A}$. We shall follow the procedure established in Ref. 13 to make this requirement precise.

To achieve an embedding of $\mathscr{H}_*$ into $\mathscr{H}_c$, we seek orthogonal projections in the algebra which will generate subspaces that are closed under $\mathbb{C}(1,e_7)$, the primitive idempotents for minimal complex manifolds. These idempotents will be invariant under the action of operators with sufficient linearity (linear over $\mathfrak{A}$) and will therefore generate superselection subspaces. Starting with the primitive idempotents for the real field [defined by Eq. (2.14)], we can find the complex primitive idempotents as follows. For $f\in\mathscr{H}_*$,

$$fP_ke_7=fe_7\cdot(e_7P_ke_7^*)$$
$$=fe_7\begin{cases}P_{k+3}, & k=1,2,3,\\ P_{k-3}, & k=4,5,6,\\ P_0, & k=7,\\ P_7, & k=0.\end{cases} \tag{4.1}$$

Hence

$$f(P_k+P_{k+3})e_7=fe_7(P_k+P_{k+3}),\quad k=1,2,3,$$
$$f(P_0+P_7)e_7=fe_7(P_0+P_7\ \ ), \tag{4.2}$$

and therefore the smallest linear manifolds invariant under $\mathbb{C}(1,e_7)$ are the subspaces corresponding to

$$\hat{P}_\alpha = e_\alpha \hat{P}_0 e_\alpha^*, \quad \alpha=0,1,2,3, \tag{4.3}$$

where

$$\begin{aligned}\hat{P}_0 &= P_0 + P_7 \\ &= \tfrac{1}{4}(I - e_4e_7e_1)(I - e_5e_7e_2).\end{aligned} \tag{4.4}$$

The last equality may be verified by multiplying out the factors of Eqs. (4.4) and (2.13) (conjugating the latter with e_7 to obtain P_7) and using Eq. (2.11) for the subspace $\mathscr{H}_+ = \mathscr{H}P_+$. The structure of $\hat{P}_0$ is due to the fact that there are just two algebraically independent, symmetric, commuting elements of C_7, $e_4e_7e_1$, and $e_5e_7e_2$, that commute with e_7 (in $\mathscr{H}_+$). The spectral representations are found by noting that fP_k, $k = 0,1,...,7$, are eigenfunctions of these operators; since they commute with e_7, the (0,7), $(k,k+3)$, $k = 1,2,3$ subspaces are degenerate, and hence the $\hat{P}_\alpha$ are the spectral family. Finally, $\hat{P}_0e_4e_7e_1 = -\hat{P}_0$, for example; conjugating with the e_α, we obtain similar relations for all the $\hat{P}_\alpha$. Since

$$\sum_{\alpha=0}^{3} \hat{P}_\alpha = I, \tag{4.5}$$

we find

$$e_4e_7e_1 = -\hat{P}_0 - \hat{P}_1 + \hat{P}_2 + \hat{P}_3 \tag{4.6}$$

and similarly

$$\begin{aligned} e_5e_7e_2 &= -\hat{P}_0 + \hat{P}_1 - \hat{P}_2 + \hat{P}_3, \\ e_6e_7e_3 &= -\hat{P}_0 + \hat{P}_1 + \hat{P}_2 - \hat{P}_3. \end{aligned} \tag{4.7}$$

The operator $e_6e_7e_3$ is not independent of the other two, but we record its spectral representation here to make explicit the spectral content of all of the operators associated with multiplication rules [Eq. (2.20)] involving e_7. In particular, note that the three multiplication rules involving e_7 in Eq. (2.20) are valid in the subspace corresponding to $\hat{P}_0$.

Gleason[29] has shown that for an irreducible system of propositions realized by the projections P_M of a separable Hilbert space, real or complex but of dimension $\geqslant 3$, there exists a density matrix ρ for every continuous state[30] such that

$$\omega(M) = \mathrm{Tr}(\rho P_M), \tag{4.8}$$

where by Tr we mean a trace over the full Hilbert space as well as over the algebraic indices. We have recognized that the linear manifolds, closed over $\mathbb{C}(1,e_7)$, of $\mathscr{H}_+$ do not form an irreducible system of propositions, since $\mathscr{H}_+$ decomposes into a direct sum over the superselection subspaces $\mathscr{H}_\alpha$ corresponding to the minimal subspaces invariant over $\mathbb{C}(1,e_7)$ (if one were willing to accept the further resolution of these linear manifolds to the real field, the minimal subspaces would be $\mathscr{H}_k = \mathscr{H}_+P_k$, $k=0,1,...,7$).[13] In each of the minimal subspaces $\mathscr{H}_\alpha$, we may express the state (4.8) as a function of the linear manifolds [over $\mathbb{C}(1,e_7)$] in that subspace, as

$$\omega_\alpha(M_c^\alpha) = \sum_i \gamma_i^\alpha (f_i^\alpha, P_{M_c^\alpha} f_i^\alpha)_c, \tag{4.9}$$

where $\gamma_i^\alpha > 0$, $\Sigma_i \gamma_i^\alpha = 1$, and $f_i^\alpha \in \mathscr{H}_\alpha$. We are now in a position to state and prove the principal result of this section.

Statement 1: *There exists a maximal (star) subalgebra $\mathfrak{A} \subset C_7$ such that to every M_c^α there corresponds a linear manifold M closed over $\mathfrak{A}$, which is a proper extension of M_c^α, for which $P_M f^\alpha = P_{M_c^\alpha} f^\alpha$, $f^\alpha \in \mathscr{H}_\alpha$, and contained in every M there is a complex M_c^α of which M is a proper extension of this type.* There are eight solutions for the algebra $\mathfrak{A}$.

As a consequence of this result, we may assert the following:

Statement 2: *States $\omega(M)$ can be defined on the set of linear manifolds M closed over $\mathfrak{A}$; the pure states correspond to vectors multiplied by the right ideals of $\mathfrak{A}$ generated by the $\hat{P}_\alpha$.*

In the following paragraphs, we shall prove Statement 1 and define the structure of the algebra $\mathfrak{A}$. We shall then give a proof of Statement 2.

Let M_c^α be spanned by $g_1\hat{P}_\alpha, g_2\hat{P}_\alpha, \cdots$ over $\mathbb{C}(1,e_7)$, and let M be the closed linear extension of M_c^α over $\mathfrak{A}$, i.e., M is spanned over $\mathfrak{A}$ by the same elements. For each f_i for which the f_i^α of Eq. (4.9) is given by $f_i^\alpha = f_i\hat{P}_\alpha$, there is a unique decomposition into a part in M and a part orthogonal to M in the sense[12]

$$f_i = g_M + h,$$

$$g_M = g_1\hat{P}_\alpha \mathbf{a}_1 + g_2\hat{P}_\alpha \mathbf{a}_2 + \cdots, \tag{4.10}$$

and

$$\mathrm{tr}((h, g_j\hat{P}_\alpha)\mathbf{a}) = 0 \quad \text{all } \mathbf{a} \in \mathfrak{A} \tag{4.11}$$

Now,

$$\begin{aligned} f_i^\alpha &= f_i\hat{P}_\alpha \\ &= g_M\hat{P}_\alpha + h\hat{P}_\alpha \\ &= g_1\hat{P}_\alpha \mathbf{a}_1\hat{P}_\alpha + g_2\hat{P}_\alpha \mathbf{a}_2\hat{P}_\alpha + \cdots + h\hat{P}_\alpha \end{aligned} \tag{4.12}$$

is a unique decomposition of f_i^α into a part in M_c^α and a part complex orthogonal to M_c^α provided that

$$\hat{P}_\alpha \mathbf{a}\hat{P}_\alpha = z\hat{P}_\alpha, \quad z \in \mathbb{C}(1,e_7), \quad \text{all } \mathbf{a} \in \mathfrak{A}, \tag{4.13}$$

and

$$(h\hat{P}_\alpha, g_j\hat{P}_\alpha)_c = 0. \tag{4.14}$$

The last condition requires

$$\begin{aligned} &\mathrm{tr}(h\hat{P}_\alpha, g_j\hat{P}_\alpha) = \mathrm{tr}(h, g_j\hat{P}_\alpha) = 0, \\ &\mathrm{tr}((h\hat{P}_\alpha, g_j\hat{P}_\alpha)e_7) = \mathrm{tr}((h, g_j\hat{P}_\alpha)e_7) = 0; \end{aligned} \tag{4.15}$$

both are satisfied as a consequence of Eq. (4.11), since $\mathfrak{A}$ must include $\mathbb{C}(1,e_7)$ if M is a proper extension of M_c^α. It then follows from Eqs. (4.13) and (4.15) that

$$\begin{aligned} \mathrm{tr}((g_M\hat{P}_\alpha, h\hat{P}_\alpha)\mathbf{a}) &= \mathrm{tr}((g_M, h)\hat{P}_\alpha \mathbf{a}\hat{P}_\alpha) \\ &= \mathrm{tr}((g_M, h\hat{P}_\alpha)z) = 0 \end{aligned}$$

and hence

$$P_M(f_i\hat{P}_\alpha)=P_M f_i^\alpha = g_M\hat{P}_\alpha=(P_M f_i)\hat{P}_\alpha, \tag{4.16}$$

i.e., P_M is linear over the $\hat{P}_\alpha$. In fact, $\hat{P}_\alpha$ must be contained in $\mathfrak{A}$ if operators linear over $\mathfrak{A}$ are to leave the superselection subspaces generated by the $\hat{P}_\alpha$ invariant. Furthermore, from Eqs. (4.12), (4.13), and (4.16), we see that for every M_c^α there exists an M closed over $\mathfrak{A}$ such that

$$P_{M_c^\alpha} f_i^\alpha = P_M f_i^\alpha .$$

For the second part of Statement 1, let us take for M the closed linear manifold spanned by $g_1, g_2, \cdots$ over $\mathfrak{A}$ and for M_c^α the elements of M multiplied on the right by $\hat{P}_\alpha$. Since $\mathbb{C}(1,e_7)\subset\mathfrak{A}$, and the complex numbers commute with the $\hat{P}_\alpha$, M_c^α is clearly closed over $\mathbb{C}(1,e_7)$. We must now demand that the extension of M_c^α by right multiplication with elements of $\mathfrak{A}$ fully reconstruct the manifold M, i.e., that (linear combinations of)

$$\{ g_1\mathbf{a}_1\hat{P}_\alpha\mathbf{a}_1' + g_2\mathbf{a}_2\hat{P}_\alpha\mathbf{a}_2' + \cdots\}$$

over all $\mathbf{a}_j$, $\mathbf{a}_j' \in \mathfrak{A}$, is equivalent to

$$\{ g_1\mathbf{a}_1'' + g_2\mathbf{a}_2'' + \cdots\}$$

over all $\mathbf{a}_j'' \in \mathfrak{A}$. We therefore obtain the following condition: *For every* $\mathbf{a}$, $\mathbf{a}' \in \mathfrak{A}$, and for any $\alpha = 0,1,2,3$,

$$\mathbf{a}\hat{P}_\alpha\mathbf{a}' = \mathbf{a}'' \tag{4.17}$$

belongs to $\mathfrak{A}$, *and for each* α, *all* $\mathfrak{A}$ *is spanned by sums of elements of this type.*

The first part is equivalent to imposing that $\mathfrak{A}$ be an algebra, since we have already admitted $\hat{P}_\alpha \in \mathfrak{A}$ (and $\Sigma_\alpha \hat{P}_\alpha = I$). The form (4.17) is a convenient starting point for our demonstration, however. With Eq. (4.13), it serves to completely characterize the algebra $\mathfrak{A}$.

The most general structure of the $\hat{P}_\alpha, \hat{P}_\beta$ part of an element of C_7 is [we write $a_{ij}\in\mathbb{R}$ for the $K_{ij}(a)$ of Eq. (2.23)]

$$\hat{P}_\alpha a\hat{P}_\beta = a_{\alpha\beta}\rho_{\alpha\beta} + a_{\alpha+3,\beta}\rho_{\alpha+3,\beta} + a_{\alpha,\beta+3}\rho_{\alpha,\beta+3} + a_{\alpha+3,\beta+3}\rho_{\alpha+3,\beta+3}, \tag{4.18}$$

where $\alpha = 0,1,2,3$ and the index $\alpha+3$ is to be read as 7 for $\alpha = 0$. With the help of the multiplication rules (2.20), valid to the left or right of P_0, Eq. (4.18) can be rewritten as

$$\hat{P}_\alpha a\hat{P}_\beta = \rho_{\alpha\beta}(a_{\alpha\beta} - e_7 a_{\alpha,\beta+3}) + \rho_{\alpha+3,\beta+3}(a_{\alpha+3,\beta+3} + e_7 a_{\alpha+3,\beta}). \tag{4.19}$$

In particular, the diagonal components of Eq. (4.19) are

$$\hat{P}_\alpha a\hat{P}_\alpha = P_\alpha(a_{\alpha\alpha} - e_7 a_{\alpha,\alpha+3}) + P_{\alpha+3,\alpha+3} \times (a_{\alpha+3,\alpha+3} + e_7 a_{\alpha+3,\alpha}). \tag{4.20}$$

The general element of C_7, restricted to the $\mathscr{H}_\alpha$ subspace, is therefore parametrized by two complex numbers, and does not satisfy the requirement (4.13). Hence $\mathfrak{A}$ cannot be all of C_7. For example, one sees that $\hat{P}_0 a\hat{P}_0$ is not simply proportional to $\hat{P}_0 = P_0 + P_7$, which generates one of the minimal subspaces invariant under $\mathbb{C}(1,e_7)$, but has the form $P_0 z_1 + P_7 z_2$.

The restriction implied by Eq. (4.13) requires that for $a\in\mathfrak{A}$,

$$a_{\alpha\alpha} = a_{\alpha+3,\alpha+3}, \quad a_{\alpha,\alpha+3} = -a_{\alpha+3,\alpha}. \tag{4.21}$$

Since $\mathbf{a}''$ in Eq. (4.17) must also satisfy Eqs. (4.21), we make use of Eq. (4.18) to write

$$\hat{P}_\beta\mathbf{a}\hat{P}_\alpha\mathbf{a}'\hat{P}_\beta = \hat{P}_\beta\mathbf{a}''\hat{P}_\beta \tag{4.22}$$

explicitly; imposing the constraint (4.21), we obtain

$$a_{\beta\alpha}a'_{\alpha\beta} + a_{\beta\alpha+3}a'_{\alpha+3,\beta} = a_{\beta+3,\alpha}a'_{\alpha,\beta+3} + a_{\beta+3,\alpha+3}a'_{\alpha+3,\beta+3} \tag{4.23}$$

and

$$a_{\beta+3,\alpha}a'_{\alpha\beta} + a_{\beta+3,\alpha+3}a'_{\alpha+3,\beta} = -(a_{\beta\alpha}a'_{\alpha\beta+3} + a_{\beta\alpha+3}a'_{\alpha+3,\beta+3}). \tag{4.24}$$

Equations (4.21), (4.23), and (4.24), with the requirement that it be a (star) algebra, contain all of the information necessary to define $\mathfrak{A}$.

In particular, the subalgebra of C_7 for which $\{a_{\alpha,\beta+3}\} = \{a_{\alpha+3,\beta}\} = 0$ satisfies Eqs. (4.21), (4.23), and (4.24) if

$$\xi_{\alpha\beta} = \frac{a_{\beta\alpha}}{a_{\beta+3,\alpha+3}} = \frac{a'_{\alpha+3,\beta+3}}{a'_{\alpha\beta}} \tag{4.25}$$

is a universal constant (with $\xi_{\alpha\alpha} = +1$). Since the right-hand side is then equal to $\xi_{\alpha\beta}^{-1}$, it follows that

$$\xi_{\beta\alpha}\xi_{\alpha\beta} = 1. \tag{4.26}$$

Furthermore, this subalgebra is also a subalgebra of $\mathfrak{A}$ when

$$\xi_{\beta\gamma}\xi_{\gamma\alpha} = \xi_{\beta\alpha}, \tag{4.27}$$

since for such a, a'

$$aa' = \sum_{\alpha\beta\gamma} a_{\beta+3,\gamma+3}a'_{\gamma+3,\alpha+3}(\xi_{\beta\gamma}\xi_{\gamma\alpha}\rho_{\beta\alpha} + \rho_{\beta+3,\alpha+3}).$$

Returning to the general expressions (4.23) and (4.24), let us choose a' to belong to the subalgebra of $\mathfrak{A}$ for which $\{a_{\alpha,\beta+3}\} = \{a_{\alpha+3,\beta}\} = 0$. Then, for general $a\in\mathfrak{A}$,

$$a_{\beta\alpha}a'_{\alpha\beta} = a_{\beta+3,\alpha+3}a'_{\alpha+3,\beta+3},$$

$$a_{\beta+3,\alpha}a'_{\alpha\beta} = -a_{\beta,\alpha+3}a'_{\alpha+3,\beta+3}. \tag{4.28}$$

Hence, Eq. (4.25) is valid in the general case of $a\in\mathfrak{A}$; from Eqs. (4.28) it also follows that

$$a_{\beta,\alpha+3} = -\xi_{\alpha\beta}a_{\beta+3,\alpha}. \tag{4.29}$$

Let us now consider Eqs. (4.23) and (4.24) for both a, a' general elements of $\mathfrak{A}$ satisfying Eqs. (4.25) and (4.29). Then [with Eq. (4.26)] it follows from Eqs. (4.23) and (4.29) that

$$\xi_{\alpha\beta}=\xi_{\beta\alpha} \tag{4.30}$$

and from Eqs. (4.24) and (4.26) that

$$\xi^2_{\alpha\beta}=1, \quad \xi_{\alpha\beta}=\pm 1. \tag{4.31}$$

The conditions (4.25) and (4.29), with (4.27), (4.30), and (4.31) define a closed algebra. For a,a' satisfying these conditions

$$aa'=\sum_{\alpha\beta\gamma}[\rho_{\alpha\beta}(a_{\alpha\gamma}a'_{\gamma\beta}+a_{\alpha\gamma+3}a'_{\gamma+3\beta})+\rho_{\alpha\beta+3}(a_{\alpha\gamma}a'_{\gamma\beta+3}+a_{\alpha\gamma+3}a'_{\gamma+3,\beta+3})$$

$$+\rho_{\alpha+3,\beta}(a_{\alpha+3,\gamma}a'_{\gamma\beta}+a_{\alpha+3,\gamma+3}a'_{\gamma+3,\beta})+\rho_{\alpha+3,\beta+3}(a_{\alpha+3,\gamma}a'_{\gamma\beta+3}+a_{\alpha+3,\gamma+3}a'_{\gamma+3,\beta+3})]$$

$$=\sum_{\alpha\beta\gamma}[\rho_{\alpha\beta}(\xi_{\alpha\beta}a_{\alpha+3,\gamma+3}a'_{\gamma+3,\beta+3}-\xi_{\gamma\alpha}a_{\alpha+3,\gamma}a'_{\gamma+3,\beta})-\rho_{\alpha,\beta+3}(\xi_{\alpha\beta}a_{\alpha+3,\gamma+3}a'_{\gamma+3,\beta}+\xi_{\gamma\alpha}a_{\alpha+3,\gamma}a'_{\gamma+3,\beta+3})$$

$$+\rho_{\alpha+3,\beta}(\xi_{\gamma\beta}a_{\alpha+3,\gamma}a'_{\gamma+3,\beta+3}+a_{\alpha+3,\gamma+3}a'_{\gamma+3,\beta})+\rho_{\alpha+3,\beta+3}(-\xi_{\beta\gamma}a_{\alpha+3,\gamma}a'_{\gamma+3,\beta}+a_{\alpha+3,\gamma+3}a'_{\gamma+3,\beta+3})]. \tag{4.32}$$

The coefficient of $\rho_{\alpha+3,\beta+3}$, when multiplied by $\xi_{\alpha\beta}$, becomes equal to the coefficient of $\rho_{\alpha\beta}$; similarly, the coefficient of $\rho_{\alpha+3,\beta}$, when multiplied by $\xi_{\alpha\beta}$, becomes equal to the negative of the coefficient of $\rho_{\alpha,\beta+3}$.

The terms of Eq. (4.32), with index γ fixed, correspond to $a\hat{P}_\gamma a'$. It is clear that sums of such terms, e.g., $\Sigma a_j\hat{P}_\gamma a'_j$ span the entire algebra $\mathfrak{A}$, since the sum over γ in Eq. (4.32), with suitable choice of a_j,a'_j, can be effectively replaced by the sum over j [e.g., define $(a_j)_{\alpha\gamma}=a_{\alpha\gamma_j}$ and similarly for other components].

Separating out explicitly the index 0 from 1, 2, 3, Eq. (4.27) may be written

$$\xi_{0\alpha}\xi_{0\beta}=\xi_{\alpha\beta},$$
$$\xi_{\alpha\gamma}\xi_{\gamma\beta}=\xi_{\alpha\beta}. \tag{4.33}$$

There are eight solutions, comprising four inequivalent types. With Eqs. (4.21) ($\xi_{\alpha\alpha}=+1,\xi_{00}=+1$), representatives of these types are

$\mathfrak{A}_{+++}$:	$\xi_{0\alpha}=+1,$	$\xi_{\alpha\beta}=+1$			(type 0),
$\mathfrak{A}_{++-}$:	$\xi_{01}=\xi_{02}=+1,$	$\xi_{03}=-1,$	$\xi_{12}=+1,$	$\xi_{13}=\xi_{23}=1$	(type 1),
$\mathfrak{A}_{--+}$:	$\xi_{01}=\xi_{02}=-1,$	$\xi_{03}=+1,$	$\xi_{12}=+1,$	$\xi_{13}=\xi_{23}=-1$	(type 2),
$\mathfrak{A}_{---}$:	$\xi_{01}=\xi_{02}=\xi_{03}=-1,$	$\xi_{\alpha\beta}=+1$			(type 3).

(4.34)

There are three solutions of type 1, with one of the $\xi_{0\alpha}$ set equal to -1 and three solutions of type 2, with one of the $\xi_{0\alpha}$ set equal to $+1$. There is only one solution of type 0, and one of type 3, as given in Eqs. (4.34).

The conditions (4.25) and (4.29) imply that the $\hat{P}_\alpha,\hat{P}_\beta$ part of every element of $\mathfrak{A}$ [in the sequel we shall refer to properties of all of the algebras specified by Eqs. (4.34) as those of the algebra $\mathfrak{A}$, unless otherwise stated] is parametrized by a single complex number. From Eq. (4.19) one sees that

$$\hat{P}_\alpha \mathbf{a}\hat{P}_\beta=(\rho_{\alpha\beta}\xi_{\alpha\beta}+\rho_{\alpha+3,\beta+3})(a_{\alpha+3,\beta+3}+e_7a_{\alpha+3,\beta}). \tag{4.35}$$

This result is a proper extension of Eq. (4.13) to the nondiagonal parts. We shall discuss the structure of the algebras $\mathfrak{A}_\xi$, where ξ stands for the sign distribution of $\xi_{0\alpha}$, $\alpha=1,2,3$, in detail in Sec. V.

We remark that the solution of type 0 corresponds to the largest subalgebra of C_7 that commutes with e_7.

Since

$$e_7=\sum_{\alpha=0}^{3}\hat{P}_\alpha e_7=\sum_{\alpha=0}^{3}(\rho_{\alpha+3,\alpha}-\rho_{\alpha,\alpha+3}), \tag{4.36}$$

we see that for $a\in C_7$

$$ae_7=\sum_{i,\alpha=0}^{3}(a_{i,\alpha+3}\rho_{i\alpha}-a_{i\alpha}\rho_{i,\alpha+3}),$$

$$e_7a=\sum_{i,\alpha=0}^{3}(a_{\alpha i}\rho_{\alpha+3,i}-a_{\alpha+3,i}\rho_{\alpha i}), \tag{4.37}$$

Comparing corresponding terms, one finds that for a and e_7 to commute, the conditions (4.25) and (4.29) must be valid with $\xi_{0\alpha}=\xi_{\alpha\beta}=+1$. Therefore, there are elements of each $\mathfrak{A}_\xi$ (type 1,2,3) which *anticommute* with e_7. In the next section, we shall show that there are other forms for the imaginary unit (also associated with the direction 7), each of which commute with one of the algebras $\mathfrak{A}_\xi$, and provide an interpretation.

We now turn to a proof of Statement 2, valid for any of the solutions given in Eqs. (4.34).

Let M correspond to M^α_c, in the sense given in Statement 1. Then

$$\omega_\alpha(M)=\sum_i\gamma^\alpha_i(f^\alpha_i,P_Mf^\alpha_i)_c. \tag{4.38}$$

For each M there exists an M_c^α in $\mathscr{H}_\alpha$ for which this correspondence is valid, and therefore $\omega_\alpha(M)$ is defined over all M closed over $\mathfrak{A}$.[31] Since $\omega_\alpha(M)$ is a state,

$$\omega(M)=\sum \lambda_\alpha \omega_\alpha(M)=\sum_\alpha \lambda_\alpha \gamma_i^\alpha (f_i^\alpha, P_M f_i^\alpha)_c \quad (4.39)$$

is also a state if $\lambda_\alpha \geqslant 0$, $\Sigma_\alpha \lambda_\alpha = 1$. From the definition (2.43) of the complex scalar product and the fact that the $\hat{P}_\alpha$ commute with e_7, $(f_i^\beta, P_M f_i^\alpha)_c = 0$ for $\beta \neq \alpha$. Hence

$$\omega(M)=\sum_{\alpha\beta} (\lambda_\alpha \gamma_i^\alpha)^{1/2} (\lambda_\beta \gamma_i^\beta)^{1/2} (f_i^\alpha, P_M f_i^\beta)_c$$

$$=\sum_i \mu_i (f_i, P_M f_i)_c, \quad (4.40)$$

where $\mu_i = \Sigma_\alpha \lambda_\alpha \gamma_i^\alpha > 0$, $\Sigma \mu_i = 1$, is a state of the form given in Eq. (4.8).

To construct a pure state, we utilize the converse of the Gleason theorem in the form given by Mackey.[32] Let M_0 be the closed linear manifold defined by $\{f_0 \hat{P}_\alpha \mathbf{a}\}$, for $\mathbf{a} \in \mathfrak{A}$, and suppose that $\omega_0(M_0) = 1$. Then, since $\|f_i\|^2 = 1$, it follows that

$$0=\sum \mu_i \|(I - P_{M_0}) f_i\|^2$$

or

$$P_{M_0} f_i = f_i = f_0 \hat{P}_\alpha \mathbf{a}_i. \quad (4.41)$$

Substituting this result in Eq. (4.40), we find

$$\omega_0(M)=\sum \mu_i (f_0 \hat{P}_\alpha \mathbf{a}_i, P_M f_0 \hat{P}_\alpha \mathbf{a}_i)_c$$

$$=\sum \mu_i (\operatorname{tr}(f_0, P_M f_0)\ \hat{P}_\alpha \mathbf{a}_i \mathbf{a}_i^* \hat{P}_\alpha). \quad (4.42)$$

Now,

$$\|f_i\|^2 = \operatorname{tr}(f_0 \hat{P}_\alpha \mathbf{a}_i, f_0 \hat{P}_\alpha \mathbf{a}_i)$$

$$=\operatorname{tr}((f_0, f_0)\ \hat{P}_\alpha \mathbf{a}_i \mathbf{a}_i^* \hat{P}_\alpha). \quad (4.43)$$

According to Eq. (4.13) ($\mathbf{aa}^*$ is symmetric)

$$\hat{P}_\alpha \mathbf{a}_i \mathbf{a}_i^* \hat{P}_\alpha = \eta_i^\alpha \hat{P}_\alpha, \quad (4.44)$$

η_i^α is real and nonnegative. Hence ($\alpha = 0,1,2$, or 3)

$$\|f_i\|^2 = \eta_i^\alpha \operatorname{tr}((f_0, f_0)\ \hat{P}_\alpha)$$

$$=\eta_i^\alpha [(f_0, f_0)_{\alpha,\alpha} + (f_0, f_0)_{\alpha+3,\alpha+3}]. \quad (4.45)$$

We define f_0 to satisfy the relation

$$(f_0, f_0)_{\alpha,\alpha} + (f_0, f_0)_{\alpha+3,\alpha+3} = 1, \quad (4.46)$$

and hence, if $\|f_i\|^2 = 1$,

$$\eta_i^\alpha \equiv 1. \quad (4.47)$$

It then follows from Eqs. (4.42), (4.44), and (4.47) that

$$\omega_0(M)=\sum \mu_i \operatorname{tr}((f_0, P_M f_0)\ \hat{P}_\alpha)$$

$$=\operatorname{tr}(f_0^\alpha, P_M f_0^\alpha). \quad (4.48)$$

If we had not chosen M_0 to be $\{f_0 \hat{P}_\alpha \mathbf{a}\}$, but chosen $\{f_0 \mathbf{a}\}$ instead, we would have obtained a mixture over all of the α's in Eq. (4.48). ■ The state $\omega_0(M)$ is invariant under the replacement

$$f_0^\alpha = f_0 \hat{P}_\alpha \to f_0 \hat{P}_\alpha \mathbf{a} \quad (4.49)$$

for the $\mathbf{a} \in \mathfrak{A}$ which preserve the norm, i.e., $\mathbf{a}$ satisfying

$$\hat{P}_\alpha \mathbf{aa}^* \hat{P}_\alpha = \hat{P}_\alpha. \quad (4.50)$$

The transformations (4.49), leaving pure states invariant, are therefore analogous to the phase transformations of a complex Hilbert space. The quantities $\hat{P}_\alpha \mathbf{a}$ of the minimal right ideals $\hat{P}_\alpha \mathfrak{A}$ of $\mathfrak{A}$ [satisfying Eq. (4.50)] play the role of a *generalized phase* which could be utilized to construct local non-Abelian gauge transformations. We shall call $\mathfrak{A}$ the generalized phase algebra.

V. SUPERSELECTION RULES AND SYMMETRIES

For an operator A linear with respect to $\mathfrak{A}$ the expectation value calculated with an arbitrary vector $f \in \mathscr{H}$, in the domain of A decomposes as if the state were mixed, i.e.,

$$\langle A \rangle_f = \operatorname{tr}(f, Af) = \sum_\alpha \operatorname{tr}(f^\alpha, Af^\alpha), \quad (5.1)$$

where $f^\alpha = f\hat{P}_\alpha \in \mathscr{H}_\alpha$. In fact, we shall adapt the procedures of Ref. 13 to prove:

Statement 3: *If all observables belong to the class of self-adjoint (with respect to a scalar product linear over $\mathfrak{A}$) operators linear with respect to $\mathfrak{A}$, then the embedding of the Hilbert space $\mathscr{H}$, (as $\mathscr{H}_{\mathfrak{A}}$, with manifolds linear over the generalized phase algebra) into $\mathscr{H}_c$ results in a representation which is reducible with respect to the primitive idempotents $\hat{P}_\alpha$, i.e.,, displays superselection rules for the subspaces $\mathscr{H}_\alpha$.*

We shall proceed by first making explicit the structure of the algebras $\mathfrak{A}_\xi$, and defining a scalar product appropriate for the consideration of operators linear with respect to $\mathfrak{A}$.

From Eq. (4.35), we obtain for the algebra $\mathfrak{A}_{+++}$, defined by Eq. (4.34), for $\alpha,\beta = 0,1,2,3$,

$$\hat{P}_\alpha \mathbf{a} \hat{P}_\beta = (\rho_{\alpha\beta} + \rho_{\alpha+3,\beta+3})(a_{\alpha+3,\beta+3} + e_7 a_{\alpha+3,\beta}). \quad (5.2)$$

With the help of the multiplication rules given in Eqs. (2.20), one obtains ($\alpha,\beta = 1,2,3$, or $\alpha = \beta = 0$)

$$\rho_{\alpha\beta} + \rho_{\alpha+3,\beta+3} = e_\alpha P_0 e_\beta^* + e_7 e_\alpha P_0 e_\beta^* e_7^*$$

$$= e_\alpha \hat{P}_0 e_\beta^*, \quad (5.3)$$

and ($\alpha = 1,2,3$)

$$\rho_{0\alpha}+\rho_{7,\alpha+3}=P_0e^*_\alpha+e_7P_0e^*_\alpha e^*_7$$

$$=(P_0-P_7)\;e^*_\alpha. \tag{5.4}$$

The algebra $\mathfrak{A}_{+++}$ is therefore generated over the complex subalgebra $\mathbb{C}(1,e_7)$ according to Eqs. (5.2), (5.3), and (5.4) by

$$\hat{\rho}_{\alpha\beta}^{+++}=e_\alpha\hat{P}_0e^*_\beta,\quad \alpha,\beta=1,2,3,$$

$$\hat{\rho}_{0\alpha}^{+++}=(P_0-P_7)\;e^*_\alpha,\quad \alpha=1,2,3,$$

$$\hat{\rho}_{00}^{+++}=\hat{P}_0. \tag{5.5}$$

As already pointed out, e_7 commutes with $\mathfrak{A}_{+++}$ and hence with these $\hat{\rho}_{\alpha\beta}^{+++}$. The general element of $\mathfrak{A}_{+++}$ has the representation

$$\mathbf{a}=\sum_{\alpha,\beta=0}^{3}\mathbf{a}_{\alpha\beta}\hat{\rho}_{\alpha\beta}^{+++}, \tag{5.6}$$

where $\mathbf{a}_{\alpha\beta}\in\mathbb{C}(1,e_7)$, and the $\hat{\rho}_{\alpha\beta}^{+++}$ satisfy $(\alpha,\beta,\gamma,\delta=0,1,2,3)$

$$\hat{\rho}_{\alpha\beta}^{+++}\hat{\rho}_{\gamma\delta}^{+++}=\delta_{\beta\gamma}\hat{\rho}_{\alpha\delta}^{+++},$$

$$\hat{\rho}_{\alpha\alpha}^{+++}=\hat{P}_\alpha. \tag{5.7}$$

It is evident from Eq. (5.6) that Eqs. (4.13) and (4.17) are satisfied, i.e.,

$$\hat{P}_\alpha\mathbf{a}\hat{P}_\alpha=\mathbf{a}_{\alpha\alpha}\hat{P}_\alpha,\quad \mathbf{a}_{\alpha\alpha}\in\mathbb{C}(1,e_7),$$

and

$$\mathbf{a}\hat{P}_\alpha\mathbf{a}'=\sum_{\beta\gamma}\mathbf{a}_{\beta\alpha}\mathbf{a}'_{\alpha\gamma}\hat{\rho}_{\beta\gamma}^{+++},$$

where for arbitrary $\mathbf{a},\mathbf{a}'\in\mathfrak{A}_{+++}$, $\mathbf{a}\hat{P}_\alpha\mathbf{a}'\in\mathfrak{A}_{+++}$, and there exists a set of $\mathbf{a}_j,\mathbf{a}'_j$ such that $\Sigma_j\mathbf{a}_j\hat{P}_\alpha\mathbf{a}'_j=\mathbf{a}''$ for every $\mathbf{a}''\in\mathfrak{A}_{+++}$.

We now turn to the general case, and study the algebras $\mathfrak{A}_\xi$ (types 0,1,2,3) simultaneously, with the help of:

Lemma 3.1: *The algebra $\mathfrak{A}_\xi$ is characterized by the fact that it commutes with*

$$e^\xi_7=\sum_{\alpha=0}^{3}\xi_{0\alpha}\hat{P}_\alpha e_7. \tag{5.8}$$

The set of imaginary units e^ξ_7 $[(e^\xi_7)^2=-1, e^{\xi*}_7=-e^\xi_7]$ are linear combinations of e_7, e_3e_6, e_1e_4, and e_2e_5.

The products e_3e_6, e_1e_4, and e_2e_5 all form $+e_7$ on the (right) subspace generated by the projection operator $\hat{P}_0$, and e_7 with various signs on the other $\mathscr{H}_\alpha$. In fact, according to Eq. (5.8) e^ξ_7 takes on the value $\xi_{0\alpha}e_7$ on each of the $\mathscr{H}_\alpha$.

We may interpret this phenomenon as follows. Since each of the $\mathscr{H}_\alpha$ are (according to Statement 3) superselection subspaces, no observable can connect them, and the sign of the imaginary unit may be chosen independently in each. The eight solutions e^ξ_7 correspond to the eight ways of making this choice (up to an over-all sign). The relation between each of these possible solutions for the generalized phase algebra then corresponds to a set of "time reversal" transformations, taken independently in each of the $\mathscr{H}_\alpha$. The algebra $\mathfrak{A}$ is therefore unique up to the choice of the sign of e_7 in each superselection subspace.

We now prove the Lemma 3.1. Following the method used in Eqs. (4.36) and (4.37), we note that ($\xi_{00}=+1$)

$$e^\xi_7=\sum_{\alpha=0}^{3}\xi_{0\alpha}(\rho_{\alpha+3,\alpha}-\rho_{\alpha,\alpha+3}), \tag{5.9}$$

and hence

$$a\,e^\xi_7=\sum_{\alpha=0}^{3}\xi_{0\alpha}(a_{i\alpha+3}\rho_{i\alpha}-a_{i\alpha}\rho_{i\alpha+3}),$$

$$e^\xi_7a=\sum_{\alpha=0}^{3}\xi_{0\alpha}(a_{\alpha i}\rho_{\alpha+3,i}-a_{\alpha+3,i}\rho_{\alpha i}). \tag{5.10}$$

Comparing corresponding terms, we obtain the relations [Eqs. (4.21), (4.25), and (4.29)]

$$a_{\beta,\alpha+3}=-\xi_{\alpha\beta}a_{\beta+3,\alpha},$$

$$a_{\beta,\alpha}=\xi_{\alpha\beta}a_{\beta+3,\alpha+3},$$

where $\xi_{\alpha\beta}=\xi_{0\alpha}\xi_{0\beta}$, characterizing the algebra $\mathfrak{A}_\xi$. From Eqs. (4.4), (2.11) (defining $\mathscr{H}_+$), and (4.3) [or from Eqs. (4.7)], we find

$$\hat{P}_0=\tfrac{1}{4}(I-e_6e_7e_3-e_4e_7e_1-e_5e_7e_2),$$

$$\hat{P}_1=\tfrac{1}{4}(I+e_6e_7e_3-e_4e_7e_1+e_5e_7e_2),$$

$$\hat{P}_2=\tfrac{1}{4}(I+e_6e_7e_3+e_4e_7e_1-e_5e_7e_2),$$

$$\hat{P}_3=\tfrac{1}{4}(I-e_6e_7e_3+e_4e_7e_1+e_5e_7e_2). \tag{5.11}$$

The expressions (5.8) defining e^ξ_7 therefore all lead to linear combinations of e_7, e_3e_6, e_1e_4, and e_2e_5. For the examples given in Eqs. (4.34) [note that $\mathrm{tr}(e_7e_7^{--+})=0$]

$$e_7^{+++}=e_7,$$

$$e_7^{++-}=\tfrac{1}{2}(e_7+e_1e_4+e_2e_5-e_3e_6),$$

$$e_7^{--+}=+e_3e_6,$$

$$e_7^{---}=\tfrac{1}{2}(e_7-e_1e_4-e_2e_5-e_3e_6).\quad\blacksquare \tag{5.12}$$

For the general case $\mathfrak{A}_\xi$, we follow the procedure used above for $\xi=(+++)$ to find that $\mathfrak{A}_\xi$ is generated over $\mathbb{C}(1,e_7)$ by

$$\hat{\rho}^\xi_{\alpha\beta}=e_\alpha(P_0\xi_{0\alpha}\xi_{0\beta}+P_7)\;e^*_\beta,\quad \alpha,\beta=1,2,3,$$

$$\hat{\rho}^\xi_{0\alpha}=(P_0\xi_{0\alpha}-P_7)\;e^*_\alpha,\quad \alpha=1,2,3,$$

$$\hat{\rho}^\xi_{00}=\hat{P}_0. \tag{5.13}$$

This basis satisfies the relations (5.7) as well. For example, for $\alpha,\beta,\gamma,\delta\neq0$

$$\hat{\rho}^\xi_{\alpha\beta}\hat{\rho}^\xi_{\gamma\delta}=e_\alpha(P_0\xi_{0\alpha}\xi_{0\beta}+P_7)\;e^*_\beta e_\gamma(P_0\xi_{0\gamma}\xi_{0\delta}+P_7)\;e^*_\delta$$

$$=\begin{cases}0, & \beta\neq\gamma,\\ e_\alpha(P_0\xi_{0\alpha}\xi_{0\delta}+P_7)\;e^*_\delta, & \beta=\gamma\,,\end{cases}$$

and

$$\hat{\rho}^\xi_{\alpha\beta}\hat{\rho}^\xi_{\gamma0}=e_\alpha(P_0\xi_{0\alpha}\xi_{0\beta}+P_7)\;e^*_\beta e_\gamma(P_0\xi_{0\gamma}-P_7)$$

$$=\begin{cases}0, & \beta\neq\gamma,\\ e_\alpha(P_0\xi_{0\alpha}-P_7), & \beta=\gamma\,.\end{cases}$$

The basis elements $\hat{\rho}^{\xi}_{\alpha\beta}$ therefore satisfy the relations (5.7) and the representation (5.6) can be used for elements of $\mathfrak{A}_{\xi}$ if $\hat{\rho}^{+++}_{\alpha\beta}$ is replaced by $\hat{\rho}^{\xi}_{\alpha\beta}$. Algebraic manipulations are complicated, however, by the fact that basis elements for which $\xi_{\alpha\beta}=-1$ or $\xi_{0\alpha}=-1$ anticommute with e_7. By replacing e_7 in $\mathbf{a}_{\alpha\beta}$ by e^{ξ}_{7}, however, one obtains a representation in which the complex scalars commute with the basis. Since.

$$e^{\xi}_{7}\hat{\rho}^{\xi}_{\alpha\beta}=\xi_{0\alpha}e_7\hat{\rho}^{\xi}_{\alpha\beta}, \tag{5.14}$$

with similar relations for right multiplication, it is clear that the use of elements in $\mathbb{C}(1,e^{\xi}_{7})$, i.e.,

$$\mathbf{a}^{\xi}_{\alpha\beta}=\lambda_{\alpha\beta}+e^{\xi}_{7}\mu_{\alpha\beta}, \tag{5.15}$$

$\lambda_{\alpha\beta},\mu_{\alpha\beta}\in\mathbb{R}$, in the representation

$$\mathbf{a}=\sum_{\alpha\beta}\mathbf{a}^{\xi}_{\alpha\beta}\hat{\rho}^{\xi}_{\alpha\beta} \tag{5.16}$$

corresponds to an effective alteration of signs of the coefficients $\mu_{\alpha\beta}$ relative to those used in Eq. (5.6). Just as for $\mathfrak{A}_{+++}$, Eqs. (4.13) and (4.17) are satisfied for $\mathfrak{A}_{\xi}$ [one may use $\mathbb{C}(1,e_7)$ or $\mathbb{C}(1,e^{\xi}_{7})$].

We shall now follow a procedure similar to that used in Sec. II to construct a scalar product sesquilinear in $\mathfrak{A}$. We say that f and g are $\mathfrak{A}$-orthogonal if

$$\operatorname{tr}((f,g)\mathbf{a})=0 \tag{5.17}$$

for all $\mathbf{a}\in\mathfrak{A}$. Since $\mathbf{a}$ has the representation (5.16) (we suppress the designation ξ in the following) it suffices that

$$\begin{aligned}&\operatorname{tr}((f,g)\hat{\rho}_{\alpha\beta})=0,\\ &\operatorname{tr}((f,g)\hat{\rho}_{\alpha\beta}e_7)=0\end{aligned} \tag{5.18}$$

for every α,β. Let us define the scalar product as

$$(f,g)_{\mathfrak{A}}=\sum_{\alpha\beta}(f,g\,\hat{\rho}_{\beta\alpha})_c\hat{\rho}_{\alpha\beta}, \tag{5.19}$$

where we shall use e^{ξ}_{7} in the definitionn of the complex scalar product in place of e_7 when dealing with the algebra $\mathfrak{A}_{\xi}$ [for $\xi\neq(+\ +\ +)$], for reasons to be discussed below.

Clearly $(f,g)_{\mathfrak{A}}=0$ implies Eqs. (5.18). From Eq. (2.26) we have

$$\operatorname{tr}(f,f)=\|f\|^2=\tfrac{1}{8}\sum_i(f,g)_{ii};$$

on the other hand, from the definition (5.19),

$$\tfrac{1}{2}\operatorname{tr}(f,f)_{\mathfrak{A}}=\tfrac{1}{8}\sum_{\alpha}\left[(f,f)_{\alpha\alpha}+(f,f)_{\alpha+3,\alpha+3}\right], \tag{5.20}$$

which is just $\|f\|^2$. Hence the topology of $\mathscr{H}_{\mathfrak{A}}$ is the same as that of $\mathscr{H}_{*}$ (or $\mathscr{H}_c$). Furthermore, conjugating the definition (5.19), we obtain

$$\begin{aligned}(f,g)^{*}_{\mathfrak{A}}&=\sum_{\alpha\beta}\hat{\rho}_{\beta\alpha}(g\,\hat{\rho}_{\beta\alpha}f)_c\\ &=\sum\hat{\rho}_{\beta\alpha}[\operatorname{tr}(g,f\hat{\rho}_{\alpha\beta})-e_7\operatorname{tr}(g,fe_7\hat{\rho}_{\alpha\beta})].\end{aligned} \tag{5.21}$$

To restore the form of the scalar product, two operations are required in Eq. (5.21]. The basis element $\hat{\rho}_{\alpha\beta}$ must be interchanged with e_7 in the second term, and the basis element outside of the brackets must be brought to the right. For the algebra $\mathfrak{A}_{+++}$ these operations are trivial. For the algebras $\mathfrak{A}_{\xi}$, $\xi\neq(+++)$, compensating sign changes occur. If, however, we use in Eq. (5.19) the following definition for the complex scalar product [in place of Eq. (2.43)],

$$(f,g)_{c_{\xi}}=\operatorname{tr}(f,g)-e^{\xi}_{7}\operatorname{tr}(f,g\,e^{\xi}_{7}), \tag{5.22}$$

these operations become trivial for the algebras $\mathfrak{A}_{\xi}$ also. The scalar product defined by Eq. (5.22) results in a Hilbert space $\mathscr{H}_{c_{\xi}}$ with the same properties as $\mathscr{H}_c$, for which $(f,g)_{c_{\xi}}\in\mathbb{C}(1,e^{\xi}_{7})$. We shall discuss the structure of this space later in this section.

Hence, we find

$$(f,g)^{*}_{\mathfrak{A}}=(g,f)_{\mathfrak{A}}. \tag{5.23}$$

We now show that the form defined by Eq. (5.19) is right linear, i.e., for $a\in\mathfrak{A}$,

$$(f,g\ \mathbf{a})_{\mathfrak{A}}=(f,g)_{\mathfrak{A}}\mathbf{a}, \tag{5.24}$$

and is therefore a proper sesquilinear form, appropriate for the consideration of operators linear over $\mathfrak{A}$.

Consider

$$\begin{aligned}(f,g\ \mathbf{a})_{\mathfrak{A}}&=\sum_{\alpha\beta\gamma\delta}(f,g\ \mathbf{a}_{\gamma\delta}\hat{\rho}_{\gamma\delta}\hat{\rho}_{\beta\alpha})_c\hat{\rho}_{\alpha\beta}\\ &=\sum_{\alpha\beta\gamma}(f,g\ \mathbf{a}_{\gamma\beta}\hat{\rho}_{\gamma\alpha})_c\hat{\rho}_{\alpha\beta}.\end{aligned} \tag{5.25}$$

For $\mathfrak{A}_{+++}$, Eq. (5.25) is (due to the linearity of the complex scalar product)

$$(f,g\ \mathbf{a})_{\mathfrak{A}_{+++}}=\sum_{\alpha\beta\gamma}(f,g\ \hat{\rho}^{+++}_{\gamma\alpha})_c\hat{\rho}^{+++}_{\alpha\beta}\mathbf{a}_{\gamma\beta}. \tag{5.26}$$

Since right multiplication provides the following relation,

$$\begin{aligned}(f,g)_{\mathfrak{A}_{+++}}\mathbf{a}&=\sum_{\alpha\beta\gamma\delta}(f,g\ \hat{\rho}^{+++}_{\beta\alpha})_c\hat{\rho}^{+++}_{\alpha\beta}\mathbf{a}_{\gamma\delta}\hat{\rho}^{+++}_{\gamma\delta}\\ &=\sum_{\alpha\gamma\delta}(f,g\ \hat{\rho}^{+++}_{\gamma\alpha})_c\hat{\rho}^{+++}_{\alpha\delta}\mathbf{a}_{\gamma\delta},\end{aligned} \tag{5.27}$$

linearity for $\mathfrak{A}_{+++}$ is evident. The demonstration for $\mathfrak{A}_{\xi}$, for $\xi\neq(+++)$, is a little more involved, using a representation of the form $\mathbf{a}=\Sigma\,\mathbf{a}_{\alpha\beta}\hat{\rho}^{\xi}_{\alpha\beta}$, $a_{\alpha\beta}\in\mathbb{C}(1,e_7)$, and requires keeping track of an exceptional index set. Using, however, the scalar product defined by Eq. (5.22) and the representation (5.16) for $\mathbf{a}\in\mathfrak{A}_{\xi}$, the arguments leading to Eqs. (5.26) and (5.27) are immediately available; hence it is evident that Eq. (5.24) holds in these cases as well.

Together with the conjugation symmetry Eq. (5.23), these results demonstrate that the scalar product defined by Eq. (5.19) [with, for convenience, the complex scalar product Eq. (5.22) in case the algebra $\mathfrak{A}_{\xi}$ is to be used] is a sesquilinear form.

Having defined $\mathscr{H}_{\mathfrak{A}}$ in terms of the algebra $\mathfrak{A}$ and the scalar product Eq. (5.19), we may now consider embedding $\mathscr{H}_{\mathfrak{A}}$ into $\mathscr{H}_c$, and complete the proof of Statement 3.

Operators linear with respect to $\mathfrak{A}$, satisfying on appropriate domains,

$$(f,Ag)_{\mathfrak{A}} = (Af,g)_{\mathfrak{A}}, \tag{5.28}$$

are said to be Hermitian in $\mathscr{H}_{\mathfrak{A}}$, and have a spectral resolution containing projections on manifolds closed over $\mathfrak{A}$.[12] In the treatment of $\mathscr{H}_{\mathfrak{A}}$ as a Hilbert space over $\mathbb{C}(1,e_7)$ [or $\mathbb{C}(1,e_7^{\xi})$; we shall suppress reference to the designation ξ in the following whenever parallel arguments apply], it is appropriate to use projections into manifolds closed over $\mathbb{C}(1,e_7)$.

If M_c^{α} is a linear manifold closed over $\mathbb{C}(1,e_7)$ (as in Statement 1) in $\mathscr{H}_{\alpha}$, and M_c^{β} in $\mathscr{H}_{\beta}$ $(\beta\neq\alpha)$, then

$$P_{M_c^{\alpha}} P_{M_c^{\beta}} = 0. \tag{5.29}$$

This relation follows from the fact that

$$P_{M_c^{\alpha}} f = h\,\hat{P}_{\alpha} \in \mathscr{H}_{\alpha},$$

$$P_{M_c^{\beta}} g = h'\,\hat{P}_{\beta} \in \mathscr{H}_{\beta},$$

and hence

$$(P_{M_c^{\alpha}} f, P_{M_c^{\beta}} g)_c = (f, P_{M_c^{\alpha}} P_{M_c^{\beta}} g)_c$$

$$= (h\,\hat{P}_{\alpha}, h'\,\hat{P}_{\beta})_c,$$

which vanishes due to the fact that $\hat{P}_{\alpha}$, $\hat{P}_{\beta}$ commute with e_7 (and e_7^{ξ}).

According to Ref. 12, the function $f_{\lambda}(A)$ of a bounded Hermitian operator A may be defined, where

$$f_{\lambda}(x) = \max(x-\lambda, 0)$$

for $x, \lambda \in \mathbb{R}$. We now define the linear manifold

$$M_c^{\alpha}(A,\lambda) = \{g \mid f_{\lambda}(A)g = 0,\ g \in \mathscr{H}_{\alpha}\}. \tag{5.30}$$

By the procedure used in Ref. 12, it can easily be shown that

$$M_c^{\alpha}(A,\lambda) = \begin{cases} \mathscr{H}_{\alpha}, & \lambda > C, \\ \phi, & \lambda < -C, \end{cases} \tag{5.31}$$

where C is the bound of A. Since $\mathscr{H} = \Sigma_{\alpha}\mathscr{H}_{\alpha}$, the unity operator is given by

$$\sum_{\alpha} P_{M_c^{\alpha}(A,\lambda)}$$

for $\lambda > C$, and again following the procedure of Ref. 12, we find that

$$(f,Af)_c = \sum_{\alpha} \int \lambda d\,(f, P_{M_c^{\alpha}(A,\lambda)} f)_c. \tag{5.32}$$

Replacing f by $f \pm g$ and using the complex Hermitian property of the $P_{M_c^{\alpha}(A,\lambda)}$, one obtains

$$(f,Ag)_c = \sum_{\alpha} \int \lambda d\,(f, P_{M_c^{\alpha}(A,\lambda)} g)_c, \tag{5.33}$$

and hence

$$A = \sum_{\alpha} \int \lambda d\, P_{M_c^{\alpha}(A,\lambda)} \tag{5.34}$$

is the operator form of the spectral decomposition of A in $\mathscr{H}_c$. The sum $\Sigma_{\alpha} P_{M_c^{\alpha}(A,\lambda)}$ is a projection since each term is a projection and Eq. (5.29) is valid. The sum is furthermore linear over $\mathfrak{A}$ since it corresponds to the projection into a manifold $\{g\}$ satisfying $f_{\lambda}(A)g = 0$, for g unrestricted in $\mathscr{H}_{++}$, and A is linear over $\mathfrak{A}$ [so that $f_{\lambda}(A)(g\mathbf{a}) = (f_{\lambda}(A)g)\mathbf{a}$, $\mathbf{a} \in \mathfrak{A}$]. Hence, Eq. (5.34) is also the spectral resolution of A in $\mathscr{H}_{\mathfrak{A}}$, and the spectral family in $\mathscr{H}_{\mathfrak{A}}$ is

$$P(\lambda) = \sum_{\alpha} P_{M_c^{\alpha}(A,\lambda)}. \tag{5.35}$$

The form (5.34) explicitly exhibits the reduction of bounded Hermitian operators linear over $\mathfrak{A}$ in $\mathscr{H}_{\mathfrak{A}}$ when embedded in $\mathscr{H}_c$ (this construction can be extended to the unbounded case).■

We now turn to a discussion of the structure of the wavefunctions of Eq. (2.39), the action of $\mathfrak{A}$ on these wavefunctions, and the symmetry represented by the elements of $\mathfrak{A}$ which leave invariant the expectation value Eq. (5.1).

Consider now the wavefunction f in the form given in Eq. (2.39). Since $P_0\psi_{\beta}^k e_{\beta}\hat{P}_{\alpha} = P_0 e_{\beta}\hat{P}_{\alpha}\psi_{\beta}^{k*}$ and $\hat{P}_{\alpha} = P_{\alpha} + P_{\alpha+3}$, we obtain (for each $\alpha = 0,1,2,3$)

$$f\hat{P}_{\alpha} = \sum_k e_k P_0 \psi_{\alpha}^k e_{\alpha} \tag{5.36}$$

for the projection of the general wavefunction into each superselection sector.

For $\mathbf{a} \in \mathfrak{A}_{+++}$, we again use the representation (5.6) to obtain

$$f\hat{P}_0\mathbf{a}$$

$$= \sum_{k,\delta} e_k P_0 \psi_0^k \mathbf{a}_{0\delta} \hat{\rho}_{0\delta}^{+++}$$

$$= \sum_k e_k P_0 \psi_0^k \Big[\mathbf{a}_{00}(P_0 + P_7) + \sum_{\delta=1}^{3} \mathbf{a}_{0\delta}(P_0 - P_7)e_{\delta}^*\Big]$$

$$= \sum_k e_k P_0 \psi_0^k \mathbf{a}_{00} - \sum_{k,\delta=1}^{3} e_k P_0 \psi_0^{k*} \mathbf{a}_{0\delta} e_{\delta}, \tag{5.37}$$

since $P_0 - P_7$ anticommutes with e_7. Similarly, for $\alpha \neq 0$,

$$f\hat{P}_{\alpha}\mathbf{a} = \sum_{k,\delta} e_k P_0 \psi_{\alpha}^k e_{\alpha} \mathbf{a}_{\alpha\delta} \hat{\rho}_{\alpha\delta}^{+++}$$

$$= -\sum_k e_k P_0 \psi_{\alpha}^k \mathbf{a}_{\alpha 0}^*(P_0 - P_7)$$

$$- \sum_{k,\delta=1}^{3} e_k P_0 \psi_{\alpha}^k a_{\alpha\delta}^*(P_0 + P_7)e_{\delta}^*$$

$$= -\sum_k e_k P_0 \psi_{\alpha}^{k*} \mathbf{a}_{\alpha 0} + \sum_{k,\delta=1}^{3} e_k P_0 \psi_{\alpha}^k \mathbf{a}_{\alpha}^* e_{\delta}. \tag{5.38}$$

Comparison of Eq. (5.37) and (5.38) with Eq. (3.8) leads to:

Statement 4: *The generalized phase albebra $\mathfrak{A}_{+++}$ induces transformations with the same pattern of wavefunction conjugations as that induced by the general operator linear over*

$\mathbb{C}(1,e_7)$, *i.e., it contains the same type of "lepton–quark" transitions.*

Before turning to a discussion of the action of $\mathfrak{A}_\xi$ for general ξ, on the wavefunctions, we must first make explicit the structure of the Hilbert space $\mathscr{H}_{c_\xi}$, the complex space defined by the scalar product given in Eq. (5.22) with e_7^ξ as complex unit. This scalar product defines orthogonality between manifolds linear over e_7^ξ. If f, g belong to two such orthogonal linear manifolds, then [using Eqs. (2.26) and (5.9)]

$$\mathrm{tr}(f,g)=\sum_{ki}(f_{ki},g_{ki})$$

$$\mathrm{tr}(f,g)\, e_7^\xi=\sum_{k,\alpha=0}^{3}[(f_{k\alpha},g_{k\alpha+3})-(f_{k\alpha+3},g_{k\alpha})]\xi_{0\alpha} \tag{5.39}$$

must vanish. We define, in place of Eq. (2.38),

$$\psi_0^k=f_{k0}-f_{k7}e_7,$$
$$\psi_\alpha^{k\xi}=-f_{k\alpha}-f_{k\alpha+3}e_7\xi_{0\alpha}$$
$$=\psi_\alpha^k\left(\frac{1+\xi_{0\alpha}}{2}\right)+\psi_\alpha^{k*}\left(\frac{1-\xi_{0\alpha}}{2}\right). \tag{5.40}$$

Using the relation ($\alpha=1,2,3$)

$$(\psi_\alpha^{k\xi},\chi_\alpha^{k\xi})=(\psi_\alpha^k,\chi_\alpha^k)\left(\frac{1+\xi_{0\alpha}}{2}\right)+(\chi_\alpha^k,\psi_\alpha^k)\left(\frac{1-\xi_{0\alpha}}{2}\right)$$
$$=(f_{k\alpha},g_{k\alpha})+(f_{k,\alpha+3},g_{k\alpha+3})$$
$$+\,e_7\xi_{0\alpha}[(f_{k\alpha},g_{k\alpha+3})-(f_{k\alpha+3},g_{k\alpha})], \tag{5.41}$$

we obtain

$$\mathrm{tr}(f,g)=\mathrm{Re}(\psi_0^k,\chi_0^k)+\mathrm{Re}(\psi_\alpha^{k\xi},\chi_\alpha^{k\xi}),$$
$$\mathrm{tr}(g,f)e_7^\xi=-\mathrm{Im}(\psi_0^k,\chi_0^k)+\mathrm{Im}(\psi_\alpha^{k\xi},\chi_\alpha^{k\xi}). \tag{5.42}$$

Hence, Eq. (5.22) may be written as

$$(f,g)_{c_\xi}=(\psi_0^k,\chi_0^k)_\xi+\sum_{\alpha=1}^{3}(\chi_\alpha^{k\xi},\psi_\alpha^{k\xi})_\xi, \tag{5.43}$$

where, for $\alpha=0,1,2,3$,

$$(\psi_\alpha^k,\chi_\alpha^k)_\xi=\mathrm{Re}(\psi_\alpha^k,\chi_\alpha^k)+e_7^\xi\mathrm{Im}(\psi_\alpha^k,\chi_\alpha^k). \tag{5.44}$$

From the identity

$$f=\sum e_kP_0\left\{\psi_0^k+\sum_{\alpha=1}^{3}\left[\psi_\alpha^{k\xi}\left(\frac{1+\xi_{0\alpha}}{2}\right)+\psi_\alpha^{k\xi*}\left(\frac{1-\xi_{0\alpha}}{2}\right)\right]e_\alpha\right\} \tag{5.45}$$

and the fact that e_7^ξ takes on the value $\xi_{0\alpha}e_7$ on each of the $\mathscr{H}_\alpha$, it follows that the analog of Eq. (2.40) is for $f\rightarrow fa_\xi$, $a_\xi\in\mathbb{C}(1,e_7^\xi)$,

$$\psi_0^k\rightarrow\psi_0^k a,$$
$$\psi_\alpha^{k\xi}\rightarrow\psi_\alpha^{k\xi}a^*, \tag{5.46}$$

where $a\equiv a_{+++}$. Making these replacements in Eq. (5.43), one finds that, in analogy with Eq. (3.1),

$$(f,g\,a_\xi)_{c_\xi}=(f,g)_{c_\xi}a_\xi \tag{5.47}$$

and by conjugation

$$(f a_\xi,g)_{c_\xi}=a_\xi^*(f,g)_{c_\xi}. \tag{5.48}$$

Following the procedure of Sec. III, we again define the action of an operator linear over the reals according to Eq. (3.4), and requiring linearity over $\mathbb{C}(1,e_7^\xi)$, i.e., that

$$(Af)a_\xi=A(fa_\xi),$$

we find the superselection rules

$$A_{00}^{kl}=0,\quad A_{0\alpha}^{kl'}=A_{\alpha 0}^{kl'}=0\quad(\xi_{0\alpha}=+1),$$
$$A_{0\alpha}^{kl}=A_{\alpha 0}^{kl}=0\quad(\xi_{0\alpha}=-1),$$
$$A_{\alpha\beta}^{kl'}=0\quad(\xi_{\alpha\beta}=+1),\quad A_{\alpha\beta}^{kl}=0\quad(\xi_{\alpha\beta}=-1). \tag{5.49}$$

We may therefore write the action of the general operator linear over $\mathbb{C}(1,e_7^\xi)$, in a form which displays these superselection rules explicitly, as

$$Af=\sum e_kP_0\left\{A_{00}^{kl}\psi_0^l+\sum_{\alpha=1}^{3}\left[A_{0\alpha}^{kl'}\psi_\alpha^l\left(\frac{1-\xi_{0\alpha}}{2}\right)+A_{0\alpha}^{kl}\psi_\alpha^{l*}\left(\frac{1+\xi_{0\alpha}}{2}\right)+A_{\alpha 0}^{kl'}\psi_0^l\left(\frac{1-\xi_{0\alpha}}{2}\right)+A_{\alpha 0}^{kl}\psi_0^{l*}\left(\frac{1+\xi_{0\alpha}}{2}\right)+\sum_{\beta=1}^{3}\left[A_{\alpha\beta}^{kl}\psi_\beta^l\left(\frac{1+\xi_{\alpha\beta}}{2}\right)+A_{\alpha\beta}^{kl'}\psi_\beta^{l*}\left(\frac{1-\xi_{\alpha\beta}}{2}\right)\right]\right]e_\alpha\right\}. \tag{5.50}$$

Operators linear over $a_\xi\in\mathbb{C}(1,e_7)$ therefore couple lepton–quark and quark–quark spaces in a way distinct from operators linear over $\mathbb{C}(1,e_7)$ [compare Eq. (3.8)]. We shall now demonstrate:

Statement 5: The generalized phase algebra $\mathfrak{A}_\xi$ induces transformations with the same pattern of wavefunction conjugations as that induced by the general operator linear over $\mathbb{C}(1,e_7^\xi)$, i.e., it contains the same type of "lepton–quark" and "quark–quark" transitions.

To prove this statement, we use the general relations given in Eq. (5.13) and the decomposition Eq. (5.36). First, consider the action of $\hat\rho_{0\alpha}^\xi$, for $\alpha=1,2,3$:

$$f\hat P_0\hat\rho_{0\alpha}^\xi=\sum e_kP_0\psi_0^k(P_0\xi_{0\alpha}-P_7)\,e_\alpha^*.$$

Since e_7 anticommutes with P_0-P_7, and commutes with P_0+P_7, we obtain

$$P_0\psi_0^k(P_0\xi_{0\alpha}-P_7)$$
$$=P_0\left[\psi_0^{k*}\left(\frac{1+\xi_{0\alpha}}{2}\right)-\psi_0^k\left(\frac{1-\xi_{0\alpha}}{2}\right)\right]$$

and hence

$$f\hat{P}_0\hat{\rho}_{0\alpha}^{\xi}$$
$$=-\sum e_p P_0\left[\psi_0^{k*}\left(\frac{1+\xi_{0\alpha}}{2}\right)-\psi_0^k\left(\frac{1-\xi_{0\alpha}}{2}\right)e_\alpha\right]. \tag{5.51}$$

Next consider ($\alpha=1,2,3$)

$$f\hat{P}_\alpha\hat{\rho}_{\alpha 0}^{\xi}=-\sum e_k P_0\psi_\alpha^k(P_0\xi_{0\alpha}-P_7).$$

By the same argument used above, we obtain

$$f\hat{P}_\alpha\hat{\rho}_{\alpha 0}^{\xi}$$
$$=-\sum e_k P_0\left[\psi_\alpha^{k*}\left(\frac{1+\xi_{0\alpha}}{2}\right)-\psi_\alpha^k\left(\frac{1-\xi_{0\alpha}}{2}\right)\right]. \tag{5.52}$$

Finally, we study the expression ($\alpha,\beta=1,2,3$)

$$f\hat{P}_\alpha\hat{\rho}_{\alpha\beta}=+\sum_k e_k P_0\psi_\alpha^k(P_0\xi_{\alpha\beta}+P_7)\,e_\beta.$$

The argument used above is again applicable when we write

$$(P_0\xi_{\alpha\beta}+P_7)$$
$$=\left(\frac{1+\xi_{\alpha\beta}}{2}\right)(P_0+P_7)-\left(\frac{1-\xi_{\alpha\beta}}{2}\right)(P_0-P_7),$$

and we obtain

$$f\hat{P}_\alpha\hat{\rho}_{\alpha\beta}^{\xi}$$
$$=\sum_k e_k P_0\left[\left(\frac{1+\xi_{\alpha\beta}}{2}\right)\psi_\alpha^k-\left(\frac{1-\xi_{\alpha\beta}}{2}\right)\psi_\alpha^{k*}\right]e_\beta. \tag{5.53}$$

Comparing the results (5.51), (5.52), (5.53) with the form of Eq. (5.50), along with the known action of $\hat{\rho}_{00}^{\xi}=\hat{P}_0$, we see that the phase algebra $\mathfrak{A}_\xi$ induces transformations in accordance with Statement 5. ∎

We emphasize that the operators linear over e_7^{ξ}, as defined by Eq. (5.50) and the generalized phase algebra $\mathfrak{A}_\xi$, induce "lepton–quark" and "quark–quark" transformations with a pattern of linearity and antilinearity that is different for each of the eight choices of ξ. According to the remarks made in our discussion of Eq. (3.8), this choice therefore influences the structure of "leptoquark" and "diquark" currents present in the corresponding field theory.

We now turn to a discussion of the symmetries associated with the generalized phase algebra and with automorphisms of C_7. We shall confine our discussion to $\mathfrak{A}_{+++}$, although similar arguments are valid for $\mathfrak{A}_\xi$ in general, using the scalar product $(f,g)_{C_\xi}$, defined in Eq. (5.22), for the construction of the complex Hilbert space.

The transformations $\mathfrak{A}$ that leave pure states invariant, i.e., satisfying [Eq. (4.50)]

$$\hat{P}_\alpha \mathbf{a}\mathbf{a}^*\hat{P}_\alpha=\sum \mathbf{a}_{\alpha\beta}\hat{\rho}_{\alpha\beta}\mathbf{a}_{\alpha\beta}^*\hat{\rho}_{\delta\sigma}=\sum_\beta |a_{\alpha\beta}|^2\hat{P}_\alpha=\hat{P}_\alpha,$$

have the property

$$\sum_\beta |a_{\alpha\beta}|^2=1. \tag{5.54}$$

At the beginning of this section, we discussed the expectation value of an operator A linear with respect to $\mathfrak{A}$. The scalar product appropriate for the discussion of operators of this type is the one defined by Eq. (5.19). Due to the properties of the trace, however, [as for the norm (5.20)],

$$\tfrac{1}{2}\mathrm{tr}(f,Af)_{\mathfrak{A}}=\mathrm{tr}(f,Af). \tag{5.55}$$

The generalized phase algebra $\mathfrak{A}$ does not respect the superselection rules associated with the physical description of *obervables* defined to be linear over $\mathfrak{A}$ (Statement 3). The linear manifolds of $\mathscr{H}_{\mathfrak{A}}$ span all of the superselection subspaces; it is the requirement of linearity with respect to this algebra which confines the action of the observables to each minimal ideal. Although the general vector f, which has components in more than one of the $\mathscr{H}_\alpha$, corresponds to a mixed state in the sense of Eq. (4.40), we may still ask for the subalgebra of $\mathfrak{A}$ which leaves Eq. (5.55) invariant. With this, we many extend the notion of generalized phase and the scope of the associated gauge transformations.

In this extension, we obtain:

Statement 6: The subalgebra of $\mathfrak{A}$ which leaves expectation values of operators linear over $\mathfrak{A}$ invariant is its unitary subgroup $U(4)$, and is a realization of the $U(4)$ invariance of the complex scalar product.

To prove Statement 6, we first discuss in a little more detail the properties of an operator linear over $\mathfrak{A}$. We shall treat the algebra $\mathfrak{A}_{+++}$ in what follows, and suppress the index $+++$. Similar arguments are valid for any of the $\mathfrak{A}_\xi$. Comparing Eqs. (5.37) and (5.38) with Eq. (3.8), we see that the transformation $f\rightarrow f\mathbf{a}$ can be performed by the operators B (linear with respect to e_7), where ($\alpha=1,2,3$)

$$B_{00}^{kl}=\mathbf{a}_{00}\delta_{kl}\,B_{0\alpha}^{kl}=-\mathbf{a}_{\alpha 0}\delta_{kl},$$
$$B_{\alpha 0}^{kl}=-\mathbf{a}_{0\alpha}\delta_{kl}\,B_{\alpha\beta}^{kl}=\mathbf{a}_{\beta\alpha}^*\delta_{kl}. \tag{5.56}$$

The condition of linearity,

$$(Af\mathbf{a})=(Af)\mathbf{a},$$

is therefore equivalent to

$$ABf=BAf, \tag{5.57}$$

i.e., A and B commute. With the help of Eqs. (3.12), and defining B as in Eq. (5.56), we obtain

$$(BA)_{00}^{kl}=\mathbf{a}_{00}A_{00}^{kl}-\sum_{\alpha=1}^{3}\mathbf{a}_{\alpha_0}A_{\alpha_0}^{kl*},$$

$$(AB)_{00}^{kl}=\mathbf{a}_{00}A_{00}^{kl}-\sum_{\alpha=1}^{3}\mathbf{a}_{0\alpha}^*A_{0\alpha}^{kl*},$$

$$(BA)_{0\alpha}^{kl}=\mathbf{a}_{00}A_{0\alpha}^{kl}-\sum_{\beta=1}^{3}\mathbf{a}_{\beta 0}A_{\beta\alpha}^{kl*},$$

$$(AB)_{0\alpha}^{kl}=-\mathbf{a}_{\alpha 0}A_{00}^{kl}-\sum_{\beta=1}^{3}\mathbf{a}_{\beta\alpha}A_{0\beta}^{kl},$$

$$(BA)^{kl}_{\alpha 0} = -\mathbf{a}_{0\alpha} A^{kl*}_{00} + \sum_\beta \mathbf{a}^*_{\beta\alpha} A^{kl}_{\beta 0},$$

$$(AB)^{kl}_{\alpha 0} = \mathbf{a}^*_{00} A^{kl}_{\alpha 0} - \sum_\beta \mathbf{a}_{0\beta} A^{kl}_{\alpha\beta},$$

$$(BA)^{kl}_{\alpha\beta} = -\mathbf{a}_{0\alpha} A^{kl*}_{0\beta} + \sum_\gamma \mathbf{a}^*_{\gamma\alpha} A^{kl}_{\gamma\beta},$$

$$(AB)^{kl}_{\alpha\beta} = -\mathbf{a}^*_{\beta 0} A^{kl}_{\alpha 0} + \sum_\gamma \mathbf{a}^*_{\beta\gamma} A^{kl}_{\alpha\gamma} \tag{5.58}$$

With the requirement (5.57), i.e., that $[A,B]=0$, it follows from Eqs. (5.58), which must be valid for all $\mathbf{a} \in \mathfrak{A}$, that ($\alpha = 1,2,3$)

$$A^{kl}_{\alpha 0} = A^{kl}_{0\alpha} = 0,$$

$$A^{kl}_{\alpha\beta} = 0 \quad (\alpha \neq \beta),$$

$$A^{kl*}_{\alpha\alpha} = A^{kl}_{00}. \tag{5.59}$$

For operators A, linear over $\mathfrak{A}$, we therefore have

$$Af = \sum_{k,l} e_k P_0 \Big(A^{kl}_{00} \psi^l_0 + \sum_{\alpha=1}^{3} A^{kl*}_{00}\, \psi^l_\alpha e_\alpha \Big). \tag{5.60}$$

It will be necessary for us to examine in some detail the structure of (f,Af). Introducing the notation

$$\int d\sigma\, f^* Af = (f,Af) \quad,$$

where $d\sigma$ is a measure on the manifold over which the complex wavefunctions are defined, we find ($\alpha,\beta \neq 0$

$$(f,Af)$$

$$= \sum_{kl} \int d\sigma \Big(-\sum_\beta e_\beta \psi^{k*}_\beta + \psi^{k*}_0 \Big) P_0$$

$$\times \Big(A^{kl}_{00}\psi^l_0 + \sum_\alpha A^{kl*}_{00}\, \psi^l_\alpha e_\alpha \Big)$$

$$= \sum_{kl} \int d\sigma \Big(-\sum_\beta \psi^k_\beta e_\beta P_0 A^{kl}_{00} \psi^l_0 + \psi^{k*}_0 P_0 A^{kl}_{00} \psi^l_0$$

$$- \sum_{\beta,\alpha} \psi^k_\beta e_\beta P_0 e_\alpha A^{kl}_{00} \psi^{l*}_\alpha + \sum_\alpha \psi^{k*}_0 P_0 e_\alpha A^{kl}_{00} \psi^{l*}_\alpha \Big). \tag{5.61}$$

The first term of Eq. (5.61) contains the $(\beta,0)$, $(\beta+3,0)$, $(\beta,7)$, $(\beta+3,7)$ parts of (f,Af), the second contains the (0,0), (7,0), (0,7) and (7,7) parts, the third the (β,α), $(\beta+3,\alpha)$, $(\beta,\alpha+3)$, $(\beta+3,\alpha+3)$, and the last term, the $(0,\alpha)$, $(7,\alpha)$, $(0,\alpha+3)$, $(7,\alpha+3)$ parts. Since these all involve the real and imaginary parts of independent wavefunctions, *all of these pieces are independent*. This is the result we shall need, but we note in passing that only the second and third terms can contribute to the trace. Breaking up the wavefunctions into real and imaginary parts for the calculation, and recombining terms, we find

$$\mathrm{tr}(f,Af)$$

$$= \mathrm{tr} \int d\sigma\, \psi^{k*}_0 P_0 A^{kl}_{00} \psi^l_0 + \sum_\alpha \mathrm{tr} \int d\sigma\, \psi^k_\alpha P_\alpha A^{kl}_{00} \psi^{l*}_\alpha$$

$$= \tfrac{1}{8} \mathrm{Re} \int d\sigma\, \psi^{k*}_0 A^{kl}_{00} \psi^l_0 + \tfrac{1}{8} \sum_\alpha \mathrm{Re} \int d\sigma\, \psi^k_\alpha A^{kl}_{00} \psi^{l*}_\alpha. \tag{5.62}$$

We now return to the question of the invariance of the expectation value of an operator linear over $\mathfrak{A}$. We require that

$$\mathrm{tr}(f\mathbf{a}), A(f\mathbf{a}))_{\mathfrak{A}} = \mathrm{tr}((f,Af)\,\mathfrak{A}\mathbf{a}\mathbf{a}^*)$$

$$= \mathrm{tr}(f,Af)\,\mathfrak{A}. \tag{5.63}$$

Calling $\mathbf{b} = \mathbf{a}\mathbf{a}^* - I\epsilon$, Eq. (5.63) can be written as

$$\sum_{\alpha\beta=0}^{3} \mathrm{tr}((f,Af\,\hat\rho_{\beta\alpha})_c \hat\rho_{\alpha\beta}\mathbf{b}) = 0. \tag{5.64}$$

Now, due to the linearity of A, one finds

$$(f,Af\,\hat\rho_{\beta\alpha})_c = (f,Af)_{\alpha\beta} + (f,Af)_{\alpha+3,\beta+3}$$

$$+ e_7[(f,Af)_{\alpha,\beta+3} - (f,Af)_{\alpha+3,\beta}], \tag{5.65}$$

and hence, Eq. (5.64) implies

$$\sum_{\alpha\beta=0}^{3} \{ [(f,Af)_{\alpha\beta} + (f,Af)_{\alpha+3,\beta+3}](\mathrm{Re}\mathbf{b}_{\beta\alpha})$$

$$- [(f,Af)_{\alpha,\beta+3} - (f,Af)_{\alpha+3,\beta}](\mathrm{Im}\mathbf{b}_{\beta\alpha})\} = 0. \tag{5.66}$$

The independence of the pieces of (f,Af) appearing in Eq. (5.66) which we have already established, then implies that $\mathbf{b}=0$ (if A is restricted to operators self-adjoint in $\mathscr{H}_{\mathfrak{A}}$, then Eq. (5.66) implies only that the symmetric part of $\mathrm{Re}\mathbf{b}_{\alpha\beta}$ and the antisymmetric part of $\mathrm{Im}\mathbf{b}_{\beta\alpha}$ vanish; since $\mathbf{b} = \mathbf{a}\mathbf{a}^* - I$, this is, however, all of $\mathbf{b}_{\beta\alpha}$). In terms of the components of $\mathbf{a}$ in $\mathbb{C}(1,e_7)$, the expectation value of an operator linear with respect to $\mathfrak{A}$ is invariant under transformation induced by $\mathbf{a} \in \mathfrak{A}$ satisfying

$$\sum_\beta \mathbf{a}_{\alpha\beta} \mathbf{a}^*_{\gamma\beta} = \delta_{\alpha\gamma}, \tag{5.67}$$

that is, the transformations of U(4).

Since $\mathbf{a} \in \mathfrak{A}$ commutes with e_7, it is easy to see that the complex scalar product is invariant under these transformations:

$$(f\mathbf{a}, g\mathbf{a})_c = \mathrm{tr}((f,g)\ \mathbf{a}\mathbf{a}^*) + e_7 \mathrm{tr}((f,g)e_7\, \mathbf{a}\mathbf{a}^*)$$

$$= (f,g)_c. \ \blacksquare \tag{5.68}$$

In the last part of this section, we discuss a class of automorphisms of the algebra and its connection to the U(4) symmetry of the states. It is well known that G_2 is the group of automorphisms of the Cayley algebra. The fact that P_0 determines, through the equivalence relations given by Eqs. (2.20), multiplication rules for elements of the Clifford algebra which are equivalent to that of the octonions implies:

Statement 7: The automorphisms of C_7 which leave P_0 invariant provide a realization of the group G_2 in the ideal generated by P_0.

We wish to make this realization explicit. Let

$$U=\sum u_{kl}\rho_{kl}, \tag{5.69}$$

where $u_{kl}\in\mathbb{R}$ and

$$\sum_{l=0}^{7} u_{kl}u_{ml}=\delta_{km},$$

represent the O(8) of automorphisms of C_7 (in $\mathscr{H}_+$). These transformations, acting as a generalized phase in $\mathscr{H}_+$, leave invariant the real scalar product $\mathrm{tr}(f,g)$, and correspond to orthogonal transformations among $\{fP_k\}$, $k=0,1,\dots,7$. If

$$UP_0U^*=\sum_{kl=0}^{7} u_{k0}u_{l0}\rho_{kl}=P_0, \tag{5.70}$$

then $u_{k0}=\pm\delta_{k0}$. The transformations which leave P_0 invariant, up to an over-all sign, are therefore

$$U=P_0+\sum_{kl=1}^{7} t_{kl}\rho_{kl}. \tag{5.71}$$

The requirement $UU^*=I$ implies

$$\sum_{l=1}^{7} t_{kl}t_{ml}=\delta_{km}, \tag{5.72}$$

so that U generates transformations (reducibly) in O(7). Since P_0 is invariant, the equivalence relations for the $\{e'_k\}$, defined by

$$e'_k=Ue_kU^*, \tag{5.73}$$

are the same as those for the $\{e_k\}$. The action of this O(7) projected into the minimal ideal generated by P_0 is therefore that of G_2. We shall discuss, in the following, the mechanism for this projected action.

From Eq. (5.71), for $k=1,2,\dots,7$,

$$\begin{aligned} e'_k&=\Big(P_0+\sum_{lm=1}^{7} t_{lm}\rho_{lm}\Big)e_k\Big(P_0+\sum_{kj=1}^{7} t_{ij}\rho_{ji}\Big)\\ &=\sum_{l=1}^{7}(e_lt_{lk}P_0+P_0e_lt_{lk})\\ &\quad+\sum_{ijlm=1}^{7} t_{lm}t_{ij}\sigma_{mkj}\rho_{li}, \end{aligned} \tag{5.74}$$

where we have defined

$$\rho_{lm}e_k\rho_{ji}=\sigma_{mkj}\rho_{li}, \tag{5.75}$$

so that $\sigma_{mkj}=0,\pm1$ according to whether the multiplication rules $P_0e_me_k=\pm P_0e_j$ are realized or not for the indices m,k,j (cyclic). We note that

$$UP_0e_kU^*=P_0e'_k=P_0e''_k, \tag{5.76}$$

where

$$e''_k=\sum_{l=1}^{7} e_lt_{lk} \tag{5.77}$$

corresponds to the O(7) of automorphisms that leave $e_0=1$ invariant. It is the more highly structured function $\{e'_k\}$ which satisfy the equivalence relations Eqs. (2.20), by construction, and not the $\{e''_k\}$. In terms of the $\{e''_k\}$, Eq. (5.74) can be written as

$$e'_k=e''_kP_0+P_0e''_k+\sum_{mj=1}^{7}\sigma_{mkj}e''_mP_0e''^*_j, \tag{5.78}$$

which is, in fact, evident from Eq. (5.76) and the fact that the

$$\rho'_{ij}=e'_iP_0e'^*_j=e''_iP_0e''^*_j,\quad i,j=0,1,\dots,7, \tag{5.79}$$

form a complete basis for C_7. For $k\neq l$,

$$P_0e'_ke'_l=P_0e''_k\sum_{ijnm=1}^{7} t_{ij}t_{nm}\sigma_{ljm}e_nP_0e^*_i. \tag{5.80}$$

Choosing a subset of the t_{ij} which belongs to the group G_2, characterized by the invariance of the antisymmetric tensor

$$\sigma_{ijk}=\sum_{lmn=1}^{7} t_{il}t_{jm}t_{kn}\sigma_{lmn}, \tag{5.81}$$

we may use the relation

$$\sum_{mj=1}^{7} t_{ij}t_{nm}\sigma_{ljm}=\sum_{q=1}^{7}\sigma_{inq}t_{ql}, \tag{5.82}$$

and hence Eq. (5.80) becomes

$$\begin{aligned} P_0e'_ke'_l&=P_0e''_k\sum_{inq=1}^{7}\sigma_{inq}t_{ql}e_nP_0e^*_i\\ &=P_0\sum_{inq=1}^{7} t_{ql}t_{nk}e_i\sigma_{inq}. \end{aligned} \tag{5.83}$$

Again making use of Eq. (5.81), we obtain

$$\sum_{nq=1}^{7}\sigma_{inq}t_{nk}t_{ql}=\sum_{j=1}^{7} t_{ij}\sigma_{jkl} \tag{5.84}$$

so that the required equivalence relation,

$$\begin{aligned} P_0e'_ke'_l&=\sum_{ij}P_0\sigma_{klj}e_it_{ij}\\ &=P_0\sigma_{klj}e''_j, \end{aligned} \tag{5.85}$$

follows.

We have derived Eq. (5.85) by assuming that the t_{ij} are elements of G_2, but is, in fact, *not* necessary that the t_{ij} satisfy Eq. (5.81) in order to obtain this result. Replacing e''_k in Eq. (5.80) by the relation given in Eq. (5.77), one obtains

$$\begin{aligned} &P_0e'_ke'_l\\ &=P_0\sum_{ijnmq=1}^{7} e_qt_{qk}t_{ij}t_{nm}\sigma_{ljm}e_nP_0e^*_i\\ &=P_0\sum_{ijmn=1}^{7} t_{nk}t_{nm}t_{ij}\sigma_{ljm}e_i, \end{aligned} \tag{5.86}$$

which reduces to Eq. (5.85) after utilizing the orthogonality of the t_{ij} in the first two factors. The G_2 symmetry, appearing in the behavior of the mapping $\{e_k\}\to\{e'_k\}$ in the equivalence relations obtained from the ideal generated by P_0, leads to an algebraic structure which would be obtained if the t_{ij} were restricted to G_2, but is, in fact, valid for all t_{ij} in O(7).

We now turn to a discussion of automorphisms on wavefunctions of the form Eq. (2.39). In particular, we shall be interested in automorphisms which leave the 7 direction invariant. If $U = U_0$ leaves e_7 invariant, it must belong to $\mathfrak{A}$. Furtherfore, let U_0 leave P_0 invariant, i.e., $U_0 P_0 U_0^* = U_0^* P_0 U_0 = P_0$; in this case, the representation Eq. (5.6) has the form

$$U_0 = \hat{P}_0 + \sum_{\alpha,\beta=1}^{3} \mathbf{a}_{\alpha\beta}\hat{\rho}_{\alpha\beta}, \tag{5.87}$$

where

$$\sum_{\beta=1}^{3} \mathbf{a}_{\alpha\beta}\mathbf{a}^*_{\gamma\beta} = \delta_{\alpha\gamma}, \tag{5.88}$$

i.e., $\mathbf{a}_{\alpha\beta}$ belongs to U(3). This result is parallel to that of Eq. (5.72), where t_{ij} was found to be an element of O(7); just as U defined by Eq. (5.71) acts as an element of G_2 in the ideal generated by P_0, U_0 defined by Eq. (5.87) must act as an element of SU(3) in the ideal generated by P_0, since it belongs to a subgroup of G_2. We shall make this result explicit in the demonstration of:

Statement 8: *The automorphisms of C_7 which leave P_0 and e_7 invariant provide a realization of SU(3) in the ideal generated by P_0, and coincides with the SU(3) subgroup of the generalized phase U(4).*

Under the automorphisms generated by U_0, $e_\alpha \rightarrow e'_\alpha$, where

$$\begin{aligned} e'_\alpha &= U_0 e_\alpha U_0^* \\ &= \Big(\hat{P}_0 + \sum_{\sigma\beta=1}^{3} \mathbf{a}_{\sigma\beta}\hat{\rho}_{\sigma\beta}\Big) e_\alpha \Big(\hat{P}_0 + \sum_{\gamma\delta=1}^{3} \mathbf{a}^*_{\gamma\delta}\hat{\rho}_{\delta\gamma}\Big) \\ &= \sum_{\gamma=1}^{3} (\hat{P}_0 e_\alpha \hat{\rho}_{\alpha\gamma}\mathbf{a}^*_{\gamma\alpha} + \mathbf{a}_{\gamma\alpha}\hat{\rho}_{\gamma\alpha} e_\alpha \hat{P}_0) \\ &\quad + \sum_{\gamma\delta\sigma\beta=1}^{3} \mathbf{a}_{\sigma\beta}\hat{\rho}_{\sigma\beta} e_\alpha \hat{\rho}_{\delta\gamma}\mathbf{a}_{\gamma\delta}. \end{aligned} \tag{5.89}$$

For the evaluation of the last term, we note that $\hat{P}_0 e^*_\beta e_\alpha e_\delta \hat{P}_0 \neq 0$ only when $\alpha \neq \beta \neq \delta$, and then $P_0 e^*_\beta e_\alpha e_\delta = \pm P_0$. Evaluated on P_7, the sign must be opposite since e_7 anticommutes with the $\{e_\alpha\}$. From Eq. (2.13), (for $\epsilon_{\beta\alpha\delta}$ the usual totally antisymmetric tensor in three dimensions)

$$P_0 e^*_\beta e_\alpha e_\delta = P_0 \epsilon_{\beta\alpha\delta} \quad \text{(cyclic)}, \tag{5.90}$$

and therefore we obtain

$$\hat{P}_0 e^*_\beta e_\alpha e_\delta \hat{P}_0 = \epsilon_{\beta\alpha\delta}(P_0 - P_7). \tag{5.91}$$

As in Eq. (5.74) we therefore have

$$e'_\alpha = \hat{P}_0 \sum_{\beta=1}^{3} \mathbf{a}_{\beta\alpha} e_\beta + \sum_{\beta=1}^{3} \mathbf{a}_{\beta\alpha} e_\beta \hat{P}_0 + \sum_{\beta\delta\sigma\gamma=1}^{3} \epsilon_{\beta\alpha\delta}\mathbf{a}_{\sigma\beta} e_\sigma (P_0 - P_7)\, e^*_\gamma \mathbf{a}^*_{\gamma\delta}. \tag{5.92}$$

From Eq. (5.92), it is clear that

$$\begin{aligned} P_0 e'_\alpha &= P_0 e''_\alpha, \\ P_7 e'_\alpha &= P_7 e''_\alpha, \end{aligned} \tag{5.93}$$

where, as in Eq. (5.77),

$$e''_\alpha = \sum_{\beta=1}^{3} \mathbf{a}_{\beta\alpha} e_\beta, \tag{5.94}$$

where the $\mathbf{a}_{\alpha\beta}$ are elements of U(3). Again, it is the more highly structured functions $\{e'_\alpha\}$ which satisfy the equivalence relations

$$(e_\beta e_\alpha)_0^+ = \sum_{\delta=1}^{3} \epsilon_{\beta\alpha\delta} e_\delta, \tag{5.95}$$

and not the $\{e''_\alpha\}$. In terms of the $\{e''_\alpha\}$, Eq. (5.92) can be written, in a form similar to that of Eq. (5.78), as

$$e'_\alpha = \hat{P}_0 e''_\alpha + e''_\alpha \hat{P}_0 + \sum_{\beta\delta=1}^{3} \epsilon_{\beta\alpha\delta} e''_\beta (P_0 - P_7)\, e''^*_\delta. \tag{5.96}$$

It is clear, from Eq. (5.90) and the fact that the transformation that brings e_α to e'_α is carried out by a unitary operator which leaves P_0 invariant, that

$$P_0 e'_\alpha e'_\beta = P_0 \epsilon_{\alpha\beta\gamma} e'_\gamma, \tag{5.97}$$

i.e., that the automorphism acts like SU(3) in the ideal generated by P_0. To obtain a little more insight into the mechanisms involved, we use Eq. (5.96) to obtain

$$\hat{P}_0 e'_\alpha e'_\beta = \hat{P}_0 e''_\alpha \sum_{\sigma\gamma=1}^{3} \epsilon_{\sigma\beta\gamma} e''_\sigma (P_0 - P_7)\, e''^*_\gamma. \tag{5.98}$$

Now,

$$\begin{aligned} \hat{P}_0 e''_\alpha e''_\sigma (P_0 - P_7) &= \hat{P}_0 \sum_{\beta\gamma=1}^{3} \mathbf{a}_{\beta\alpha} e_\beta \mathbf{a}_{\gamma\sigma} e_\gamma (P_0 - P_7) \\ &= \sum_{\beta\gamma=1}^{3} \mathbf{a}_{\beta\alpha}\mathbf{a}^*_{\gamma\sigma} P_0 e_\beta e_\gamma (P_0 - P_7) \\ &= -\delta_{\alpha\sigma}(P_0 - P_7), \end{aligned} \tag{5.99}$$

where we have used the U(3) relation Eq. (5.88). Hence,

$$\begin{aligned} \hat{P}_0 e'_\alpha e'_\beta &= \epsilon_{\alpha\beta\gamma}(P_0 - P_7)\, e''_\gamma \\ &= \epsilon_{\alpha\beta\gamma}(P_0 - P_7)\, e'_\gamma. \end{aligned} \tag{5.100}$$

As in Eq. (5.96), the occurrence of $P_0 - P_7$ is necessary due to the commutation properties of the $\{e_\alpha\}$ (and $\{e'_\alpha\}$) with e_7. The product of e_α's does not generate multiplication rules through equivalence relations defined on the single nonminimal ideal generated by $P_0 + P_7$. However, projecting out the minimal ideal generated by P_0, one obtains Eq. (5.97).

The automorphisms induced by U_0 on wavefunctions of the form given in Eq. (5.36) result in the following,

$$U_0 f \hat{P}_0 U_0^* = \sum_k e'_k P_0 \psi_0^k, \tag{5.101}$$

and for $\alpha \neq 0$,

$$U_0 f \hat{P}_\alpha U_0^* = \sum_{k,\alpha=1}^{3} e'_k P_0 \psi_\alpha^k e'_\alpha$$

$$= \sum_{k,\alpha=1}^{3} e'_k P_0 \psi_\alpha^k \mathbf{a}_{\beta\alpha} e_\beta. \tag{5.102}$$

Ignoring the transformation of the k indices for our present purposes (one could alternatively consider wavefunctions $f_k = P_0 e^*_k f$ for which this transformation is trivial), we see that the automorphisms generated by U_0 coincide (on the right) with the action of the generalized phase transformations induced by $U_0^* \in \mathfrak{A}$. As we have seen, the effect of this transformation in the ideal associated with P_0 is that of SU(3), and is the intersection of the U(4) from $\mathfrak{A}$ with the G_2 acting on this ideal due to the automorphisms [Eq. (5.7)] of C_7 which leave P_0 invariant.

VI. TENSOR PRODUCTS

The construction of tensor product spaces is essential for the treatment of the many-body problem and for a constructive approach to field theory. In particular, in a theory of leptons and quarks, the one particle Hilbert space is supposed to contain an "unobservable" sector, corresponding to the quark states. These states play a role in the construction of two- and three-body states corresponding to mesons and nucleons, when they are combined in such a way that the combination lie in the observable parts of the two- or three-body space.

A tensor product useful for the purposes mentioned above should have the following properties:

(a) The algebraic structure of the tensor product space is the same as that of the original Hilbert space;

(b) it is well-balanced,[33] i.e., it is linear in each factor (up to automorphisms);

(c) for sufficiently well-behaved linear operators acting on each factor, there exists a linear operator on the tensor product with equivalent action.

The first of these is required in order that the definition of the structure of the Hilbert space, including the field or algebra over which it is defined (containing the scalar products), be the same for its tensor products. The definitions of orthogonality, linear manifolds, transition probabilities, etc., are therefore the same for one-particle, two-particle, ··· states and hence a Fock space can be constructed. The second requirement asserts that the tensor product of linear manifolds is a linear manifold in the tensor product space, and the third requirement corresponds to the consistency of linear mappings. These last two properties are also essential for the construction of a Fock space. The difficulties associated with achieving a tensor product of this type with Hilbert spaces over quaternions has been discussed by Jauch, Schiminovich, and Speiser.[17]

We have defined four types of Hilbert space in the framework of the vector space over C_7. These spaces have scalar products belonging, respectively, to $\mathbb{R}$, $\mathbb{C}(1,e_7)$, $\mathfrak{A}$, and C_7, and the closed linear manifolds associated with each of them are closed over the corresponding algebras. The requirements (a), (b), and (c) may therefore be imposed on tensor products appropriate to each case, and for operators linear over the corresponding algebra. They can be satisfied trivially for tensor products of $\mathscr{H}_{\mathbb{R}}$, and we have constructed solutions for tensor products of $\mathscr{H}_{\mathbb{C}}$, valid for operators linear over $\mathbb{C}(1,e_7)$. We shall show that there is no well-balanced solution for the tensor product of the spaces $\mathscr{H}_{\mathfrak{A}}$ or $\mathscr{H}_{C_7}$, representing the action of operators linear over the gauge algebra or over C_7.

A linear manifold closed over the gauge algebra $\mathfrak{A}$ is a gauge invariant object. The negative result that we have cited above implies that there is no closed linear manifold in a tensor product space which corresponds to the direct product of two such gauge invariant manifolds. On the other hand, since the algebraic structure of the Hilbert space is preserved under the tensor products which take products of complex linear manifolds into complex linear manifolds, such product manifolds can always be extended to their closure under the gauge algebra $\mathfrak{A}$ in this tensor product space; i.e., the actions of the gauge group can be divided out only after the tensor products are carried out in a nongauge invariant way. This situation is somewhat reminiscent of the structure of the Fock space for the electromagnetic field, for which annihilation–creation operators are constructed in some specified gauge.

According to condition (a), the tensor product of vectors f, g, belonging respectively to two (isomorphic) Hilbert spaces, has the same algebraic structure as that of a vector in one of these spaces, i.e., in $(\mathscr{H} \otimes \mathscr{H})_*$,

$$f \otimes g = \sum_{ij=0}^{7} (f \otimes g)_{ij} \rho_{ij}. \tag{6.1}$$

Real linearity (the reals are a subalgebra of every algebra that we shall consider) then requires that

$$(f \otimes g)_{ij} = \sum_{klmn} T_{ij}(kl|mn) f_{kl} g_{mn}, \tag{6.2}$$

where the tensor product of real functions is defined in the usual way (direct product). Since we shall require property (c) as well, there is no loss of generality in taking the set of coefficients $\{T_{ij}(kl\,mn)\}$ to be real (and not real linear operator) valued. Although the existence of tensor products for $\mathscr{H}_{\mathbb{R}}$ follows trivially from the existence of tensor products for $\mathscr{H}_{\mathbb{C}}$, some points can be clarified more easily by first demonstrating:

Statement 9: There exist tensor products, satisfying the properties (a), (b), and (c), for Hilbert spaces of type $\mathscr{H}_{\mathbb{R}}$; in this case, linearity refers to the real subalgebra $\mathbb{R}$ of C_7.

Since $f_{ij} \to f_{ij}\lambda$ when $f \to f\lambda$, and $\lambda \in \mathbb{R}$, condition (b) is trivially satisfied. Using Eq. (3.3) to represent the action of the general operator linear over $\mathbb{R}$, requirement (c), that is, that there exists an operator $A^{(12)}$ such that

$$A^{(12)}(f \otimes g)=A^{(1)}f \otimes A^{(2)}g \tag{6.3}$$

is represented by

$$\begin{aligned}(A^{(1)}f \otimes A^{(2)}g)_{ij} &= \sum_{\substack{klmn\\ pq,p'q'}} T_{ij}(kl|mn)(\mathscr{A}^{(1)}_{kl,pq}f_{pq})(\mathscr{A}^{(2)}_{mn,p'q'}g_{p'q'})\\ &= \sum_{\substack{kl,pq\\ p'q'}} \mathscr{A}^{(12)}_{ij,k}T_{kl}(pq|p'q')\, f_{pq}g_{p'q'},\end{aligned} \tag{6.4}$$

where the action of $\mathscr{A}^{(12)}$ on the real tensor product $f_{pq}g_{p'q'}$ is defined by this equation. Since we are not interested at present in questions of domains and closure, we shall assume that $A^{(1)}$ and $A^{(2)}$ are bounded operators in their respective spaces. Hence Eq. (6.4) must be valid for every f,g, and therefore

$$\sum_{kl} \mathscr{A}^{(12)}_{ijkl}T_{kl}(pq|p'q') = \sum_{kl,mn} T_{ij}(kl|mn)\ \mathscr{A}^{(1)}_{kl,pq}\ \mathscr{A}^{(2)}_{mn,p'q'} \tag{6.5}$$

must define the operator $A^{(12)}$. The coefficients T_{kl} can be chosen with a property analogous to that of Clebsch–Gordan coefficients, with all explicit indices corresponding to magnetic quantum numbers, e.g.,

$$\sum_{pq,p',q'} T_{kl}(pq|p'q')\,T_{k'l'}(pq|p'q') = \delta_{kk'}\delta_{ll'}; \tag{6.6}$$

in this case, Eq. (6.5) can be solved for the operators $\mathscr{A}^{(12)}_{ij,kl}$. We now turn to the proof of:

Statement 10: *There exist tensor products, satisfying the properties (a), (b) and (c), for Hilbert spaces of type $\mathscr{H}_{\mathbb{C}}$; in this case, linearity refers to the complex subalgebra $\mathbb{C}(1,e_7)$ of C_7.*

To work with complex linearity, it will be convenient to introduce the complex-valued elements defined in Eq. (2.38) and, furthermore, to define

$$\begin{aligned}\Xi^k_0 &= (f \otimes g)_{k0} - (f \otimes g)_{k7}e_7,\\ \Xi^k_\alpha &= -(f \otimes g)_{k\alpha} - (f \otimes g)_{k,\alpha+3}e_7,\end{aligned} \tag{6.7}$$

and, for the coefficients,

$$\begin{aligned}\Gamma^k_0 &= T_{k0} - T_{k7}e_7,\\ \Gamma^k_\alpha &= -T_{k\alpha} - T_{k,\alpha+3}e_7.\end{aligned} \tag{6.8}$$

Substituting these expressions in Eq. (6.1), we obtain

$$\begin{aligned}\Xi^i_0 = \sum_{k,m} \{&\tfrac14\psi^k_0\chi^m_0[\Gamma^i_0(k,0|m,0) + e_7\Gamma^i_0(k\,7|m0) - \Gamma^i_0(k\,7|m7) + e_7\Gamma^i_0(k\,0|m7)]\\
&+ \tfrac14\psi^{k*}_0\chi^{m*}_0[\Gamma^i_0(k,0|m,0) - e_7\Gamma^i_0(k\,7|m0) - \Gamma^i_0(k\,7|m7) - e_7\Gamma^i_0(k\,0|m7)]\\
&+ \tfrac14\psi^k_0\chi^{m*}_0[\Gamma^i_0(k,0|m,0) + e_7\Gamma^i_0(k\,7|m0) + \Gamma^i_0(k\,7|m7) - e_7\Gamma^i_0(k\,0|m7)]\\
&+ \tfrac14\psi^{k*}_0\chi^m_0[\Gamma^i_0(k,0|m,0) - e_7\Gamma^i_0(k\,7|m0) + \Gamma^i_0(k\,7|m7) + e_7\Gamma^i_0(k\,0|m7)]\}\\
+ \sum_{k,m,\alpha} \{&\tfrac14\psi^k_0\chi^m_\alpha[-\Gamma^i_0(k\,0|m\alpha) + e_7\Gamma^i_0(k\,0|m,\alpha+3) - e_7\Gamma^i_0(k\,7|m\alpha) - \Gamma^i_0(k\,7|m,\alpha+3)]\\
&+ \tfrac14\psi^{k*}_0\chi^{m*}_\alpha[-\Gamma^i_0(k\,0|m\alpha) - e_7\Gamma^i_0(k\,0|m,\alpha+3) + e_7\Gamma^i_0(k\,7|m\alpha) - \Gamma^i_0(k\,7|m,\alpha+3)]\\
&+ \tfrac14\psi^k_0\chi^{m*}_\alpha[-\Gamma^i_0(k\,0|m\alpha) - e_7\Gamma^i_0(k\,0|m,\alpha+3) - e_7\Gamma^i_0(k\,7|m\alpha) + \Gamma^i_0(k\,7|m,\alpha+3)]\\
&+ \tfrac14\psi^{k*}_0\chi^m_\alpha[-\Gamma^i_0(k\,0|m\alpha) + e_7\Gamma^i_0(k\,0|m,\alpha+3) + e_7\Gamma^i_0(k\,7|m\alpha) + \Gamma^i_0(k\,7|m,\alpha+3)]\\
&+ \tfrac14\psi^k_\alpha\chi^m_0[-\Gamma^i_0(k\alpha|m0) + e_7\Gamma^i_0(k,\alpha+3|m0) - e_7\Gamma^i_0(k\alpha|m7) - \Gamma^i_0(k,\alpha+3|m7)]\\
&+ \tfrac14\psi^{k*}_\alpha\chi^{m*}_0[-\Gamma^i_0(k\alpha|m0) - e_7\Gamma^i_0(k,\alpha+3|m0) + e_7\Gamma^i_0(k\alpha|m7) - \Gamma^i_0(k,\alpha+3|m7)]\\
&+ \tfrac14\psi^k_\alpha\chi^{m*}_0[-\Gamma^i_0(k\alpha|m0) + e_7\Gamma^i_0(k,\alpha+3|m0) + e_7\Gamma^i_0(k\alpha|m7) + \Gamma^i_0(k,\alpha+3|m7)]\\
&+ \tfrac14\psi^{k*}_\alpha\chi^m_0[-\Gamma^i_0(k\alpha|m0) - e_7\Gamma^i_0(k,\alpha+3|m0) - e_7\Gamma^i_0(k\alpha|m7) + \Gamma^i_0(k,\alpha+3|m7)]\}\\
+ \sum_{k,m,\alpha,\beta} \{&\tfrac14\psi^k_\alpha\chi^m_\beta[\Gamma^i_0(k\alpha|m\beta) - e_7\,\Gamma^i_0(k,\alpha+3|m\beta) - e_7\Gamma^i_0(k\alpha|m,\beta+3) - \Gamma^i_0(k,\alpha+3|m\beta+3)]\end{aligned}$$

$$+\tfrac{1}{4}\psi_\alpha^{k*}\chi_\beta^{m*}[\Gamma_0^i(k\alpha|m\beta)+e_7\,\Gamma_0^i(k,\alpha+3|m\beta)+e_7\,\Gamma_0^i(k\alpha|m,\beta+3)-\Gamma_0^i(k,\alpha+3|m,\beta+3)]$$

$$+\tfrac{1}{4}\psi_\alpha^{k}\chi_\beta^{m*}[\Gamma_0^i(k\alpha|m\beta)-e_7\Gamma_0^i(k,\alpha+3|m\beta)+e_7\,\Gamma_0^i(k\alpha|m,\beta+3)+\Gamma_0^i(k,\alpha+3|m,\beta+3)]$$

$$+\tfrac{1}{4}\psi_\alpha^{k*}\chi_\beta^{m}[\Gamma_0^i(k\alpha|m\beta)+e_7\,\Gamma_0^i(k,\alpha+3|m\beta)-e_7\,\Gamma_0^i(k\alpha|m,\beta+3)+\Gamma_0^i(k,\alpha+3|m,\beta+3)]\}. \tag{6.9}$$

The part of the direct product in the "nonobservable" space Ξ_γ^i, is given by Eq. (6.9) with Γ_0^i replaced by Γ_γ^i, since these are distinguished only by making linear combinations with respect to subscript indices of the coefficients T_{kl}. Our purpose in displaying the structure of Eq. (6.9) is to demonstrate that the coefficients of $\psi_0^k\chi_0^m$, $\psi_0^{k*}\chi_0^{m*}$,$\cdots$ are linearly independent, and any subset may be chosen to vanish. We shall require, to satisfy property (b), first of all, that

$$\Xi(f\otimes(gz))=(\Xi(f\otimes g))z \tag{6.10}$$

and

$$\Xi((fz)\otimes g)=\Xi(f\otimes g)z, \tag{6.11}$$

for every $z\in\mathbb{C}(1,e_7)$. Then, according to Eq. (2.4),

$$\Xi_0^i\to\Xi_0^i z,\quad \Xi_\gamma^i\to\Xi_\gamma^i z^*. \tag{6.12}$$

Equations (6.10) and (6.12) require that the coefficients of terms including the factors χ_0^{m*}, χ_α^m vanish in Eqs. (6.9) and that the coefficients of the terms with factors χ_0^m, χ_α^{m*} vanish in the corresponding equations for Ξ_γ^i. Equations (6.11) and (6.12) require that the coefficients of terms including the factors ψ_0^{k*}, ψ_α^k vanish in Eqs. (6.9), and that coefficients of the terms with factors ψ_0^k, ψ_α^{k*} vanish in the corresponding equations for Σ_γ^i. These conditions are satisfied only if the following relations are valid:

$$\Gamma_0^i(k\,0|m0)=e_7\Gamma_0^i(k\,0|m7),\quad \Gamma_0^i(k\,7|m7)=-e_7\Gamma_0^i(k\,7|m0),$$
$$\Gamma_0^i(k\,0|m\alpha)=e_7\Gamma_0^i(k\,0|m\alpha+3),\quad \Gamma_0^i(k\,7|m,\alpha+3)=-e_7\Gamma_0^i(k\,7|m\alpha),$$
$$\Gamma_0^i(k\alpha|m0)=e_7\Gamma_0^i(k\alpha|m7),\quad \Gamma_0^i(k,\alpha+3|m7)=-e_7\Gamma_0^i(k,\alpha+3|m0),\quad \Gamma_0^i(k\alpha|m\beta)=e_7\,\Gamma_0^i(k\alpha|m\beta+3),$$
$$\Gamma_0^i(k,\alpha+3|m,\beta+3)=-e_7\,\Gamma_0^i(k,\alpha+3|m\beta),\quad \Gamma_\gamma^i(k\,0|m0)=-e_7\Gamma_\gamma^i(k\,0|m7),\quad \Gamma_\gamma^i(k\,7|m7)=e_7\Gamma_\gamma^i(k\,7|m0),$$
$$\Gamma_\gamma^i(k\,0|m\alpha)=-e_7\Gamma_\gamma^i(k\,0|m\alpha+3),\quad \Gamma_\gamma^i(k\,7|m,\alpha+3)=e_7\Gamma_\gamma^i(k\,7|m\alpha),\quad \Gamma_\gamma^i(k\alpha|m0)=-e_7\Gamma_\gamma^i(k\alpha|m7),$$
$$\Gamma_\gamma^i(k,\alpha+3|m7)=e_7\Gamma_\gamma^i(k,\alpha+3|m0),\quad \Gamma_\gamma^i(k\alpha|m\beta)=-e_7\,\Gamma_\gamma^i(k\alpha|m\beta+3),\quad \Gamma_\gamma^i(k,\alpha+3|m,\beta+3)=e_7\,\Gamma_\gamma^i(k,\alpha+3|m\beta),$$
$$\Gamma_0^i(k\,7|m7)=-\Gamma_0^i(k\,0|m0),\quad \Gamma_0^i(k\,0|m\alpha)=-\Gamma_0^i(k\,7|m,\alpha+3),\quad \Gamma_0^i(k\alpha|m0)=-\Gamma_0^i(k,\alpha+3|m7),$$
$$\Gamma_0^i(k\alpha|m\beta)=-\Gamma_0^i(k,\alpha+3|m,\beta+3),\quad \Gamma_\gamma^i(k\,7|m7)=-\Gamma_\gamma^i(k\,0|m0),\quad \Gamma_\gamma^i(k\,0|m\alpha)=-\Gamma_\gamma^i(k\,7|m,\alpha+3),$$
$$\Gamma_\gamma^i(k\alpha|m0)=-\Gamma_\gamma^i(k,\alpha+3|m7),\quad \Gamma_\gamma^i(k\alpha|m\beta)=-\Gamma_\gamma^i(k,\alpha+3|m,\beta+3). \tag{6.13}$$

The first 16 of these relations are adequate for Eq. (6.10) and the remaining eight additional relations are required to satisfy Eq. (6.11) as well. We therefore obtain[34] the following, expressed in terms of the remaining independent coefficients:

$$\Xi_0^i=\sum_{k,m}\psi_0^k\chi_0^m\Gamma_0^i(k0|m0)-\sum_{k,m,\alpha}[\psi_0^k\chi_\alpha^{m*}\Gamma_0^i(k0|m\alpha)+\psi_\alpha^{k*}\chi_0^m\Gamma_0^i(k\alpha|m0)]+\sum_{k,m,\alpha,\beta}\psi_\alpha^{k*}\chi_\beta^{m*}\Gamma_0^i(k\alpha|m\beta), \tag{6.14}$$

$$\Xi_\gamma^i=\sum_{k,m}\psi_0^{k*}\chi_0^{m*}\Gamma_\gamma^i(k0|m0)-\sum_{k,m,\alpha}[\psi_0^{k*}\chi_\alpha^m\Gamma_\gamma^i(k\,0|m\alpha)+\psi_\alpha^k\chi_0^{m*}\Gamma_\gamma^i(k\alpha|m0)]+\sum_{k,m,\alpha,\beta}\psi_\alpha^k\chi_\beta^m\Gamma_\gamma^i(k\alpha|m\beta). \tag{6.15}$$

The structure of this result is evident. What we have done in the course of our demonstration is to make explicit the relations which must be imposed on the independent coefficients of the general direct product in order to assure the validity of property (b). By a different choice of surviving terms in Eqs. (6.9), a tensor product linear in the second factor and antilinear in the first (thus accommodating the automorphism $z\to z^*$), for example, can be obtained.

We shall not discuss here the symmetry properties of the coefficients remaining in Eqs. (6.14) and (6.15) which would correspond to the structure of tensor products with definite quantum statistics, nor shall we discuss the question of associativity.

We shall now show that this tensor product is consistent with property (c). According to Eq. (3.8) for complex linear operators $A_{\mathbb{C}}\,B_{\mathbb{C}}$,

$$(A_{\mathbb{C}}f)^k=\sum_l\left\{A_{00}^{kl}\psi_0^l+\sum_\alpha A_{0\alpha}^{kl}\psi_\alpha^{l*},A_{\gamma 0}^{kl}\psi_0^{l*}+\sum_\beta A_{\gamma\beta}^{kl}\psi_\beta^l\right\},$$

$$(B_{\mathbb{C}}g)^k=\sum_l\left\{B_{00}^{kl}\chi_0^l+\sum_\alpha B_{0\alpha}^{k0}\chi_\alpha^{l*},B_{\gamma 0}^{kl}\chi_0^{l*}+\sum_\beta B_{\gamma\beta}^{kl}\chi_\beta^l\right\},$$

and hence, substituting these relations into Eq. (6.14) and (6.15), we require the existence of a complex linear operator C_C defined on the tensor product space such that

$$(C_C \Xi)^i_0 = \sum_{\substack{k,m\\ l,l'}} \Gamma^i_0(k\,0|m0)\Big(A^{kl}_{00}\psi^l_0 + \sum_\alpha A^{kl}_{0\alpha}\psi^{l*}_\alpha\Big)\Big(B^{ml'}_{00}\chi^{l'}_0 + \sum_\beta B^{ml'}_{0\beta}\chi^{l'*}_\beta\Big)$$

$$- \sum_{\substack{k,m,\alpha\\ l,l'}} \Big[\Gamma^i_0(k\,0|m\alpha)\Big(A^{kl}_{00}\psi^l_0 + \sum_\beta A^{kl}_{0\beta}\psi^{l*}_\beta\Big)\Big(B^{ml'}_{\alpha 0}\chi^{l'*}_0 + \sum_{\beta'} B^{ml'}_{\alpha\beta'}\chi^{l'}_{\beta'}\Big)$$

$$+ \Gamma^i_0(k\alpha|m0)\Big(A^{kl*}_{\alpha 0}\psi^l_0 + \sum_\beta A^{kl*}_{\alpha\beta}\psi^{l*}_\beta\Big)\Big(B^{ml'}_{00}\chi^{l'}_0 + \sum_{\beta'} B^{ml'}_{0\beta'}\chi^{l'*}_{\beta'}\Big)\Big]$$

$$+ \sum_{\substack{k,m,\alpha,\beta\\ l,l'}} \Gamma^i_0(k\alpha|m\beta)\Big(A^{kl*}_{\alpha 0}\psi^l_0 + \sum_{\alpha'} A^{kl*}_{\alpha\alpha'}\psi^{l*}_{\alpha'}\Big)\Big(B^{ml'*}_{\beta 0}\chi^{l'}_0 + \sum_{\alpha''} B^{ml'*}_{\beta\alpha''}\chi^{l'*}_{\alpha''}\Big), \tag{6.16}$$

$$(C_C \Xi)^i_\gamma = \sum_{\substack{k,m\\ l,l'}} \Gamma^i_\gamma(k\,0|m0)\Big(A^{kl*}_{00}\psi^{l*}_0 + \sum_\alpha A^{kl*}_{0\alpha}\psi^l_\alpha\Big)\Big(B^{ml'*}_{00}\chi^{l'*}_0 + \sum_\beta B^{ml'*}_{0\beta}\chi^{l'}_\beta\Big)$$

$$- \sum_{\substack{k,m,\alpha\\ l,l'}} \Big[\Gamma^i_\gamma(k\,0|m\alpha)\Big(A^{kl*}_{00}\psi^{l*}_0 + \sum_\beta A^{kl*}_{0\beta}\psi^l_\beta\Big)\Big(B^{ml'*}_{\alpha 0}\chi^{l'}_0 + \sum_{\beta'} B^{ml'}_{\alpha\beta'*}\chi^{l'*}_{\beta'}\Big)$$

$$+ \Gamma^i_\gamma(k\alpha|m0)\Big(A^{kl}_{\alpha 0}\psi^{l*}_0 + \sum_\beta A^{kl}_{\alpha\beta}\psi^l_\beta\Big)\Big(B^{ml'*}_{00}\chi^{l'*}_0 + \sum_{\beta'} B^{ml'*}_{0\beta'}\chi^{l'}_{\beta'}\Big)\Big]$$

$$+ \sum_{\substack{k,m,\alpha,\beta\\ l,l'}} \Gamma^i_\gamma(k\alpha|m\beta)\Big(A^{kl}_{\alpha 0}\psi^{l*}_0 + \sum_{\alpha'} A^{kl}_{\alpha\alpha'}\psi^l_{\alpha'}\Big)\Big(B^{ml'}_{\beta 0}\chi^{l'*}_0 + \sum_{\alpha''} B^{ml'}_{\beta\alpha''}\chi^{l'}_{\alpha''}\Big). \tag{6.17}$$

On the other hand, the action of C_C on the vector Ξ is also specified by Eq. (3.8), i.e., using Eqs. (6.14) and (6.15) in Eq. (3.8),

$$(C_C \Xi)^i_0 = \sum_{k,m,l} C^{il}_{00}\Big[\Gamma^l_0(k\,0|m0)\psi^k_0\chi^m_0 - \sum_\alpha \{\Gamma^l_0(k\,0|m\alpha)\psi^k_0\chi^{m*}_\alpha + \Gamma^l_0(k\alpha|m0)\psi^{k*}_\alpha\chi^m_0\} + \sum_{\alpha\beta}\Gamma^l_0(k\alpha|m\beta)\psi^{k*}_\alpha\chi^{m*}_\beta\Big]$$

$$+ \sum_{k,m,l,\alpha'} C^{il}_{0\alpha'}\Big[\Gamma^l_{\alpha'}(k\,0|m0)^*\psi^k_0\chi^m_0 - \sum_\alpha \{\Gamma^l_{\alpha'}(k\,0|m\alpha)^*\psi^k_0\chi^{m*}_\alpha + \Gamma^l_{\alpha'}(k\alpha|m0)^*\psi^{k*}_\alpha\chi^m_0\}$$

$$+ \sum_{\alpha\beta}\Gamma^l_{\alpha'}(k\alpha|m\beta)^*\psi^{k*}_\alpha\chi^{m*}_\beta\Big], \tag{6.18}$$

$$(C_C \Xi)^i_\gamma = \sum_{k,m,l} C^{il}_{\gamma 0}\Big[\Gamma^l_0(k0|m0)^*\,\psi^{k*}_0\chi^{m*}_0 - \sum_\alpha \{\Gamma^l_0(k0|m\alpha)\;\psi^{k*}_0\chi^m_\alpha + \Gamma^l_0(k\alpha|m0)\,\psi^k_\alpha\chi^{m*}_0\} + \sum_{\alpha\beta}\Gamma^l_0(k\alpha|m\beta)\psi^k_\alpha\chi^m_\beta\Big]$$

$$+ \sum_{k,m,l,\beta} C^{il}_{\gamma\beta}\Big[\Gamma^l_\beta(k0|m0)\;\psi^{k*}_0\chi^{m*}_0 - \sum_\alpha \{\Gamma^l_\beta(k0|m\alpha)\;\psi^{k*}_0\chi^m_\alpha + \Gamma^l_\beta(k\alpha|m0)\,\psi^k_\alpha\chi^{m*}_0\}$$

$$+ \sum_{\alpha,\beta'}\Gamma^l_\beta(k\alpha|m\beta')\;\psi^k_\alpha\chi^m_{\beta'}\Big]. \tag{6.19}$$

Equating the operators defined by their action on $\psi_0^l\chi_0^{l'}$, $\psi_0^l\chi_\alpha^{l'*}$, $\psi_\alpha^{l*}\chi_0^{l'}$, and $\psi_\alpha^{l*}\chi_\beta^{l'*}$ in Eqs. (6.16) and (6.18), we find

$$\sum_n \{\Gamma_0^n(l0|l'0)\; C_{00}^{in}+\sum_{\alpha'}\Gamma_{\alpha'}^n(l0|l'0)^*\; C_{0\alpha'}^{in}$$

$$=\sum_{k,m}\{\Gamma_0^i(k0|m0)\; A_{00}^{kl}B_{00}^{ml'}-\sum_{\alpha'}\Gamma_0^i(k\alpha'|m0)\,A_{\alpha'0}^{kl*}B_{00}^{ml'}+\sum_{\alpha'\alpha''}\Gamma_0^i(k\alpha'|m\alpha'')\,A_{\alpha'0}^{kl*}B_{\alpha''0}^{ml'*}\},\tag{6.20}$$

$$\Gamma_0^n(l0|l'\alpha)\; C_{00}^{in}+\sum_{\alpha'}\Gamma_{\alpha'}^n(l0|l'\alpha)^*\; C_{0\alpha'}^{in}$$

$$=-\Gamma_0^i(k0|m0)\; A_{00}^{kl}B_{0\alpha}^{ml'}+\sum_{\alpha'}\Gamma_0^i(k\alpha'|m\,0)\,A_{\alpha'0}^{kl*}B_{0\alpha}^{ml'}-\sum_{\alpha'\alpha''}\Gamma_0^i(k\alpha'|m\alpha'')\,A_{\alpha'0}^{kl*}B_{\alpha''\alpha}^{ml'*},\tag{6.21}$$

$$\Gamma_0^n(l\alpha|l'0)\; C_{00}^{in}+\sum_{\alpha'}\Gamma_{\alpha'}^n(l\alpha|l'0)^*\, C_{0\alpha'}^{in}$$

$$=-\Gamma_0^i(k0|m0)\; A_{0\alpha}^{kl}B_{00}^{ml'}+\sum_{\alpha'}\Gamma_0^i(k\alpha'|m0)\,A_{\alpha'\alpha}^{kl*}B_{00}^{ml'}-\sum_{\alpha'\alpha''}\Gamma_0^i(k\alpha'|m\alpha'')\,A_{\alpha'\alpha}^{kl*}B_{\alpha''0}^{ml'*},\tag{6.22}$$

$$\Gamma_0^n(l\alpha|l'\beta)\, C_{00}^{in}+\sum_{\alpha'}\Gamma_{\alpha'}^n(l\alpha|l'\beta)^*C_{0\alpha'}^{in}$$

$$=\Gamma_0^i(k0|m0)\; A_{0\alpha}^{kl}B_{0\beta}^{ml'}-\sum_{\alpha'}\Gamma_0^i(k\alpha'|m0)\,A_{\alpha'\alpha}^{kl*}B_{0\beta}^{ml'}+\sum_{\alpha'\alpha''}\Gamma_0^i(k\alpha'|m\alpha'')\,A_{\alpha'\alpha}^{kl*}B_{\alpha''\beta}^{ml'*}.\tag{6.23}$$

Comparing coefficients of $\psi_0^{l*}\chi_0^{l'*},\cdots$ in Eqs. (6.17) and (6.19) [in $(C_{\mathbb{C}}\Xi)_\gamma^i$], one finds a set of relations with structure similar to that of Eqs. (6.20)–(6.23). In this case, the coefficients Γ_0^n on the left-hand side appear conjugated, but the $\Gamma_{\alpha'}^n$ are not; in place of $C_{0\alpha'}^{in}$, the operators $C_{\gamma 0}^{in}$, $C_{\gamma\alpha'}^{in}$ appear, and on the right hand side, one finds Γ_γ^i in place of Γ_0^i, and the complex-valued representatives of the complex linear operators $A_{\mathbb{C}}$, $B_{\mathbb{C}}$ appear conjugated relative to their structure in Eqs. (6.20)–(6.23). As in Eq. (6.6), we may suppose that the Γ_α^n's are chosen so that, for $\alpha',\alpha'',\gamma,\gamma'=0,1,2,3,$

$$\sum_{l\alpha',l'\alpha''}\Gamma_\gamma^{n*}(l\alpha'|l'\alpha'')\;\Gamma_{\gamma'}^{n'}(l\alpha'|l'\alpha'')=\delta_{\gamma\gamma'}\delta_{nn'}.\tag{6.24}$$

Equations (6.20)–(6.23) and the additional equations obtained from $(C_{\mathbb{C}}\Xi)_\gamma^i$ may then be solved explicitly for the operators $C_{\alpha'\alpha''}^{in}$.

In fact, the conditions of Eq. (6.6) and (6.24) are not necessary; it is only required that the equations be invertable.

A tensor product defined, for example, as (suppressing superscripts)

$$\Xi_0=\psi_0\chi_0-\psi_\alpha\chi_\alpha^*,\quad \Xi_\alpha=-\psi_\alpha\chi_0^*+\psi_0\chi_\alpha,\tag{6.25}$$

obtained in a natural way with the use of octonions by Günaydin,[16] does not satisfy this condition. By following the procedure outlined above, one finds the relations

$$\begin{aligned}
&C_{00}=A_{00}B_{00},\quad 0=-\sum_\alpha A_{\alpha0}B_{\alpha0}^*,\\
&0=A_{00}B_{0\alpha'},\quad C_{0\alpha'}=\sum_\alpha A_{\alpha0}B_{\alpha\alpha'}^*,\\
&-C_{0\alpha'}=A_{0\alpha'}B_{00},\quad 0=\sum_\alpha A_{\alpha\alpha'}B_{\alpha0}^*,\\
&0=A_{0\alpha'}B_{\alpha\alpha''}^*,\quad -C_{00}\delta_{\beta\alpha'}=\sum_\alpha A_{\alpha\beta}B_{\alpha\alpha'}^*
\end{aligned}\tag{6.26}$$

from the alternative forms for $(C_{\mathbb{C}}\Xi)_0$, and

$$\begin{aligned}
&C_{\alpha0}=-A_{\alpha0}B_{00}^*,\quad 0=A_{00}B_{\alpha0},\\
&0=-A_{\alpha0}B_{0\alpha'}^*,\quad C_{\alpha\alpha'}=A_{00}B_{\alpha\alpha'},\\
&-C_{\alpha\alpha'}=A_{\alpha\alpha'}B_{00}^*,\quad 0=A_{0\alpha'}B_{\alpha0},\\
&0=A_{\alpha\alpha'}B_{00}^*,\quad -C_{\alpha0}\delta_{\beta\alpha'}=A_{0\beta}B_{\alpha\alpha'}
\end{aligned}\tag{6.27}$$

from the alternative forms for $(C_{\mathbb{C}}\Xi)_\alpha$. Hence, even if the complex linear operators in question do not induce "quark–lepton" transitions (these appear, in any case, to be necessary in a theory which attempts to unify weak, electromagnetic and strong interaction[18,20,21]), one still finds the conditions

$$A_{00}B_{00}=-\sum_{\alpha} A_{\alpha\alpha'}B_{\alpha\alpha'},$$

for any α', and

$$A_{00}B_{\alpha\alpha'}=-A_{\alpha\alpha'}B^{*}_{00}. \tag{6.28}$$

These conditions cannot be valid for arbitrary operators A,B, and hence condition (c) cannot be satisfied for a tensor product of the type given by Eq. (6.25) (contrary to the assertion made in our earlier study[34]).

We now turn to a consideration of tensor products for the spaces $\mathscr{H}_{\mathfrak{A}}$, and prove

Statement 11: *There is no nontrivial choice of coefficients T_{ij} in the product defined in Eq.* (6.2) *which can be used to construct a tensor product of spaces of type $\mathscr{H}_{\mathfrak{A}}$ (with gauge invariant closed linear manifolds) or $\mathscr{H}_{C}$, with the properties* (*a*), (*b*), *and* (*c*).

Since $e_7 \in \mathfrak{A}$, a well-balanced tensor product of spaces of type $\mathscr{H}_{\mathfrak{A}}$ must be a restriction of the definition given by Eqs. (6.14) and (6.15). We shall require first of all that

$$\Xi(f\otimes g\mathbf{a})=\Xi(f\mathbf{a}\otimes g)=\Xi(f\otimes g)\mathbf{a}, \tag{6.29}$$

where $\mathbf{a}\in\mathfrak{A}$ is of the form given in Eq. (5.6). According to Eqs. (5.37) and (5.38), when $f\to f\mathbf{a}$ ($f=\{\psi^k_0,\psi^k_\alpha\}$),

$$\psi^k_0\to\psi^k_0\mathbf{a}_{00}-\sum_{\alpha'}\psi^{k*}_{\alpha'}\mathbf{a}_{\alpha'0},$$

$$\psi^k_\alpha\to\sum_{\alpha'}\psi^k_{\alpha'}\mathbf{a}^{*}_{\alpha'\alpha}-\psi^{k*}_0\mathbf{a}_{0\alpha}. \tag{6.30}$$

Substituting this transformation into the first two parts of Eq. (6.29), i.e.,

$$\Xi(f\otimes g\ \mathbf{a})^i_0=\Big[\Gamma^i_0(k0|m0)\ \psi^k_0-\sum_{\alpha'}(k\alpha'|m0)\,\psi^{k*}_{\alpha'}\Big]\Big(\chi^m_0\mathbf{a}_{00}-\sum_\gamma\chi^{m*}_\gamma\mathbf{a}_{\gamma0}\Big)$$

$$+\sum_{\alpha'}\Big[\sum_\beta\Gamma^i_0(k\beta|m\ \alpha')\,\psi^{k*}_\beta-\Gamma^i_0(k0|m\alpha')\ \psi^k_0\Big]\Big(\sum_\gamma\chi^{m*}_\gamma\mathbf{a}_{\gamma\alpha'}-\chi^m_0\mathbf{a}^{*}_{0\alpha'}\Big),$$

$$\Xi(f\otimes g\ \mathbf{a})^i_\alpha=\Big[\Gamma^i_\alpha(k0|m0)\ \psi^{k*}_0-\sum_{\alpha'}\Gamma^i_\alpha(k\alpha'|m0)\,\psi^k_{\alpha'}\Big]\Big(\chi^{m*}_0\mathbf{a}^{*}_{00}-\sum_\gamma\chi^m_\gamma\mathbf{a}^{*}_{\gamma_0}\Big)$$

$$+\sum_{\alpha'}\Big[\sum_\beta\Gamma^i_\alpha(k\beta|m\alpha')\ \psi^k_\beta-\Gamma^i_\alpha(k0|m\alpha')\ \psi^{k*}_0\Big]\Big(\sum_\gamma\chi^m_\gamma\mathbf{a}^{*}_{\gamma\alpha'}-\chi^{m*}_0\mathbf{a}_{0\alpha'}\Big) \tag{6.31}$$

and

$$\Xi(f\mathbf{a}\otimes g)^i_0=\Big(\psi^k_0\mathbf{a}_{00}-\sum_\alpha\psi^{k*}_{\alpha'}\mathbf{a}_{\alpha'0}\Big)\Big[\Gamma^i_0(k0|m0)\ \chi^m_0-\sum_\gamma\Gamma^i_0(k0|m\gamma)\ \chi^{m*}_\gamma\Big]$$

$$+\sum_{\alpha'}\Big(\sum_\beta\psi^{k*}_\beta\mathbf{a}_{\beta\alpha'}-\psi^k_0\mathbf{a}^{*}_{0\alpha'}\Big)\Big[\sum_\gamma\Gamma^i_0(k\alpha'|m\gamma)\,\chi^{m*}_\gamma-\Gamma^i_0(k\alpha'|m0)\,\chi^m_0\Big]$$

$$\Xi(f\mathbf{a}\otimes g)^i_\alpha=\Big(\psi^{k*}_0\mathbf{a}^{*}_{00}-\sum_{\alpha'}\psi^k_{\alpha'}\mathbf{a}^{*}_{\alpha'}\Big)\Big[\Gamma^i_\alpha(k0|m0)\ \chi^{m*}_0-\sum_\gamma\Gamma^i_\alpha(k0|m\gamma)\ \chi^m_\gamma\Big]$$

$$+\sum_{\alpha'}\Big(\sum_\beta\psi^k_\beta\mathbf{a}^{*}_{\beta\alpha'}-\psi^{k*}_0\mathbf{a}_{0\alpha'}\Big)\Big[\sum_\gamma\Gamma^i_\alpha(k\alpha'|m\gamma)\,\chi^m_\gamma-\Gamma^i_\alpha(k\alpha'|m0)\,\chi^{m*}_0\Big] \tag{6.32}$$

Comparing the coefficients of $\psi^k_0\chi^{m*}_\gamma$ and $\psi^{k*}_{\alpha'}\chi^m_0$ from Ξ^i_0, and the coefficients of $\psi^{k*}_0\chi^m_\gamma$ and $\psi^k_\alpha\chi^{m*}_0$ from Ξ^i_α, in Eqs. (6.31) and (6.32), we obtain

$$+\mathbf{a}_{\gamma0}\,\Gamma^i_0(k0|m0)+\sum_{\alpha'}\Gamma^i_0(k0|m\alpha')\ \mathbf{a}_{\gamma\alpha'}=+\Gamma^i_0(k0|m\gamma)\ \mathbf{a}_{00}+\sum_{\alpha'}\Gamma^i_0(k\alpha'|m\gamma)\,\mathbf{a}^{*}_{0\alpha'},$$

$$+\mathbf{a}_{00}\Gamma^{i}_{0}(k\alpha'|m0) + \sum_{\gamma} \Gamma^{i}_{0}(k\alpha'|m\gamma)\,\mathbf{a}^{*}_{0\gamma} = +\Gamma^{i}_{0}(k0|m0)\;\mathbf{a}_{\alpha'0} + \sum_{\gamma} \mathbf{a}_{\alpha'\gamma}\Gamma^{i}_{0}(k\gamma|m0),$$

$$+\Gamma^{i}_{\alpha}(k0|m0)\;\mathbf{a}^{*}_{\gamma0} + \sum_{\alpha'} \Gamma^{i}_{\alpha}(k0|m\alpha')\;\mathbf{a}^{*}_{\gamma\alpha'} = +\Gamma^{i}_{\alpha}(k0|m\gamma)\;\mathbf{a}^{*}_{00} + \sum_{\alpha'} \Gamma^{i}_{\alpha}(k\alpha'|m\gamma)\,\mathbf{a}_{0\alpha'},$$

$$+\Gamma^{i}_{\alpha}(k\alpha'|m0)\,\mathbf{a}^{*}_{00} + \sum_{\gamma} \Gamma^{i}_{\alpha}(k\alpha'|m\gamma)\,\mathbf{a}_{0\gamma} = +\Gamma^{i}_{\alpha}(k0|m0)\;\mathbf{a}^{*}_{\alpha'0} + \sum_{\gamma} \Gamma^{i}_{\alpha}(k\gamma|m0)\,\mathbf{a}^{*}_{\alpha'\gamma}. \tag{6.33}$$

Since all of the 16 complex numbers representing $\mathbf{a} \in \mathfrak{A}$ are independent, it follows from the first of these equations that $\Gamma^{i}_{0}(k0|m0) = \Gamma^{i}_{0}(k\alpha|m\beta) = \Gamma^{i}_{0}(k0|m\beta) = 0$, and from the second that $\Gamma^{i}_{0}(k\alpha|m0) = 0$ as well. The same conclusion follows for the Γ^{i}_{α} from the last two equations, and hence the first equality of Eqs. (6.29) admits only the trivial solution for the coefficients of the tensor product [comparing only Eqs. (6.31) with $\Xi\mathbf{a}$, one finds a nontrivial solution corresponding to a tensor product linear only on the right factor; such a solution would not be useful in constructing a Fock space].

Finally, if we admit that the $\{a_{\alpha\beta}\}$ on the right side Eq. (6.33) differ from those on the left by an automorphism, the conclusion is the same (the proof is a little more involved).

Since $\mathfrak{A}$ is a subalgebra of C_7, and we have shown that there is no well-balanced tensor product over $\mathfrak{A}$, there is none over C_7 either. ■

VII. CONCLUSIONS

We have discussed an approach to the construction of a Hilbert space with a gauge group (in the fundamental representation) in the simplest case in which the associative algebraic structure has a restriction to the nonassociative normed division algebra of octonions and its automorphism group to the exceptional Lie group G_2.

Given the structure of the linear manifolds, corresponding to the elements of the lattice of propositions of the associated quantum theory,[15] the corresponding gauge symmetry (generalized phase algebra) is determined by the application of Gleason's theorem to the construction of physical states. One finds that *the generalized phase algebra is the commutant in the full algebra of the subalgebra over which the linear manifolds are closed*. The action of the generalized phase algebra on vectors of the Hilbert space has the same algebraic pattern as that of operators linear over the corresponding subalgebra because these operators must also commute with the subalgebra, in the sense $A\,(fa) = (Af)a$, for a an element of the subalgebra. This commutant relation leads to a kind of duality between the structure of linear manifolds and the corresponding gauge symmetry. We have, in particular studied four types of linear operators, $A_{\mathbb{R}}, A_{\mathbb{C}}, A_{\mathfrak{A}}$, and A_{C_7}, associated with manifolds linear over the corresponding subalgebras. In the form "linear manifolds↔generalized phase algebra," we may express this duality for the cases we have considered here as

$$\mathscr{H}_{\mathbb{R}} \leftrightarrow C_7, \quad \mathscr{H}_{\mathbb{C}} \leftrightarrow \mathfrak{A},$$
$$\mathscr{H}_{\mathfrak{A}} \leftrightarrow \mathbb{C}, \quad \mathscr{H}_{C} \leftrightarrow \mathbb{R}. \tag{7.1}$$

The construction of tensor products was examined for each type of space listed in the relations (7.1). The requirement that the tensor product be well balanced, i.e., that the product of linear manifolds correspond to a linear manifold in the tensor product space, precludes the construction of a consistent tensor product for $\mathscr{H}_{\mathfrak{A}}$ or $\mathscr{H}_{C_7}$. The linear manifolds of $\mathscr{H}_{\mathbb{C}}$ are invariant under complex phase (but not under $\mathfrak{A}$ or C_7), and this is the largest invariance that can be carried along with tensor products. Manifolds invariant under $\mathfrak{A}$ can, of course, be constructed in the image space of $\mathscr{H}_{\mathbb{C}} \otimes \mathscr{H}_{\mathbb{C}} \to \mathscr{H}_{\mathbb{C}}$, since its algebraic structure is the same as that of the constituent spaces.

If the algebra $\mathfrak{A}$ is to function as the gauge degrees of freedom of a physical theory, the self-adjoint operators representing physical observables should be invariant under its action, i.e. these operators should be linear over $\mathfrak{A}$. The representation of the quantum theory of such a system in $\mathscr{H}_{\mathbb{C}}$ then, as pointed out in Statement 3, displays superselection rules for the subspaces $\mathscr{H}_{\alpha}$, $\alpha = 0,1,2,3$. The U(4) from $\mathfrak{A}$ which leaves invariant the expectation values of operators linear over $\mathfrak{A}$ (Statement 6) contains an SU(3) subgroup which coincides with the *automorphisms* of the algebra which leave P_0 and e_7 invariant (Statement 8). As long as this gauge subgroup remains unbroken, the subspaces $\mathscr{H}_{\alpha}$, $\alpha = 1,2,3$, transform coherently among each other, but no measurement of an observable can put their phase relation in evidence. The fact that pure states in this subspace cannot be prepared or detected suggests that some difficulty may arise in the formulation of a second quantized theory in which individual quanta are observable.

The physical manifestation of superselection rules, on the other hand, is usually associated with some dynamical phenomenon, such as a parameter that becomes very large (e.g., the size of a ferromagnet), and our algebraic formulation here makes no reference to dynamical constraints which could lead to the algebraic structure we have used. We should emphasize, however, that the SU(3) of algebraic automorphisms which enforce the preservation of this part of the gauge symmetry is effected in an ideal of the algebra, and in the same ideal, the elements of the C_7 algebra behave as the elements of the nonassociative Cayley or octonian algebra. Günaydin, Piron, and Ruegg[8] have recently shown that the projective geometry associated with the octonian algebra is consistent with the axioms of the quantum theory although it does not satisfy the Desargues theorem and hence cannot be embedded in the larger projective geometry that may be represented by a Hilbert space.[15] As Biedenharn and van Dam[35] have explained, the Moufang projective plane, coordinatized by octonians, has translations but not dilatations, while a Desarguesian projective plane has both. Coordinatizing the Galilean null plane subdynamics [SL(3,R)] with octonians, an embedding into the Poincaré group is then not accessible since the required dilatation is not available.

There are, therefore, indications that the nonassociativity of the octonian algebra is associated, as Gürsey[5] proposed some time ago, with the nonobservability of isolated quarks. The connection between these algebraic constraints, in the context of a unified theory rich enough in structure to support them, and the attempts[36] to achieve confinement through the dynamical properties of non-Abelian gauge fields remains a challenging problem.

ACKNOWLEDGMENTS

The authors wish to express their gratitude to Duke University and Tel Aviv University for their hospitality during our exchange of visits. We wish to thank Professor. F. Gürsey for several discussions during the Fifth International Colloquium on Group Theoretical Methods in Physics at Montreal in July, 1976, and Professor L. Michel for discussions of a preliminary report of this work and for helping us to clarify our thinking on the structure of tensor products during the Sixth Colloquium at Tübingen in July, 1977. One of us (L.H.) would like to thank C. Piron and H. Ruegg for a stimulating exchange of ideas during the early stages of this work, and C. Saclioglu for his explanation of the diquark and leptoquark transitions present in theories of this type. We also wish to thank H. Bacry and Y. Ilamed-Lehrer for helpful conversations.

APPENDIX A: GENERAL FORM OF AN OPERATOR LINEAR OVER THE REALS

In Sec. III, the structure of an operator linear over the complex subalgebra $\mathbb{C}(1,e_7)$ was obtained from the general form of an operator linear over the reals. In this Appendix we shall verify Eq. (3.3).

Let $\{\varphi_n\}$ be a complete orthonormal set in the (separable) Hilbert space $\mathscr{H}_{\mathbb{R}}$, so that for $f\in\mathscr{H}_{\mathbb{R}}$

$$f=\sum_n \varphi_n\lambda_n,\quad \lambda_n\in\mathbb{R}, \tag{A1}$$

and, since

$$\begin{aligned}&\operatorname{tr}(\varphi_n,\varphi_m)=\delta_{nm},\\ &\lambda_n=\operatorname{tr}(\varphi_n,f)\end{aligned} \tag{A2}$$

Using the representation of Eq. (2.25), Eq. (A2) may be written as

$$\lambda_n=\sum_{ij}(\varphi_{nij}f_{ij}). \tag{A3}$$

If A is an operator linear over the reals, and is defined on the φ_n, then

$$Af=\sum_n A(\varphi_n\lambda_n)=\sum_n (A\varphi_n)\lambda_n.$$

Using Eq. (A1) again, we obtain the representation

$$A\varphi_n=\sum_m \varphi_m\alpha_{mn},$$

where $\alpha_{mn}\in\mathbb{R}$. WIth the help of Eq. (A3), we then find

$$\begin{aligned}Af&=\sum_{mn}\varphi_m\alpha_{mn}\lambda_n\\ &=\sum_{ijmn}\varphi_m\alpha_{mn}(\varphi_{nij}f_{ij}),\end{aligned}$$

or

$$(Af)_{ij}=\sum_{klmn}\varphi_{mij}\alpha_{mn}(\varphi_{nkl}f_{kl}). \tag{A4}$$

The linear mapping defined by Eq. (A4) is explicitly of the form of Eq. (3.3).

APPENDIX B: ANTILINEAR STRUCTURE

If $\{\psi_0,\psi_\alpha\}$ represents the vector f in the sense of Eq. (2.39), Eq. (3.8) indicates that a complex linear operator A can bring ψ_α into the "lepton" sector of Af, and ψ_0 into the "quark" sector of Af, but with complex conjugation. This structure is induced by the fact that under the operation $f\to fz$, $\psi_0\to\psi_0 z$, and $\psi_\alpha\to\psi_\alpha z^*$. The antilinear structure of such a space can be exhibited by representing f by the column matrix (we shall use an index convention)

$$f=\begin{pmatrix}\psi_0\\ \psi_\alpha K\end{pmatrix} \tag{B1}$$

where K is the operator of complex conjugation,[22] satisfying

$$K^\dagger=K,\quad K^2=1,\quad KzK=z^*.$$

The "natural" scalar product for vectors of the type (B1) is [with a notation of the type introduced in Eq. (5.61)]

$$\begin{aligned}(f,g)_K&=\int d\sigma\,(\psi_0^*,K\psi_\alpha^*)\begin{pmatrix}\chi_0\\ \chi_\alpha K\end{pmatrix}\\ &=\int d\sigma(\psi_0^*\chi_0+\psi_\alpha\chi_\alpha^*).\end{aligned} \tag{B2}$$

This is exactly the complex scalar product defined in Eq. (2.45). Complex linear operators then have the following structure [see Eq. (3.8)]:

$$A=\begin{pmatrix}A_{00} & A_{0\beta}K\\ A_{\alpha 0}K & A_{\alpha\beta}\end{pmatrix}. \tag{B3}$$

In a similar way, a "natural" tensor product for vectors of type (B1) is

$$f\otimes g=\begin{pmatrix}\psi_0\\ \psi_\alpha K\end{pmatrix}\otimes i\begin{pmatrix}\chi_0\\ \chi_\alpha K\end{pmatrix}, \tag{B4}$$

where i is the imaginary unit. Multiplying out the relation (B4), one obtains

$$f\otimes g=\begin{pmatrix}i(\psi_0\chi_0-\psi_\alpha\chi_\alpha^*)\\ i(\psi_0\chi_\alpha-\psi_\alpha\chi_0^*)K\end{pmatrix}, \tag{B5}$$

corresponding precisely to the tensor product given by Günaydin.[16] The tensor product which satisfies the conditions stated in Sec. VI, however, can be obtained with the help of some matrix algebra in the space defined by Eq. (B1).

Equations (6.14) and (6.15) can be written in matrix form as

$$\Xi^0=(\psi_0,\psi_\alpha^*)\begin{pmatrix}\Gamma_0(0|0) & \Gamma_0(0|\beta)\\ \Gamma_0(\alpha|0) & \Gamma_0(\alpha|\beta)\end{pmatrix}\begin{pmatrix}\chi_0\\ \chi_\beta^*\end{pmatrix}, \tag{B6}$$

$$\Xi^\gamma=(\psi_0^*,\psi_\alpha)\begin{pmatrix}\Gamma_\gamma(0|0) & \Gamma_\gamma(0|\beta)\\ \Gamma_\gamma(\alpha|0) & \Gamma_\gamma(\alpha|\beta)\end{pmatrix}\begin{pmatrix}\chi_0^*\\ \chi_\beta\end{pmatrix}, \tag{B7}$$

It follows from the relation

$$\begin{pmatrix}\psi_0\\ \psi_\alpha^*\end{pmatrix}=\begin{pmatrix}1 & 0\\ 0 & K\end{pmatrix}\begin{pmatrix}\psi_0\\ \psi_\alpha K\end{pmatrix} \tag{B8}$$

that the transpose has the following property:

$$(\psi_0,\psi_\alpha^*)=\left[\begin{pmatrix}1 & 0\\ 0 & K\end{pmatrix}\begin{pmatrix}\psi_0\\ \psi_\alpha K\end{pmatrix}\right]^{\mathrm{T}}$$

$$=(\psi_0,K\psi_\alpha)\begin{pmatrix}1 & 0\\ 0 & K\end{pmatrix}. \tag{B9}$$

It then follows that Eq. (B6) can be written as

$$\Xi_0=(\psi_0,K\psi_\alpha)\begin{pmatrix}1 & 0\\ 0 & K\end{pmatrix}\begin{pmatrix}\Gamma_0(0|0) & \Gamma_0(0|\beta)\\ \Gamma_0(\alpha|0) & \Gamma_0(\alpha|\beta)\end{pmatrix}$$

$$\times\begin{pmatrix}1 & 0\\ 0 & K\end{pmatrix}\begin{pmatrix}\chi_0\\ \chi_\beta K\end{pmatrix}$$

or

$$\Xi^0=(\psi_0,K\psi_\alpha)\begin{pmatrix}\Gamma_0(0|0) & \Gamma_0(0|\beta)K\\ K\Gamma_0(\alpha|0) & K\Gamma_0(\alpha|\beta)K\end{pmatrix}\begin{pmatrix}\chi_0\\ \chi_\beta K\end{pmatrix}. \tag{B10}$$

Since the structure of Eq. (B10) assures complex linearity in both factors of the tensor product in the "lepton" part, the additional conjugation supplied by the quaternion factors imply that Ξ^γ must be essentially the complex conjugate of a form similar to Eq. (B10) [this can be explicitly seen from Eq. (B7)]. In terms of the matrix algebra,

$$\begin{pmatrix}\psi_0^*\\ \psi_\alpha\end{pmatrix}=K\begin{pmatrix}\psi_0\\ \psi_\alpha^*\end{pmatrix}K$$

$$=K\begin{pmatrix}1 & 0\\ 0 & K\end{pmatrix}\begin{pmatrix}\psi_0\\ \psi_\alpha K\end{pmatrix}K$$

$$=\begin{pmatrix}K & 0\\ 0 & 1\end{pmatrix}\begin{pmatrix}\psi_0 K\\ \psi_\alpha\end{pmatrix} \tag{B11}$$

and

$$(\psi_0^*,\psi_\alpha)=(K\psi_0,\psi_\alpha)\begin{pmatrix}K & 0\\ 0 & 1\end{pmatrix}. \tag{B12}$$

With the help of Eqs. (B11) and (B12), Eq. (B7) can be written as

$$\Xi^\gamma=(K\psi_0,\psi_\alpha)\begin{pmatrix}K & 0\\ 0 & 1\end{pmatrix}\begin{pmatrix}\Gamma_\gamma(0|0) & \Gamma_\gamma(0|\beta)\\ \Gamma_\gamma(\alpha|0) & \Gamma_\gamma(\alpha|\beta)\end{pmatrix}$$

$$\times\begin{pmatrix}K & 0\\ 0 & 1\end{pmatrix}\begin{pmatrix}\chi_0 K\\ \chi_\alpha\end{pmatrix}$$

or

$$\Xi^\gamma=(K\psi_0,\psi_\alpha)\begin{pmatrix}K\Gamma_\gamma(0|0)K & K\Gamma_\gamma(0|\beta)\\ \Gamma_\gamma(\alpha|0)K & \Gamma_\gamma(\alpha|\beta)\end{pmatrix}\begin{pmatrix}\chi_0 K\\ \chi_\alpha\end{pmatrix}. \tag{B13}$$

Note that the coupling matrices for the tensor product defined by Eqs. (B10) and (B13) have the form of the general complex linear operator [Eq. (B3)].

[1]C.N. Yang and R.L. Mills, Phys. Rev. **96**, 191 (1954).

[2]D. Gross and F. Wilczek, Phys. Rev. Lett. **30**, 1343 (1973); H. Politzer, *ibid.*, **30**, 1346 (1973); S. Coleman and D.J. Gross. *ibid.* **31**, 851 (1973).

[3]M. Gell-Mann, Acta Phys. Austriaca Suppl. **9**, 733 (1972); H. Fritzsch and M. Gell-Mann, in *Proceedings of the XVI International Conference on High Energy Physics, Chicago-Batavia, Ill.*, 1972, edited by J.D. Jackson and A. Roberts (NAL, Batavia, Ill., 1973), Vol. 2, p. 135; M. Gell-Mann, *ibid.* Vol. 4, p. 333.

[4]S. Weinberg, Phys. Rev. Lett. **19**, 1264 (1967); A. Salam, in *Elementary Particle Physics*, edited by N. Svartholm (Stockholm, 1968). See also S. Weinberg, Rev. Mod. Phys. **46**, 255 (1974); M.K. Gaillard, B.W. Lee, and J.L. Rosner, *ibid.* **47**, 277 (1975).

[5]F. Gürsey, Orbis Scientiae Conference at Coral Gables, Jan. 1976. See also Refs. 16 and 35 of this paper. For earlier suggestions associating internal degrees of freedom with octonions and the exceptional quantum mechanical states, see Refs. 8, 9, 10.

[6]P. Jordan, J. von Neumann and E.P. Wigner, Ann. Math. **35**, 29 (1934).

[7]J. Tits, Proc. Colloq. Utrecht, p. 175 (1962); H. Freudenthal, Advan. Math. **I**, 145 (1965).

[8]They do, however, satisfy the axioms of general quantum theory; M. Günaydin, C. Piron and H. Ruegg, Commun. Math. Phys. **61**, 69 (1978).

[9]A. Pais, Phys. Rev. Lett. **7**, 291 (1961). See also A. Gamba, in *High Energy Physics and Elementary Particles*, edited by A. Salam (IAEA, Vienna, 1965), p. 641; J. Math. Phys. **8**, 775 (1967).

[10]H.H. Goldstine and L.P. Horwitz, Proc. Nat. Acad. Sci. USA **48**, 1134 (1962); Math. Ann. **154**, 1 (1964).

[11]That is, self-adjoint with respect to the real scalar product, and linear over the real field. The associated spectral families are therefore associated with real linear manifolds, and the Hilbert space is isomorphic to a real Hilbert space.

[12]H.H. Goldstine and L.P. Horwitz, Math. Ann. **164**, 291 (1966).

[13]L.P. Horwitz and L.C. Biedenharn, Helv. Phys. Acta **38**, 385 (1965). See also the discussion in M. Jammer, *Conceptual Development of Quantum Mechanics* (McGraw-Hill, New York, 1966), p. 370.

[14]G. Birkhoff and J. von Neumann, Ann. Math. **37**, 823 (1936).

[15]C. Piron, Helv. Phys. Acta **37**, 439 (1964). See also, C. Piron, *Foundations of Quantum Physics*(Benjamin, New York, 1976).

[16]M. Günaydin and F. Gürsey, Lett. Nuovo Cimento **6**, 401 (1973); Phys. Rev. D **9**, 3387 (1974); J. Math. Phys. **14**, 1651 (1973); F. Gürsey, Johns Hopkins Workshop on Current Problems in High Energy Particle Theory, p. 15 (Johns Hopkins University, 1974); M. Günaydin, J. Math. Phys. **17**, 1875 (1976).

[17]The use of the complex field is essential to the construction of tensor products. D. Finkelstein, J.M. Jauch, S. Schiminovich, and D. Speiser, J. Math. Phys. **3**, 207 (1962), have discussed the difficulties associated with the construction of tensor products of Hilbert spaces over a non-Abelian algebra (they specifically study quaternion Hilbert spaces).

[18]C. Saclioglu, Phys. Rev. D **15**, 2267 (1977). See also C. Saclioglu, Phys. Rev. D **17**, 1598 (1978).

[19]The parameter labelling the direction of e_7 is a superselection rule since we do not contemplate observables which change this direction. Other examples of continuous superselection rules are the center of mass momentum as used by H. van Dam and L.C. Biedenharn, Phys. Rev. **14**, 405 (1976); see also L.C. Biedenharn and H. van Dam in *Group Theoretical Methods in Physics*, edited by R.T. Sharp and B. Kolman (Academic, New York, (1977). p. 416, and in a closely analogous way, a unit timelike 4-vector n^μ, introduced in order to describe particles with spin in relativistic quantum theory, by L.P. Horwitz, C. Piron, and F. Reuse, Helv. Physica Acta **48**, 546 (1975). C. Piron, Helv. Physica Acta **42**, 330 (1969), has given a general discussion and other examples. See also G. Emch, Helv. Physica Acta **36**, 739, 770 (1963) and in *Lectures in Theoretical Physics*, edited by W.E. Brittin (Univ. Colorado Press, Boulder, 1965), for a discussion of the choice of a complex unit in a theory with non-Abelian "scalars."

[20]H. Georgi and S. Glashow, Phys. Rev. Lett. **32**, 438 (1974); H. Georgi, H.R. Quinn, and S. Weinberg, Phys. Rev. Lett. **33**, 451 (1974).

[21]M. Gell-Mann, P. Ramond and R. Slansky, Los Alamos Preprint LA-UR-77-2059.

[22]E.P. Wigner, *Group Theory*, (Academic, New York, 1959), p. 334.

[23]A.M. Gleason, J. Math. Mech. **6**, 885 (1957). See C. Piron, Ref. 15, for a simple geometrical proof of the central lemma for this theorem.

[24]See, for example, H. Boerner, *Representations of Groups*(North-Holland, Amsterdam, 1963).

[25]We use the same terminology for symmetric algebraic idempotents as for the operator valued projections discussed in connection with Eq. (2.6). These projection have analogous properties [note, however, that $(f,gP) = (f,g)P \neq P(f,g) = (fP,g)$; the symmetry of P must be defined on a carrier space for C_7, which is also the carrier space for (f,g)].

[26]We shall not discuss here possible interpretations or applications of the additional symmetry implied by the sum over k.

[27]Y. Nambu, lectures given at the International Summer School in High Energy Physics, Erice, Italy, 1972 (unpublished).

[28]Note that although the transitions $q^* \to l$, $l^* \to q$, $l \to l$, $q \to q$ occur in Eq. (3.8) (corresponding to the action of currents of the form l^*l, q^*q, ql), the transitions $q^* \to q$, $l^* \to l$, $q \to l$, $l \to q$ corresponding to currents of the form qq, ll, lq^*) are ruled out by complex linearity. The complex linear operators cannot generate G_2 since diquark operators are present in its generators, but, as stated in the sequel, the U(4) transformations which are a symmetry of the complex scalar product can be induced by operators of complex linear type. This group contains the SU(3) (color) subgroup of G_2.

[29]Ref. 23.

[30]The function $\omega(M)$ has the properties

$0 < \omega(M) < 1$;

$\omega(\phi) = 0, \quad \omega(I) = 1$;

if M is compatible with $N, \omega(M) + \omega(N)$
$= \omega(M \cap N) + \omega(M \cup N)$;

if $\omega(M) = \omega(N) = 1$, then $\omega(M \cap N) = 1$;

if $M \neq \phi$, there exists a state ω such that $\omega(M) \neq 0$.

Two states are different if there exists a proposition M such that $\omega_1(M) \neq \omega_2(M)$. If ω_1 and ω_2 are two different states, then $\omega(M) = \lambda_1\omega_1(M) + \lambda_2\omega_2(M)$, with $\lambda_1, \lambda_2 > 0$ and $\lambda_1 + \lambda_2 = 1$, defines a new state. A state which can be represented in this way with two different states is called a *mixture*; a state which is not a mixture is called *pure*. In the embedding of the propositional system in a Hilbert space, the closed linear manifolds are in correspondence with the propositions. See J.M. Jauch and C. Piron, Helv. Phys. Acta **36**, 827 (1963); J.M. Jauch, *Foundations of Quantum Mechanics* (Addison-Wesley, Reading, Mass., 1968), and Ref. 15 for further discussion of this notion.

[31]Note that P_M is self-adjoint in $H_{\mathbb{C}}$ since $e_7 \in \mathfrak{A}$. To see this, let $P_M f = f_M$
$f = f_M + h_f, \quad \mathrm{tr}((f_M, h_f)\mathbf{a}) = 0 \quad \forall \quad \mathbf{a} \in \mathfrak{A}$.
Then,

$$\begin{aligned}(g, P_M f)_c &= (g, f_M)_c = (g_M + h_g, f_M)_c \\ &= \mathrm{tr}(g_M + h_g, f_M) + e_7\mathrm{tr}((g_M + h_g, f_M)e_7) \\ &= \mathrm{tr}(g_M, f_M) + e_7\mathrm{tr}((g_M, f_M)e_7) = (P_M g, f)_c,\end{aligned}$$

since $\mathrm{tr}(h_g, f_M)e_7) = 0$.

[32]G.W. Mackey, *Mathematical Foundations of Quantum Mechanics* (Benjamin, New York, 1963), p. 75.

[33]We thank Professor L. Michel for a discussion of this point.

[34]This includes the results given in our preliminary report, L.C. Biedenharn and L.P. Horwitz, VI International Colloquium on Group Theoretical Methods in Physics, Tubingen, July 18–22, 1977, Lecture Notes in Physics, edited by P. Kramer and A. Rieckers (Springer, Berlin, 1978), p. 367, as special cases.

[35]L.C. Biedenharn and H. van Dam, Proc. of the International Symposium on Mathematical Physics, Mexico City, Jan. 5–8, 1976, Vol. 1, p. 5.

[36]For example, A. Casher, J. Kogut and L. Susskind, Phys. Rev. Lett. **31**, 792 (1973); K. Wilson Phys. Rev. D **10**, 2445 (1974); J. Kogut and L. Susskind, Phys. Rev. D **9**, 3501 (1974). For early discussions of theories with confinement, see H. Fritzsch and M. Gell-Mann, Proc. of XVI International Conference on High Energy Physics, Vol. 2, p. 135 (1972); H. Fritzsch, M. Gell-Mann, and H. Leutwyler, Phys. Lett. B **74**, 365 (1973); S. Weinberg, Phys. Rev. Lett. **31**, 494 (1973) and Phys. Rev. D **8**, 4482 (1973).

Scattering theory and polynomials orthogonal on the unit circle

J. S. Geronimo[a] and K. M. Case

The Rockefeller University, New York, New York 10021
(Received 28 December 1977)

The techniques of scattering theory are used to investigate polynomials orthogonal on the unit circle. The discrete analog of the Jost function, which has been shown to play an important role in the theory of polynomials orthogonal on a segment of the real line, is defined for this system and its properties are investigated. The relation between the Jost function and the weight function is discussed. The techniques of inverse scattering theory are developed and used to obtain new asymptotic formulas satisfied by the polynomials. A set of sum rules satisfied by the coefficients in the recurrence relaxation is exhibited. Finally, Szegö's theorem on Toeplitz determinants is proved using the recurrence formulas and the Jost function. The techniques of inverse scattering theory are used to find the correction terms.

I. INTRODUCTION

The techniques of scattering theory have been used recently to study the properties of polynomials orthogonal on a segment of the real line.[1] These techniques have formed a unified basis for obtaining information about various properties of orthogonal polynomials. It is natural to ask whether the same techniques can be applied to other orthogonal systems.

In this paper we extend the theory to polynomials orthogonal on the unit circle. Very little in the way of new results are obtained. However, we hope to show (1) that as with polynomials orthogonal on a segment of the real line, the methods of scattering theory form a unified basis for obtaining various properties of polynomials orthogonal on the unit circle, and (2) that these techniques exhibit close parallels between the theory of polynomials orthogonal on the unit circle and those orthogonal on a segment of the real line.

Our program is the following: In Sec. II we define the polynomials and derive the recurrence relations they satisfy. These formulas plus the initial conditions are taken as fundamental. Next (Sec. III) the Jost function, which has been shown to play an important role in various properties of polynomials orthogonal on a segment of the real line,[1] is defined and many of its properties are examined. Since we have started with the recurrence formulas, we must show that the polynomials are orthogonal. This is done in Sec. IV which also contains a formula relating the spectral function to the Jost function. In Sec. V the techniques of inverse scattering theory are developed. These techniques are used (Sec. VI) to develop a new asymptotic formula satisfied by the polynomials. A set of sum rules satisfied by the coefficients in the recurrence relation is also presented in this section. Finally, (Sec. VII) a proof of Szegö's theorem on Toeplitz determinants is given. The proof depends only upon the recurrence formulas and some properties of the Jost function. The techniques of inverse scattering theory are applied to find the correction terms.

---

[a]Present address: Mathematics Department, Georgia Institute of Technology, Atlanta, Georgia 30332

II. PRELIMINARIES

Our study of these polynomials begins with the spectral function and the orthogonality condition. This is done to help motivate the recurrence relation [Eqs. (II.7) and (II.9)] we want to discuss.

Let $\rho(\theta)$ be a bounded, nondecreasing function on $[-\pi,\pi]$ with an infinite number of growth points and with

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} d\rho(\theta)\neq 0. \tag{II.1}$$

We are to construct polynomials[2-4] $\phi(Z,n)$, $n=0,1,2,\cdots$, $Z=e^{i\theta}$, such that:

(i) $\phi(Z,n)$ is a polynomial of precise degree n in which the coefficient of Z^n is real and positive,

(ii)

$$\frac{1}{2\pi}\int_{-\pi}^{\pi}\phi(Z,n)\overline{\phi(Z,m)}\,d\rho(\theta)=\delta(n,m)^{5}$$
$$n,m=0,1,2,\cdots. \tag{II.2}$$

Using standard orthogonalization procedures (see Sec. IV) one finds

$$\phi(Z,n)=[D_{n-1}D_n]^{-1/2}$$
$$\times\begin{vmatrix} C_0 & C_{-1} & \cdot & \cdot & \cdot & C_{-n}\\ C_1 & \cdot & \cdot & \cdot & \cdot & C_{-n+1}\\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot\\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot\\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot\\ C_{n-1} & \cdot & \cdot & \cdot & \cdot & C_{-1}\\ 1 & Z & \cdot & \cdot & \cdot & Z^n \end{vmatrix}, \tag{II.3}$$

where

$$C_n=\frac{1}{2\pi}\int_{-\pi}^{\pi}e^{-in\theta}d\rho(\theta),\quad n=0,\pm1,\pm2,\cdots \tag{II.4}$$

 0022-2488/79/020299-12$01.00

and

$$D_n=\begin{vmatrix} C_0 & C_{-1} & C_{-2} & \cdot & \cdot & \cdot & C_{-n} \\ C_1 & C_0 & C_{-1} & \cdot & \cdot & \cdot & C_{-n+1} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ C_n & C_{n-1} & C_{n-2} & \cdot & \cdot & \cdot & C_0 \end{vmatrix}, \tag{II.5}$$

$n=1,2,3,,\cdots$.

[Equation (II.3) also holds for $n=0$ provided one defines $D_{-1}=1$.]

The coefficients of Z^n in $\phi(Z,n)$ can be shown from Eq. (II.3) to be

$$K(n)=\left(\frac{D_{n-1}}{D_N}\right)^{1/2}, \quad n=0,1,2,\cdots. \tag{II.6}$$

The theory of positive Hermitian forms tells us that the D_n are positive (see Ref. 2).

From the spectral function $\rho(\theta)$ and the orthogonality condition one can construct the following set of recurrence relations[3,4]:

$$\phi(Z,n+1)=\frac{K(n+1)}{K(n)}Z\phi(Z,n) + \frac{\alpha(n+1)}{K(n)}Z^n\bar{\phi}(1/Z,n), \quad n\geqslant 0 \tag{II.7}$$

and

$$\bar{\phi}(1/Z,n+1)=\frac{K(n+1)}{K(n)}\frac{1}{Z}\bar{\phi}(1/Z,n) + \frac{\overline{\alpha(n+1)}}{K(n)}Z^{-n}\phi(Z,n), \ n\geqslant 0, \tag{II.8}$$

where

$$\alpha(n)=\phi(0,n). \tag{II.9}$$

From Eq. (II.7) and condition (ii) it is not hard to show that the leading coefficients of $\phi(Z,n),\bar{\phi}(1/Z,n+1)$, and

$\phi(Z,n+1)$ satisfy the relation

$$K(n+1)^2-K(n)^2=|\alpha(n+1)|^2. \tag{II.10}$$

It is convenient to view $\phi(Z,n)$ and $\bar{\phi}(1/Z,n)$ as the components[6] of a vector function $\Psi(Z,n)$, where

$$\Psi(Z,n)=\begin{pmatrix}\phi(Z,n)\\ \bar{\phi}(1/Z,n)\end{pmatrix}. \tag{II.11}$$

The recurrence relations now assume the simple form

$$\Psi(Z,n+1)=A(n)\Psi(Z,n) \tag{II.12}$$

with

$$A(n)=a(n)\begin{bmatrix} Z & b(n+1)Z^n \\ \bar{b}(n+1)Z^{-n} & Z^{-1}\end{bmatrix}, \tag{II.13}$$

where

$$a(n)=\frac{K(n+1)}{K(n)} \tag{II.14a}$$

and

$$b(n+1)=\frac{\alpha(n+1)}{K(n+1)}. \tag{II.14b}$$

At times we will find it more useful to work with

$$\phi^*(Z,n)=Z^n\bar{\phi}(1/Z,n) \tag{II.15}$$

instead of $\bar{\phi}(1/Z,n)$. Equations (II.11), (II.12), and (II.13) now become

$$\Psi^*(Z,n)=\begin{pmatrix}\phi(Z,n)\\ \phi^*(Z,n)\end{pmatrix} \tag{II.11'}$$

satisfying

$$\Psi^*(Z,n+1)=A^*(n)\Psi^*(Z,n) \tag{II.12'}$$

with

$$A^*(n)=a(n)\begin{bmatrix} Z & b(n+1)\\ \bar{b}(n+1)Z & 1\end{bmatrix}. \tag{II.13'}$$

One should note, because of Eq. (II.10), the $a(n)$'s and $b(n)$'s are not independent. In fact

$$\frac{1}{a(n)^2}=1-|b(n+1)|^2 \tag{II.16}$$

and since $a(n)^2$ is positive from Eq. (II.14) and condition (i),

$$|b(n)|<1. \tag{II.17}$$

Thus given a sequence of complex numbers $\{b(n)\}$ satisfying Eq. (II.17) one can construct a sequence of polynomials using Eqs. (II.11), (II.12), (II.13), (II.16) and the initial condition

$$\phi(Z,0)=1/\sqrt{C_0}>0. \tag{II.18}$$

We will now take the above equations and the initial condition as the fundamental equations in our discussion of orthogonal polynomials.[4,7-10]

As a first application, let us examine special cases of the following equation,

$$\Psi^{*(1)}(Z,n+1)\begin{pmatrix}0 & -1\\ 1 & 0\end{pmatrix}\Psi^{*(2)}(Z',n+1)$$

$$=\Psi^{*(1)}(Z,n)A^{*(1)T}(n)\begin{pmatrix}0 & -1\\ 1 & 0\end{pmatrix}$$

$$\times A^{*(2)}(n)\Psi^{*(2)}(Z',n)$$

$$= \Psi^{*(1)}(Z,n)a(n)^2\begin{bmatrix} Z & \bar{b}(n+1)Z \\ b(n+1) & 1 \end{bmatrix}\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

$$\times\begin{bmatrix} Z' & b(n+1) \\ \bar{b}(n+1)Z' & 1 \end{bmatrix}\Psi^{*(2)}(Z',n), \tag{II.19}$$

which is

$$= \Psi^{*(1)}(Z,n)\begin{pmatrix} 0 & -Z \\ Z' & 0 \end{pmatrix}\Psi^{*(2)}(Z',n). \tag{II.20}$$

(A^T means a transpose here.) Let us set $Z = Z'$ (Wronskian theorem). Equation (II.20) becomes, using Eq. (II.15),

$$\Psi^{(1)}(Z,n+1)\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}\Psi^{(2)}(Z,n+1)$$

$$= \Psi^{(1)}(Z,n)\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}\Psi^{(2)}(Z,n)$$

$$\equiv W[\Psi^{(1)},\Psi^{(2)}]. \tag{II.21}$$

Thus the Wronskian W is independent of n.

With $\phi^{(1)} = \phi^{(2)}$ (Christoffel–Darboux),[3,4] Eq. (II.20) becomes

$$\phi^*(Z,n+1)\phi(Z',n+1) - \phi(Z,n+1)\phi^*(Z',n+1)$$

$$= Z'\phi^*(Z,n)\phi(Z',n) - Z\phi(Z,n)\phi^*(Z',n). \tag{II.22}$$

For now let us assume that $|Z'| = 1$. Multiply the above equation by $1/Z'^{n+1} = \bar{Z}'^{n+1}$ and divide by $1 - Z\bar{Z}'$. This gives us

$$\frac{\phi^*(Z,n+1)\overline{\phi^*(Z',n+1)} - \phi(Z,n+1)\overline{\phi(Z',n+1)}}{1 - Z\bar{Z}'}$$

$$= \frac{\phi^*(Z,n)\overline{\phi^*(Z,n)} - Z\bar{Z}'\phi(Z,n)\overline{\phi(Z',n)}}{1 - Z\bar{Z}'}.$$

Adding and substracting $\phi(Z,n)\overline{\phi(Z,n)}$ to the above equation, then taking the complex conjugate, yields

$$\frac{\phi^*(Z',n+1)\overline{\phi^*(Z,n+1)} - \phi(Z',n+1)\overline{\phi(Z,n+1)}}{1 - \bar{Z}Z'}$$

$$= \frac{\phi^*(Z',n)\overline{\phi^*(Z,n)} - \phi(Z',n)\overline{\phi(Z,n)}}{1 - \bar{Z}Z'}$$

$$+ \overline{\phi(Z,n)}\phi(Z',n). \tag{II.23}$$

Since the numerator and denominator in the above formula are polynomials in Z', Eq. (II.23) can be continued to $|Z'| \neq 1$ Iteration down yields the Christoffel–Darboux formula,

$$\frac{\phi^*(Z',n+1)\overline{\phi^*(Z,n+1)} - \phi(Z',n+1)\overline{\phi(Z,n+1)}}{1 - \bar{Z}Z'}$$

$$= \sum_{i=0}^{n} \overline{\phi(Z,i)}\phi(Z',i). \tag{II.24}$$

III. JOST FUNCTION

In scattering theory an important role is played by the Jost function. It can be shown that the Jost function is identical to the Fredholm determinant of the radial integral equation for the $l=0$ radial wavefunction.[7] As such, its zeros in the Im$k > 0$ plane give the values of the bound state energies, and its derivative evaluated at these zeros is proportional to the bound state normalization constants.

As analogous function can be defined in the discrete scattering theory as applied to polynomials orthogonal on a segment of the real line and it has been shown that this function plays an important role in the theory of these orthogonal polynomials.[1]

In this section we define the Jost function and describe some of its properties.

Assuming that $\lim_{n\to\infty} b(n) = 0$, Eq. (II.13) becomes for large n,

$$A^0 = \lim_{n\to\infty} A(n) = \begin{bmatrix} Z & 0 \\ 0 & Z^{-1} \end{bmatrix}. \tag{III.1}$$

Proceeding formally, we introduce two auxiliary solutions,

$$\Psi_+(Z,n) = \begin{pmatrix} \phi_+(Z,n) \\ \hat{\phi}_+(Z,n) \end{pmatrix} \tag{III.2}$$

and

$$\Psi_-(Z,n) = \begin{pmatrix} \hat{\phi}_-(Z,n) \\ \phi_-(Z,n) \end{pmatrix}, \tag{III.3}$$

satisfying Eqs. (II.12) and (II.13) where the components of $\Psi_\pm$ are defined by the boundary conditions

$$\lim_{n\to\infty} |\phi_\pm(Z,n) - Z^{\pm n}| = 0, \quad |Z| \leqslant 1 \tag{III.4}$$

and

$$\lim_{n\to\infty} |\hat{\phi}_\pm(Z,n)| = 0, \quad |Z| \leqslant 1. \tag{III.5}$$

This is possible since the components of $\Psi_\pm$ uncouple for large n. From Eq. (II.20)

$$W[\Psi_-,\Psi_+] = 1. \tag{III.6}$$

Thus $\Psi_+(Z,n)$ and $\Psi_-(Z,n)$ are linearly independent and

$$\Psi(Z,n) = f_-(Z)\Psi_+(Z,n) + f_+(Z)\Psi_-(Z,n), \quad |Z| = 1, \tag{III.7}$$

where

$$f_\pm = \pm\Psi(Z,n)\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\Psi_\pm(Z,n) \tag{III.8}$$

which in component form is

$$f_+(Z) = \phi_+(Z,n)\bar{\phi}(1/Z,n) - \hat{\phi}_+(Z,n)\phi(Z,n) \tag{III.9}$$

and

$$f_-(Z) = \phi_-(Z,n)\phi(1/Z,n) - \hat{\phi}_-(Z,n)\bar{\phi}(1/Z,n). \tag{III.10}$$

For $|Z| = 1$ Eqs. (II.13), (II.14), (III.4), and (III.5) show that

$$\bar{\phi}(1/Z,n) = \overline{\phi(Z,n)}, \quad \overline{\phi_+(Z,n)} = \phi_-(Z,n), \tag{III.11}$$

and

$$\overline{\hat{\phi}_+(Z,n)} = \hat{\phi}_-(Z,n).$$

Therefore,

$$\overline{f_+(Z)} = f_-(Z) \tag{III.12}$$

for $|Z| = 1$. Because $f_+(Z)$ is independent of n, it is convenient to evaluate Eq. (III.8) in the limit $n\to\infty$ since there $\Psi_+(Z,n)$ assumes a simple form. In particular

$$f_+(Z)=\lim_{n\to\infty} Z^n\bar{\phi}(1/Z,n)=\lim_{n\to\infty}\phi^*(Z,n). \qquad \text{(III.13)}$$

We shall call $f_+(Z)$ the Jost function for polynomials orthogonal on the unit circle since it will be shown to play the same role in the theory of these polynomials as the analogous function does in scattering theory.

In order to investigate the properties of the Jost function, we will find it convenient at this point to introduce the techniques of Banach algebras. Thus, let A denote the class of functions integrable on $[-\pi,\pi]$ with

$$f(\theta)\approx \sum_{K=-\infty}^{\infty} f(K)\, e^{iK\theta}, \qquad \text{(III.14)}$$

where

$$\Sigma|f(K)|<\infty.$$

A is a Banach algebra[8] with norm

$$\|f\|=\Sigma|f(K)|. \qquad \text{(III.15)}$$

Let A^+ and A^- denote those functions in A which are of the form

$$g(\theta)=\sum_{K=o}^{\infty} g(K)e^{iK\theta} \qquad \text{(III.16)}$$

and

$$h(\theta)=\sum_{K=-\infty}^{0} h(K)\, e^{iK\theta} \qquad \text{(III.17)}$$

respectively. A^+ and A^- are also Banach algebras.

If

$$\sum^{\infty} |b(n)|<\infty, \qquad \text{(III.18)}$$

then it has been shown that $f_+(Z)$ is

(i) analytic inside the unit circle and continuous on it,[4]

(ii) nonzero inside and on the unit circle,[3,4,9,11]

(iii) an element of A^+.[10]

As to the properties of $\phi_+(Z,n)$ and $\hat{\phi}_+(Z,n)$, we show in Appendix A that if $\{b(n)\}$ satisfy Eq. (III.18), then $\phi_+(Z,n)$ and $\hat{\phi}_+(Z,n)$ are analytic inside and continuous on the unit circle and they are elements of A^+. Throughout the rest of this paper, we will assume that $\{b(n)\}$ satisfy Eqs. (III.18) and (II.16).

IV. CONSEQUENCES

Having defined the Jost function and examined some of its properties in the previous section, we now investigate the role it plays in various properties of orthogonal polynomials. In particular, we derive the following:

(i) (orthogonality relation)

$$\frac{1}{2\pi}\int_{-\pi}^{\pi}\phi(Z,n)\overline{(Z,m)}d\rho(\theta)$$
$$=\delta(n,m),\quad Z=e^{i\theta},\quad m,n=0,1,2,\cdots, \qquad \text{(IV.1)}$$

where

$$d\rho(\theta)=\sigma(\theta)d\theta\approx\frac{d\theta}{|f_+(Z)|^2}; \qquad \text{(IV.2)}$$

(ii)

$$f_+(Z)=\exp(-1/4\pi)\int_{-\pi}^{\pi}\ln\sigma(\theta)\left(\frac{\exp(i\theta)+Z}{\exp(i\theta)-Z}\right)d\theta,$$
$$|Z|<1. \qquad \text{(IV.3)}$$

To obtain the orthogonality relation [condition (i)],[3,4] examine the following integral

$$I=\frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{\phi_+(Z,n)\overline{\phi(Z,m)}}{f_+(Z)}d\theta,\qquad z=e^{i\theta},$$
$$n\geqslant m. \qquad \text{(IV.4)}$$

Solving for $\phi_+(Z,n)$ in Eq. (III.7) and then substituting the result into Eq. (IV.14) using Eqs. (III.11) and (III.12) yields

$$\frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{\phi_+(Z,n)\overline{\phi(Z,m)}}{f_+(Z)}d\theta$$
$$=\frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{\phi(Z,n)\phi(Z,m)}{|f_+(Z)|^2}d\theta$$
$$+\frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{\overline{\hat{\phi}(Z,n)\phi(Z,m)}}{\overline{f_+(Z)}}d\theta,$$
$$z=e^{i\theta},\quad n\geqslant m. \qquad \text{(IV.5)}$$

In order to evaluate these integrals, we need the following limits:

(a) $\lim_{Z\to 0}\phi(Z,n)=b(n)K(0)\prod_{i=1}^{n-1}a(i)=\alpha(n),$

(b) $\lim_{Z\to 0}\bar{\phi}(1/Z,n)=Z^{-n}K(0)\prod_{i=1}^{n-1}a(i)=Z^{-n}K(n),$ (IV.6)

(c) $\lim_{Z\to 0}\phi_+(Z,n)=Z^n\prod_{i=n}^{\infty}a(i)=Z^n\dfrac{K(\infty)}{K(n)}$

(d)

$$\lim_{Z\to 0}\hat{\phi}_+(Z,n)=-Z\bar{b}(n+1)\prod_{i=n}^{\infty}a(i)$$
$$=-Z\frac{K(\infty)}{K(n)}\frac{\bar{\alpha}(n+1)}{K(n+1)},$$

where Eq. (II.14) has been used. These limits are easily obtained using Eqs. (II.11) and (II.12) and the boundary conditions the functions satisfy.

Returning to Eq. (IV.5), take the complex conjugate of the second term on the right-hand side. Now using limits a and b above and the fact $f_+(Z)\neq 0$ for $|Z|\leqslant 1$, it is easy to see that this term is equal to zero. The term on the left-hand side is evaluated using limits b and c and is equal to zero for $n>m$ and one if $n=m$. Combining the above results, one finds

$$\frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{\phi(Z,n)\overline{\phi(Z,m)}}{|f_+(Z)|^2}d\theta=\delta(n,m)$$
$$Z=e^{i\theta},\quad n\geqslant m. \qquad \text{(IV.7)}$$

For $m\geqslant n$ examine the integral

$$I = \frac{1}{2\pi}\int_{-\pi}^{\pi} \frac{\phi_+(Z,m)\overline{\phi(Z,n)}}{f_+(Z)}\, d\theta,$$

$$z=e^{i\theta},\quad m\geqslant n. \qquad \text{(IV.8)}$$

Using the above procedures, we obtain

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} \frac{\phi(Z,m)\overline{\phi(Z,n)}}{|f_+(Z)|^2}\, d\theta = \delta(n,m)$$

$$Z=e^{i\theta},\quad m\geqslant n. \qquad \text{(IV.9)}$$

Taking the complex conjugate of the above equation and combining it with Eq. (IV.7), gives

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} \frac{\phi(Z,n)\overline{\phi(Z,m)}}{|f_+(Z)|^2}\, d\theta = \delta(n,m),$$

$$Z=e^{i\theta}\quad n,m=0,1,2,\cdots. \qquad \text{(IV.10)}$$

This result allows us to identify the weight function with the Jost function in the following manner

$$\sigma(\theta)\approx |f_+(Z)|^{-2}.^{3,4} \qquad \text{(IV.11)}$$

Note that from the above equation, properties III(ii) and III(iii) and the Wiener–Levy theorem[12] $\sigma(\theta)$ is an element of A. [Baxter[10,13] has shown that this is a necessary and sufficient condition on $\sigma(\theta)$.]

Usually in the theory of orthogonal polynomials $\sigma(\theta)$ is given and $f_+(Z)$ must be determined. This can be done using a modification of the Poisson integral formula to give[3,4,14]

$$f_+(Z)=\exp(1/2\pi)\int_{-\pi}^{\pi}\ln\sigma(\theta)\left\{\frac{\exp(i\phi)+Z}{\exp(i\phi)-Z}\right\}d\phi,$$

$$|Z|<1. \qquad \text{(IV.12)}$$

V. INVERSE SCATTERING THEORY

In the previous section the methods of scattering theory have been used to study orthogonal polynomials. In this section we introduce the techniques of inverse scattering theory. Besides being of interest in their own right, the results obtained in this chapter will prove to be very useful in our discussion of the asymptotic properties of orthogonal polynomials (Sec. VI) and Szegö's theorem (Sec. VII).

We begin with the derivation of the discrete analog of the Marchenko equations. Since $\phi_+(Z,n)$ and $\hat{\phi}_+(Z,n)$ are elements of A^+ (see Appendix A), they can be written as

$$\phi_+(Z,n)=\sum_{n'=n}^{\infty} A_1(n,n')Z^{n'} \qquad \text{(V.1)}$$

and

$$\hat{\phi}_+(Z,n)=\sum_{n'=1}^{\infty} A_2(n,n')Z^{n'}, \qquad \text{(V.2)}$$

where Eq. (IV.6) has been used as a guide.

Substituting Eqs. (V.1) and (V.2) into Eq. (III.7) and using Eq. (III.11) yields

$$\frac{\phi(Z,n)}{f_+(Z)}=\sum_{n'=1}^{\infty}\overline{A_2(n,n')}Z^{-n'}$$

$$+\sum_{n'=n}^{\infty}A_1(n,n')S(Z)Z^{n'},\quad |Z|=1 \qquad \text{(V.3)}$$

and

$$\frac{\bar{\phi}(1/Z,n)}{f_+(Z)}=\sum_{n'=n}^{\infty}\overline{A_1(n,n')}Z^{-n'}$$

$$+\sum_{n'=1}^{\infty}A_2(n,n')S(Z)Z^{n'}, \qquad \text{(V.4)}$$

where

$$S(Z)=\frac{f_-(Z)}{f_+(Z)}=\frac{\overline{f_+(Z)}}{f_+(Z)},\quad |Z|=1. \qquad \text{(V.5)}$$

Note from properties III(ii) and III(iii) and the Wiener–Levy theorem, $S(Z)$ is a element of A. Multiplying Eqs. (V.3) and (V.4) by $Z^{m-1}/2\pi i$ and integrating around the unit circle gives

$$0=\overline{A_2(n,m)}+\sum_{n'=n}^{\infty}A_1(n,n')F(n'+m),$$

$$m,n>0 \qquad \text{(V.6)}$$

and

$$\frac{\delta(n,m)}{A_1(n,m)}=\overline{A_1(n,m)}+\sum_{n'=1}^{\infty}A_2(n,n')F(n'+m),$$

$$m\geqslant n\geqslant 1, \qquad \text{(V.7)}$$

where

$$F(n',m)=\oint S(Z)Z^{n'+m-1}\frac{dz}{2\pi i},\quad |Z|=1. \qquad \text{(V.8)}$$

Solving for $A_2(n,m)$ in Eq. (V.6) and substituting the result into the complex conjugate of Eq. (V.7), gives the discrete analogs of the Marchenko equations

$$0=a(n,m)+G(n,m)+\sum_{l=n+1}^{\infty}a(n,l)G(l,m),$$

$$m>n \qquad \text{(V.9)}$$

and

$$\frac{1}{A_1(n,n)^2}=1+G(n,n)+\sum_{l=n+1}^{\infty}a(n,l)G(l,n),$$

$$n=m, \qquad \text{(V.10)}$$

where

$$G(l,m)=-\sum_{n'=1}^{\infty}F(l+n')\overline{F(n'+m)} \qquad \text{(V.11)}$$

and

$$a(n,l)=\frac{A_1(n,l)}{A_1(n,n)}. \qquad \text{(V.12)}$$

Since $S(Z)$ is an element of A, and $\phi_+(Z,n)$ and $\hat{\phi}_+(Z,n)$ are elements of A^+, all manipulations leading to Eqs. (V.9) and (V.10) are justifiable.

Solving Eq. (V.11) for $a(n,m)$, using Cramer's rules yields

$$a(n,m)=\frac{|\;|_{n+m}}{\det[1+G]_{n+1}^{\infty}}, \qquad \text{(V.13)}$$

with

$$\det[1+G]_{n+1}^{\infty}=\begin{vmatrix} 1+G(n+1,n+1) & G(n+1,n+2) & G(n+1,n+3) & \cdots \\ G(n+2,n+1) & 1+G(n+2,n+2) & G(n+2,n+3) & \cdots \\ G(n+3,n+1) & G(n+3,n+2) & 1+G(n+3,n+3) & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{vmatrix}. \quad \text{(V.14)}$$

$|\,|_{n+m}$ is the same except the mth row is replaced by $-G(n,n+1), -G(n,n+2),\cdots$. For example,

$$|\,|_{n+1}=\begin{vmatrix} -G(n,n+1) & -G(n,n+2) & -G(n,n+3) & \cdots \\ G(n+2,n+1) & 1+G(n+2,n+2) & G(n+2,n+3) & \cdots \\ G(n+3,n+1) & G(n+3,n+2) & 1+G(n+3,n+3) & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{vmatrix}. \quad \text{(V.15)}$$

Subsituting these results into Eq. (V.10), yields

$$\frac{1}{A(n,n)^2} = \frac{[1+G(n,n)]\det[1+G]_{n+1}^{\infty}+\Sigma_{l=n+1}^{\infty}|\,|_l G(l,n)}{\det[1+G]_{n+1}^{\infty}}, \quad \text{(V.16)}$$

which is[14,15]

$$\frac{1}{A_1(n,n)^2}=\frac{\det[1+G]_n^{\infty}}{\det[1+G]_{n+1}^{\infty}}. \quad \text{(V.17)}$$

[This equation, Eq. (IV.6) and the fact that $S(Z)$ is an element of A imply that $\det[1+G]_n^{\infty}>0$.]

Consequently,

$$A_1(n,n)=\left(\frac{\det[1+G]_{n+1}^{\infty}}{\det[1+G]_n^{\infty}}\right)^{1/2} \quad \text{(V.18)}$$

and substituting this into Eq. (V.12) using Eq. (V.13) yields

$$A_1(n,l)=\frac{|\,|_{+n+l}}{(\det[1+G]_n^{\infty}\det[1+G]_{n+1}^{\infty})^{1/2}}. \quad \text{(V.19)}$$

Using these results in Eq. (V.11) gives

$$\phi_+(Z,n)=\frac{1}{[\det[1+G]_n^{\infty}\det[1+G]_{n+1}^{\infty}]^{1/2}}\begin{vmatrix} Z^n & Z^{n+1} & Z^{n+2} & \cdots \\ G(n+1,n) & 1+G(n+1,n+1) & G(n+1,n+2) & \cdots \\ G(n+2,n) & G(n+2,n+1) & 1+G(n+2,n+2) & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{vmatrix}. \quad \text{(V.20)}$$

To relate the coefficients in the recurrence relations to the solutions of the Marchenko equation, we begin by writing Eq. (II.12) for ϕ_+ in component form,

$$\phi_+(Z,n+1)=a(n)\{Z\phi_+(Z,n)+b(n+1)Z^n\hat{\phi}_+(Z,n)\} \quad \text{(V.21)}$$

and

$$\hat{\phi}_+(Z,n+1)=a(n)[1/Z\hat{\phi}_+(Z,n) +\bar{b}(n+1)Z^{-n}\phi_+(Z,n)]. \quad \text{(V.22)}$$

Subsituting in Eqs. (V.1) and (V.2), multiplying Eqs. (V.21) and (V.22) by $Z^{-n-2}/2\pi i$ and $Z^{-1}/2\pi i$, respectively, and integrating around the unit circle yields,

$$A_1(n+1,n+1)=a(n)[A_1(n,n)+b(n+1)A_2(n,1)] \quad \text{(V.23)}$$

and

$$0=a(n)A_2(n,1)+a(n)\bar{b}(n+1)A_1(n,n). \quad \text{(V.24)}$$

Therefore,

$$\frac{A_2(n,1)}{A_1(n,n)}=-\bar{b}(n+1). \quad \text{(V.25)}$$

Using the above equation and Eq. (II.15) in Eq. (V.23) gives

$$a(n)=\frac{A_1(n,n)}{A_1(n+1,n+1)}. \quad \text{(V.26)}$$

We now turn our attention to the derivation for another set of equations used in inverse scattering theory, the Gel'fand–Levitan equations. Given a system of orthogonal polynomials $\{\phi^0(Z,n)\}$ with weight $d\rho^0(\theta)$, satisfying Eqs. (II.11), (II.13), and (II.17), we wish to find polynomials orthogonal with respect to the weight $d\rho(\theta)$. Writing

$$\phi(Z,n)=\sum_{i=0}^{n}K(n,i)\phi^0(Z,i), \quad \text{(V.27)}$$

The orthogonality condition

$$\frac{1}{2\pi}\int \phi(Z,n)\phi(Z,m)d\rho(\theta)=\delta(n,m) \tag{V.28}$$

is equivalent to

$$\frac{1}{2\pi}\int \phi(Z,n)\overline{\phi^{0}(Z,m)}\,d\rho(\theta)=\frac{\delta(n,m)}{K(n,n)},\qquad n\geqslant m. \tag{V.29}$$

Substituting Eq. (V.27) into Eq. (V.29), yields the discrete analogs of the Gel'fand–Levitan equations,

$$h(n,m)+q(n,m)+\sum_{l=0}^{n-1}h(n,l)q(l,m)=0,\qquad n>m \tag{V.30}$$

and

$$\frac{1}{K(n,n)^2}=1+q(n,n)+\sum_{l=0}^{n-1}h(n,l)q(l,n), \tag{V.31}$$

where

$$q(m,l)=\frac{1}{2\pi}\int_{-\pi}^{\pi}\phi^{0}(Z,m)\overline{\phi^{0}(Z,l)}\,d\,[\rho(\theta)-\rho^{0}(\theta)] \tag{V.32}$$

and

$$h(n,m)=\frac{K(n,m)}{K(n,n)}. \tag{V.33}$$

Solving Eq. (V.30) using Cramer's rules, yields

$$h(n,m)=\frac{|\;|_m^{n-1}}{\det[1+q]_0^{n-1}}, \tag{V.34}$$

where

$$\det[1+q]_0^{n-1}=\begin{vmatrix} 1+q(0,0) & q(0,1) & q(0,2) & \cdots & q(0,n-1)\\ \cdot & \cdot & \cdot & \cdots & \cdot\\ \cdot & \cdot & \cdot & \cdots & \cdot\\ \cdot & \cdot & \cdot & \cdots & \cdot\\ q(n-1,0) & \cdot & \cdot & \cdots & 1+q(n-1,n-1)\end{vmatrix}. \tag{V.35}$$

$|\;|_m^{n-1}$ is the same as this except the mth row is replaced by $-q(n,0)$, $-q(n,1)$, $-q(n,2)$,.... *For example,*

$$|\;|_1^{n-1}=\begin{vmatrix} -q(n,0) & -q(n,1) & \cdots & -q(n,n-1)\\ q(1,0) & 1+q(1,1) & \cdots & q(1,n-1)\\ \cdot & \cdot & \cdots & \cdot\\ \cdot & \cdot & \cdots & \cdot\\ \cdot & \cdot & \cdots & \cdot\\ q(n-1,0) & \cdot & \cdots & 1+q(n-1,n-1)\end{vmatrix}. \tag{V.36}$$

Subsituting Eq. (V.34) in Eq. (V.31), gives

$$K(n,n)=\frac{(\det[1+q]_0^{n-1})^{1/2}}{\det[1+q]_0^{n}}. \tag{V.37}$$

Using this and Eq. (V.34) in Eq. (V.33), yields

$$K(n,m)=\frac{|\;|_m^{n-1}}{(\det[1+q]_0^{n}\det[1+q]_0^{n-1})^{1/2}}. \tag{V.38}$$

Now substituting Eq. (V.37) and Eq.(V.38) into Eq.(V.27) gives

$$\phi(Z,n)=\frac{1}{(\det[1+q]_0^{n}\det[1+q]_0^{n-1})^{1/2}}\times\begin{vmatrix} 1+q(0,0) & q(0,1) & \cdots & q(0,n)\\ q(1,0) & 1+q(1,1) & \cdots & q(1,n)\\ \cdot & \cdot & \cdots & \cdot\\ \cdot & \cdot & \cdots & \cdot\\ \cdot & \cdot & \cdots & \cdot\\ q(n-1,0) & \cdot & \cdots & q(n-1,n)\\ \phi^{0}(Z,0) & \phi^{0}(Z,1) & \cdots & \phi^{0}(Z,n)\end{vmatrix} \tag{V.39}$$

To show how the $\{K(n,m)\}$ are related to the coefficients of the recurrence formulas, we begin with Eqs. (II.11), (II.12), and (II.13) for $\phi(Z,n+1)$. Multiplying them by $\overline{\phi^{0}(Z,n+1)}$ and integrating with respect to $d\rho^{0}(\theta)$ yields

$$\begin{aligned}\frac{1}{2\pi}\int_{-\pi}^{\pi}\overline{\phi^{0}(Z,n+1)}\,\phi^{0}(Z,n+1)d\rho^{0}(\theta)&=a(n)\frac{1}{2\pi}\int_{-\pi}^{\pi}\phi^{0}(Z,n+1)Z\phi(Z,n)d\rho^{0}(\theta)\\&\quad+a(n)b(n+1)\frac{1}{2\pi}\int_{-\pi}^{\pi}\overline{\phi^{0}(Z,n+1)}\\&\quad\times Z^{n}\overline{\phi(Z,n)}d\rho^{0}(\theta).\end{aligned} \tag{V.40}$$

Using Eqs. (V.27), (V.28), and the recurrence formula for $\phi^{0}(Z,n)$ gives

$$\frac{K(n+1,n+1)}{K(n,n)}=\frac{a'(n)}{a'^{0}(n)}. \tag{V.41}$$

If $a^{0}(n)=1$ for all n, then using Eqs. (V.41) and (V.26) yields the following relation between the solutions of the Marchenko equations and the Gel'fand–Levitan equations

$$a(n)=\frac{K(n+1)}{K(n)}=\frac{K(n+1,n+1)}{K(n,n)}$$

$$=\frac{A_1(n,n)}{A_1(n+1,n+1)}. \qquad (V.42)$$

It is worthwhile to notice that

$$K(n+1,n+1)A_1(n+1,n+1)=K(n,n)A_1(n,n)=C_0. \qquad (V.43)$$

C_0 will be evaluated in Sec. VI. Since $\lim_{n\to\infty}A_1(n,n)=1$,

$$K(\infty)=K(\infty,\infty)=C_0. \qquad (V.44)$$

VI. APPLICATIONS

In this section we investigate the asymptotic form of the orthogonal polynomials using the Marchenko equations discussed in the previous section. We also derive new sum rules satisfied by the coefficients in the recurrence formulas.

A. Asymptotic formulas

We are now in position to find the behavior for large n of the polynomials associated with $\sigma(\theta)$. Starting with Eqs. (III.7)

$$\phi(Z,n)=|f_+(Z)|\,[e^{i\delta(\theta)}\phi_+(Z,n)+e^{-i\delta(\theta)}\hat{\phi}_-(Z,n)],$$
$$Z=e^{i\theta}, \qquad (VI.1)$$

where

$$f_+(Z)=|f_+(Z)|e^{-i\delta(\theta)}, \qquad (VI.2)$$

and using Eqs. (V.1) and (V.2) gives

$$\phi(Z,n)=|f_+(Z)|\Big[e^{i\delta(\theta)}\sum_{m=n}^{\infty}A_1(n,m)Z^m$$
$$+e^{-i\delta(\theta)}\sum_{n'=1}^{\infty}\overline{A_2(n,n')Z^{-n'}}\Big]. \qquad (VI.3)$$

The asymptotic behavior of $\phi(Z,n)$ can be investigated using the Marchenko equations, Eq. (V.9) and (V.10), and perturbation theory. In the first approximation $A_2(n,m)=0$ and $A_1(n,m)=\delta(n,m)$. Formula (VI.3) becomes

$$\phi(Z,n)=\overline{f_+(Z)}Z^n. \qquad (VI.4)$$

In the next approximation

$$A_1(n,n)=1-\frac{G(n,n)}{2},\quad A_1(n,m)=G(n,m), \qquad (VI.5)$$

and

$$A_2(n,m)=\overline{F(n+m)}. \qquad (VI.6)$$

Thus

$$\phi(Z,n)=|f_+(Z)|\Big[\Big(1-\frac{G(n,n)}{2}\Big)Z^ne^{i\delta(\theta)}$$
$$+e^{i\delta(\theta)}\sum_{m=n+1}^{\infty}G(n,m)Z^m-e^{-i\delta(\theta)}$$
$$\times\sum_{n'=1}^{\infty}F(n+m')Z^{-n'}\Big). \qquad (VI.7)$$

Further iteration of Eqs. (V.9) and (V.10) will give successively improved asymptotic formulas.

B. Sum rules

From the explicit form for $f_+(Z)$ in terms of $\sigma(\theta)$, one can obtain a number of identities which the coefficients in the recurrence formula satisfy. First expand $f_+(Z)$ in a power series in the vicinity of the origin

$$f_+(Z)=C_0+C_1Z+C_2Z^2+\cdots \qquad (VI.8)$$

or

$$\frac{f_+(Z)}{C_0}=1+d_1Z+d_2Z^2+\cdots, \qquad (VI.9)$$

where

$$d_i=\frac{C_i}{C_0}. \qquad (VI.10)$$

We wish to express the above coefficients in terms of the coefficents in the recurrence formula. In Ref. 14, a systematic derivation of the sum rules is given. However, here we shall be content to illustrate the first few. Thus,

$$C_0=K(\infty) \qquad (VI.11)$$

$$C_1=K(\infty)\sum_{n=0}^{\infty}\Delta(n), \qquad (VI.12)$$

and

$$C_2=K(\infty)\Big(\sum_{n=0}^{\infty}b'(n)\overline{b'(n+2)}$$
$$+\sum_{n=0}^{\infty}\Delta(n)\sum_{m=n+2}^{\infty}\Delta(m)\Big), \qquad (VI.13)$$

where

$$\Delta(n)=b'(n)\overline{b'(n+1)} \qquad (VI.14)$$

and

$$b'(0)\equiv 1. \qquad (VI.15)$$

To see how the coefficients are related to the moments of $\ln\sigma(\theta)$, we use Eq. (IV.12). Hence,

$$\ln f_+(Z)=(-1/4\pi)\int_{-\pi}^{\pi}\ln\sigma(\theta)\Big(1+2\sum_{i=1}^{\infty}(Z/Z')^i\Big)d\theta,$$
$$Z'=e^{i\theta},\quad |Z|<1. \qquad (VI.16)$$

Substituting in Eq. (VI.8) and using Eq. (VI.11) yields

$$C_0=K(\infty)=\exp(-1/4\pi)\int_{-\pi}^{\pi}\ln\sigma(\theta)d\theta. \qquad (VI.17)$$

[See Eq. (V.44).] Now using Eq. (VI.8) and the above result gives

$$\ln\frac{f_+(Z)}{C_0}=\ln\Big(1+\sum_{i=1}^{\infty}d_iZ^i\Big)=(-1/2\pi)\int_{-\pi}^{\pi}\ln\sigma(\theta)$$
$$\times\sum_{i=1}^{\infty}(Z/Z')^id\theta,\quad Z'=e^{i\theta}. \qquad (VI.18)$$

Expanding the ln and equating coefficients of Z^i yields the desired relations. Thus, for $i = 1$

$$\sum_{n=0}^{\infty} \Delta(n) = d_1 = (-1/2\pi)\int_{-\pi}^{\pi} \ln\sigma(\theta) e^{-i\theta} d\theta, \tag{VI.19}$$

and for $i = 2$,

$$\sum_{n=0}^{\infty} b'(n)\overline{b'(n+2)} - \tfrac{1}{2}\sum_{n=0}^{\infty}\Delta(n)^2 - \sum_{n=0}^{\infty}\Delta(n)\Delta(n+1)$$
$$= d_2 - \frac{d_1^2}{2} = (-1/2\pi)\int_{-\pi}^{\pi}\ln\sigma(\theta) e^{-2i\theta}d\theta. \tag{VI.20}$$

VII. SZEGÖ'S THEOREM

In this section we discuss Szegö's theorem on Teoplitz determinants. This theorem was first proved by Szegö[16] with the assumption that the derivative of the weight function, for a set of polynomials orthogonal on the unit circle, satisfied a Lipschitz condition with some positive exponent. Since then, the theorem has been proved using Banach and Hilbert space techniques with the subsequent weakening of the conditions placed on the weight function.[17-20] In this chapter the theorem is proved solely from the point of view of orthogonal polynomials. Only the recurrence relations and some properties of the Jost function are used. The techniques of inverse scattering theory are applied to the problem to find correction terms to the asymptotic formula.

We wish to prove the following: If

$$\sum_{n=1}^{\infty} |b(n)| < \infty \quad \text{and} \quad \sum_{n=1}^{\infty} n|\gamma(n)|^2 < \infty, \tag{III.18}$$

then

$$\sum_{i=1}^{\infty} i\ln a(i-1)^2 = \sum_{n=1}^{\infty} n|\gamma(n)|^2, \tag{VII.1}$$

where

$$\gamma(n) = \frac{-1}{2\pi}\int_{-\pi}^{\pi}\ln f_+(Z)e^{-in\theta},$$
$$n \geqslant 1, \quad Z = e^{i\theta}. \tag{VII.2}$$

From properties III(i) and III(ii),

$$\frac{-1}{2\pi}\int_{-\pi}^{\pi}\ln f_+(Z)e^{in\theta}d\theta = 0, \quad n\geqslant 1. \tag{VII.3}$$

Taking the complex conjugate of Eq. (VII.3) and adding it to Eq. (VII.2), then using Eq. (IV.2) yields

$$\gamma(n) = \frac{1}{2\pi}\int_{-\pi}^{\pi}\ln\sigma(\theta)e^{-in\theta}d\theta. \tag{VII.4}$$

Using the properties of $f_+(Z)$, it is possible to show that[14]

$$\frac{1}{2\pi}\int_{-\pi}^{\pi}\ln f_+(Z)\ln\overline{f_+(Z)}'Z^{-1}d\theta$$
$$= \sum_{n=1}^{\infty} n|\gamma(n)|^2, \quad Z = e^{i\theta}. \tag{VII.5}$$

Since $\phi^*(Z,n) \to f_+(Z)$, uniformly in norm[10,14]

$$\lim_{n\to\infty}\frac{1}{2\pi}\int_{-\pi}^{\pi}\ln\phi^*(e^{i\theta},n)\ln\overline{\phi^*(e^{i\theta},n)}e^{-i\theta}d\theta$$
$$= \frac{1}{2\pi}\int_{-\pi}^{\pi}\ln f_+(e^{i\theta})\ln\overline{f_+(e^{i\theta})'}e^{-i\theta}d\theta. \tag{VII.6}$$

In order to continue further, we show that

$$\frac{1}{2\pi}\int_{-\pi}^{\pi}\left(\overline{\frac{\phi^{*\prime}(Z,i)}{\phi^*(Z,i)}}\right)Z^{-1}\ln[K(i-1)\phi^*(Z,i-1)]d\theta$$
$$= \frac{1}{2\pi}\int_{-\pi}^{\pi}\left(\overline{\frac{\phi^{*\prime}(Z,i-1)}{\phi^*(Z,i-1)}}\right)$$
$$\times Z^{-1}\ln[K(i-1)\phi^*(Z,i-1)]d\theta. \tag{VII.7}$$

Evaluating the above integrals by means of residues, reduces the above expression to

$$\ln\prod_{l=1}^{i}\frac{\phi^*(Z_{i,l},i-1)}{K(i-1)} = \ln\prod_{j=1}^{i-1}\frac{\phi^*(Z_{i-1,j},i-1)}{K(i-1)}, \tag{VII.8}$$

where $Z_{i,l}$ and $Z_{i-1,j}$ are the zeros of $\phi(Z,i)/K(i)$ and $\phi(Z,i-1)/K(i-1)$, respectively. To prove Eq. (VII.8), we start with Eq. (II.7). Thus,

$$\prod_{j=1}^{i-1}\frac{\phi^*(Z_{(i-1,j)},i-1)}{K(i-1)} = \prod_{j=1}^{i-1}\frac{K(i)}{\alpha(i)}\frac{\phi(Z_{(i-1,j)},i)}{K(i)}. \tag{VII.9}$$

This is so because at $Z = Z_{(i-1,j)}$ Eq. (II.7) becomes

$$\frac{\phi^*(Z_{(i-1,j)},i-1)}{K(i-1)} = \frac{1}{\alpha(i)}\phi(Z_{(i-1,j)},i). \tag{VII.10}$$

Since the constant multiplying the highest power of Z in $\phi(Z,i)/K(i)$ and $\phi(Z,i-1)/K(i-1)$ is equal to one, the term on the right-hand side of Eq. (VII.9) can be rewritten as

$$= \left(\frac{K(i)}{\alpha(i)}\right)^{i-1}\prod_{l=1}^{i}\frac{\phi(Z_{(i,l)},i-1)}{K(i-1)}. \tag{VII.11}$$

But at a zero of $\phi(Z,i)$, we have from recurrence formula (II.7)

$$\frac{\phi(Z_{(i,l)},i-1)}{K(i-1)} = \frac{-\alpha(i)}{K(i)}\frac{1}{Z_{i,l}}\frac{\phi^*(Z_{(i,l)},i-1)}{k(i-1)}. \tag{VII.12}$$

Thus Eq. (VII.11) becomes equal to

$$= \left(\frac{K(i)}{\alpha(i)}\right)^{i-1}\prod_{l=1}^{i}\frac{\alpha(i)}{K(i)}\left(\frac{-1}{Z_{i,l}}\right)\frac{\phi^*(Z_{(i,l)},i-1)}{K(i-1)}. \tag{VII.13}$$

Now the product of the zeros of $\phi(Z,i)/K(i)$ is

$$\prod_{l=1}^{i} -Z_{i,l} = \frac{\alpha(i)}{K(i)}. \tag{VII.14}$$

Therefore, Eq. (VII.13) becomes

$$= \prod_{l=1}^{i}\frac{\phi^*(Z_{(i,l)},i-1)}{K(i-1)}, \tag{VII.15}$$

proving Eq. (VII.8). Note that

$$\frac{1}{2\pi}\int_{-\pi}^{\pi}\left(\overline{\frac{\phi^{*\prime}(Z,1)}{\phi^*(Z,1)}}\right)Z^{-1}\ln K(0)\phi^*(Z,0)d\theta = 0. \tag{VII.16}$$

This is easily seen by taking the complex conjugate and then using the residue theorem. Now the left-hand side of Eq. (VII.6) can be rewritten as

$$= -\frac{1}{2}\lim_{n\to\infty}\int_{-\pi}^{\pi}\left(\frac{\overline{\phi^{*\prime}(Z,n)}}{\phi^*(Z,n)}\right)$$

$$\times Z^{-1}\ln\phi^*(Z,n)^{-1}d\theta \qquad \text{(VII.17)}$$

which is, using Eqs. (VII.7) and (VII.16)

$$= -\lim_{n\to\infty}\sum_{i=1}^{n}(1/2\pi)\int_{-\pi}^{\pi}\left(\frac{\overline{\phi^{*\prime}(Z,i)}}{\phi^*(Z,i)}\right)Z^{-1}$$

$$\times\ln\left(\frac{K(i-1)}{K(i)}\frac{\phi^*(Z,i-1)}{\phi^*(Z,i)}\right)d\theta. \qquad \text{(VII.18)}$$

multiplying Eq. (II.8) by Z^{n+1} then solving for $Z\phi(Z,n)$ and substituting the result into Eq. (II.7) yields,

$$K(n)\phi^*(Z,n)$$
$$=K(n+1)\phi^*(Z,n+1)-\overline{\alpha(n+1)}\phi(Z,n+1) \qquad \text{(VII.19)}$$

or

$$\frac{K(n)\phi^*(Z,n)}{K(n+1)\phi^*(Z,n+1)}$$
$$=\left(1-\frac{\overline{\alpha(n+1)}}{K(n+1)}\frac{\phi(Z,n+1)}{\phi^*(Z,n+1)}\right). \qquad \text{(VII.20)}$$

Thus, Eq. (VII.18) becomes

$$= -\lim_{n\to\infty}\sum_{i=1}^{n}(1/2\pi)\int_{-\pi}^{\pi}\left(\frac{\overline{\phi^{*\prime}(Z,i)}}{\phi^*(Z,i)}\right)Z^{-1}$$

$$\times\ln\left(1-\frac{\overline{\alpha(i)}}{K(i)}\frac{\phi(Z,i)}{\phi^*(Z,i)}\right)dQ. \qquad \text{(VII.21)}$$

The only contributing residue is at $Z=0$. Thus, Eq. [VII.21) becomes

$$=-\lim_{n\to\infty}\sum_{i=1}^{n}i\ln\left(1-\left|\frac{\alpha(i)}{K(i)}\right|^2\right), \qquad \text{(VII.22)}$$

where Eq. (IV.6) has been used. Letting $n\to\infty$ yields

$$\sum_{n=1}^{\infty}i\ln\frac{K(i)^2}{K(i-1)^2}=\sum_{n=1}^{\infty}n|\gamma(n)|^2. \qquad \text{(VII.23)}$$

Using Eq. (II.14) gives us the desired result. Exponentiating each side and using formulas (II.6) and (VI.17) gives the more familiar result

$$\lim_{n\to\infty}\frac{D_n(\sigma)}{G^{n+1}(\sigma)}=\exp\sum_{n=1}^{\infty}n|\gamma(n)|^2, \qquad \text{(VII.24)}$$

where

$$G(\sigma)=K(\infty)^{-2}. \qquad \text{(VII.25)}$$

To find the correction terms to Eq. (VII.24), let us return to Eq. (VII.23).[19,21] Using Eq. (V.42), Eq. (VII.23) can be rewritten as

$$\sum_{i=1}i\ln\frac{K(i)^2}{K(i-1)^2}$$
$$=\sum_{i=n+1}^{\infty}i\ln\frac{A_1(i,i)^2}{A_1(i-1,i-1)}+\sum n|\gamma(n)|^2. \qquad \text{(VII.26)}$$

This becomes, using Eqs. (II.6) and (V.17),

$$\ln D_n(\sigma)+(n+1)\ln[K(n)A_1(n,n)]^2$$
$$=\sum n|\gamma(n)|^2+\ln\det[1+G]_n^\infty.. \qquad \text{(VII.27)}$$

Now using Eqs. (V.43), (VI.17), and (VII.25) yields

$$\ln D_n(\sigma)-(n+1)\ln G(\sigma)$$
$$=\sum_{n=1}^{\infty}n|\gamma(n)|^2+\ln\det[1+G]_n^\infty. \qquad \text{(VII.28)}$$

CONCLUSION

In this article the ideas and techniques of scattering theory have been used to investigate the properties of orthogonal polynomials. We hope that we have shown that the methods of scattering theory provide a unified basis for obtaining many results concerning the theory of orthogonal polynomials. As in the theory of polynomials orthogonal on a segment of the real line, an important role is played by the Jost function. The similarity of the roles of the Jost function in these two systems of polynomials is striking.

Using the techniques of inverse scattering theory and the properties of the Jost function, new asymptotic formulas have been developed and a set of sum rules satisfied by the coefficients in the recurrence formula has been presented.

APPENDIX A: PROPERTIES OF $\phi_+(Z,n)$ AND $\hat{\phi}_+(Z,n)$

To investigate the properties of $\phi_+(Z,n)$ and $\hat{\phi}_+(Z,n)$ define

$$\Psi(Z,n)=\begin{pmatrix}\phi_\alpha(Z,n)\\ \phi_\beta(1/Z,n)\end{pmatrix}, \quad n\geqslant 1, \qquad \text{(A1)}$$

satisfying Eq. (II.13) with

$$\phi_\beta(1/Z,n)=-\bar{\phi}_\alpha(1/Z,n) \qquad \text{(A2)}$$

$$\text{and}\ \phi_\alpha(Z,1)=\frac{a(0)}{K(0)}[Z-b(1)]. \qquad \text{(A3)}$$

Since $\hat{\Psi}(Z,n)$ satisfy the recurrence relations for $n\geqslant 1$, define

$$f_{\pm\alpha}(Z)=\pm W[\hat{\Psi},\Psi_\pm], \quad n\geqslant 1, \qquad \text{(A4)}$$

which in component form is

$$f_{+\alpha}=\phi_+(Z,n)\phi_\beta(1/Z,n)-\hat{\phi}_+(Z,n)\phi_\alpha(Z,n), \quad n\geqslant 1 \qquad \text{(A5)}$$

and

$$f_{-\alpha}=\phi_-(Z,n)\phi_\alpha(Z,n)-\hat{\phi}_-(Z,n)\phi_\beta(1/Z,n), \quad n\geqslant 1. \qquad \text{(A6)}$$

In the limit as $n\to\infty$, Eq. (A5) becomes

$$f_{+\alpha}=\lim_{n\to\infty}Z^n\phi_\beta(1/Z,n)$$

$$= -\lim_{n\to\infty} Z^n\bar{\phi}_\alpha(1/Z,n). \qquad \text{(A7)}$$

To investigate the analytic properties of $f_{+\alpha}$, notice that

$$\left|\frac{Z\phi_\beta(1/Z,n)}{K(1)}\right| = \left|\frac{Z\bar{\phi}_\alpha(1/Z,1)}{K(1)}\right|$$

$$\leqslant \frac{1}{K(0)^2}\,|1+|b(1)||. \qquad \text{(A8)}$$

Substituting this into Eq. (II.8) and using induction arguments, it is easy to see that

$$\left|\frac{Z^n\phi_\beta(1/Z,n)}{K(n)}\right| = \left|\frac{Z^n\bar{\phi}_\alpha(1/Z,n)}{K(n)}\right|$$

$$\leqslant \frac{1}{K(0)^2}\prod_{i=0}^{n-1}(1+|b(i+1)|),$$

$$|Z|\leqslant 1 \qquad \text{(A9)}$$

and using Eq. (A7) that

$$\left|\frac{f_{+\alpha}(Z)}{K(\infty)} - Z^n\frac{\phi_\beta(1/Z,n)}{K(n)}\right|$$
$$\leqslant \frac{1}{K(0)^2}\prod_{i=0}^{\infty}(1+|b(i+1)|)\Big(\prod_{j=n}^{\infty}(1+|b(j+1)|)\Big),$$
$$|Z|\leqslant 1 \qquad \text{(A10)}$$

Thus, $f_{+\alpha}(Z)$ is the uniform limit of a sequence of polynomials for $|Z|\leqslant 1$ and, therefore, is analytic inside the unit circle and continuous on it. Notice that using the above arguments

$$\left\|\frac{f_{+\alpha}(Z)}{K(\infty)} - Z^n\frac{\phi_\beta(1/Z,n)}{K(n)}\right\|$$
$$\leqslant \frac{1}{K(0)^2}\prod_{i=0}^{\infty}(1+|b(i+1)|)$$

$$\times\Big[\prod_{j=n}^{\infty}(1+|b(j+1)|)-1\Big],$$

$$Z=e^{i\theta}, \qquad \text{(A11)}$$

where $\|f\|$ is the norm defined in Sec. II. This implies that $f_{+\alpha}(e^{i\theta})$ is an element of A^+.

Multiplying Eq. (III.9) by $\phi_\alpha(Z,n)$ and Eq. (A5) by $\phi(Z,n)$, then subtracting yields,

$$[\phi_\alpha(Z,n)f_+(Z)-\phi(Z,n)f_{+\alpha}(Z)]$$
$$=\phi_+(Z,n)[\bar{\phi}(1/Z,n)\phi_\alpha(Z,n)$$

$$-\phi(Z,n)\phi_\beta(1/Z,n)], \quad |Z|=1. \qquad \text{(A12)}$$

Now from Eq.(II.21)

$$\bar{\phi}(1/Z,n)\phi_\alpha(Z,n)-\phi(Z,n)\phi_\beta(1/Z,n)=W[\Psi,\hat{\Psi}],$$
$$|Z|=1 \qquad \text{(A13)}$$

which is

$$=\bar{\phi}(1/Z,1)\phi_\alpha(Z,1)-\phi(Z,1)\phi_\beta(1/Z,1)=2. \qquad \text{(A14)}$$

Therefore,

$$\phi_+(Z,n)$$
$$=\tfrac{1}{2}[\phi_\alpha(Z,n)f_+(Z)-\phi(Z,n)f_{+\alpha}(Z)],$$
$$|Z|=1 \qquad \text{(A15)}$$

and $\phi_+(Z,n)$ is an element of A^+. From Eq. (II.12),

$$\frac{\phi^*(Z,n)}{K(n)}=\frac{\phi^*(Z,n-1)}{K(n-1)}+\bar{b}(n)Z\frac{\phi(Z,n-1)}{K(n-1)} \qquad \text{(A16)}$$

and

$$\frac{Z^n\phi_\beta(1/Z,n)}{K(n)}=\frac{Z^n\phi_\beta(1/Z,n-1)}{K(n-1)}$$

$$+\bar{b}(n)Z\frac{\phi_\alpha(Z,n-1)}{K(n-1)} \qquad \text{(A17)}$$

Thus, iterating these equations up and using Eqs. (III.13) and (A7) yields,

$$f_+(Z)=K(\infty)\Big(\frac{\phi^*(Z,n)}{K(n)}+Z\sum_{i=n}^{\infty}\bar{b}(i+1)\frac{\phi(Z,i)}{K(i)}\Big) \qquad \text{(A18)}$$

and

$$f_\alpha(Z)=K(\infty)\Big(Z^n\frac{\phi_\beta(1/Z,n)}{K(n)}+Z\sum_{i=n}^{\infty}\bar{b}(i+1)\frac{\phi_\alpha(Z,i)}{K(i)}\Big). \qquad \text{(A19)}$$

Substituting these equations into (A15) gives,

$$\phi_+(Z,n)=\frac{1}{2}\frac{K(\infty)}{K(n)}[\phi^*(Z,n)\phi_\alpha(Z,n)$$

$$-(\phi(Z,n)Z^n\phi_\beta(1/Z,n)]+\frac{1}{2}\frac{K(\infty)}{K(n)}Z$$

$$\times\sum_{i=n}^{\infty}\bar{b}(i+1)\Big(\frac{\phi(Z,i)}{K(i)}\phi_\alpha(Z,n)$$

$$-\phi(Z,n)\frac{\phi_\alpha(Z,i)}{K(i)}\Big). \qquad \text{(A20)}$$

This becomes, using Eqs. (A13) and (A14),

$$=\frac{K(\infty)}{K(n)}Z^n+\frac{k(\infty)}{2K(n)}Z\sum^{\infty}\bar{b}(i+1)$$

$$\times\Big[\frac{\phi(Z,i)}{K(i)}\phi_\alpha(Z,n)-\phi(Z,n)\frac{\phi_\alpha(Z,i)}{K(i)}\Big]. \qquad \text{(21)}$$

Since $\phi(Z,i)$ and $\phi_\alpha(Z,i)$ are uniformly bounded in $i,\phi_+(Z,n)\to Z^n$ as $n\to\infty$. Using similar procedures that led to Eq. (A15), the following formula for $\hat{\phi}_+(Z,n)$ can be derived,

$$\hat{\phi}_+(Z,n)=\tfrac{1}{2}[\phi_\beta(1/Z,n)f_+(Z)+f_{+\alpha}(Z)\bar{\phi}(1/Z,n)]. \qquad \text{(A22)}$$

Because the right-hand sides of Eqs. (A15) and (A22) obey Eq. (II.13) with boundary condition (III.4), we see that $\phi_+(Z,n)$ has analytic properties similar to as $f_+(Z)$ and $f_{+\alpha}(Z)$. Using Eq. (II.13) for $\hat{\phi}_+(Z,n)$, shows us that $\hat{\phi}_+(Z,n)$ has analytic properties similar to $f_+(Z,n)$ and that $\hat{\phi}_+(Z,n)$ is an element of A^+.

Notice that once we have found the weight function $\sigma(\theta)$ from Eq. (IV.2), it is easy to see that

$$\phi_\alpha(Z,n) = \frac{1}{2\pi}\int_{-\pi}^{\pi}\left(\frac{e^{i\theta}+Z}{e^{i\theta}-Z}\right)$$

$$\times[\phi(Z',n)-\phi(Z,n)]\sigma(\theta)d\theta,$$

$$Z'=e^{i\theta}|Z|\leqslant 1. \quad \text{(A23)}$$

Since $\sigma(\theta)$ is an element of A and

$$\frac{Z^n\phi_\beta(1/Z,n)}{\phi^*(Z,n)}$$

converges uniformly for $|Z|\leqslant 1$, it can be shown[4,14] that

$$f_{+\alpha}(Z) = -\frac{f_+(Z)}{2\pi}\int_{-\pi}^{\pi}\left(\frac{e^{i\theta}+Z}{e^{i\theta}-Z}\right)\sigma(\theta)d\theta,$$

$$|Z|\leqslant 1, \quad \text{(A24)}$$

and from Eq. (A15) that

$$\phi_+(Z,n) = \frac{f_+(Z)}{2\pi}\int_{-\pi}^{\pi}\left(\frac{e^{i\theta}+Z}{e^{i\theta}-Z}\right)$$

$$\times\phi(e^{i\theta},n)\sigma(\theta)d\theta, \quad |Z|\leqslant 1. \quad \text{(A25)}$$

[1]K.M. Case, J. Math. Phys. **15**, 2166 (1974).
[2]These polynomials were first introduced by G. Szegö in the early twenties in his studies on sections of Toeplitz forms.
[3]G. Szegö, *Orthogonal Polynomials*, American Mathematical Society, Vol. 23 (1939).
[4]Ya. L. Geronimus, *Polynomials Orthogonal on a Circle and Interval* (Pergamon, New York, 1960).
[5]Here $\bar{\phi}(Z,n)$ denotes the complex conjugate polynomial while $\tilde{\phi}(Z,n)$ is the polynomial obtained by using complex conjugate coefficients.
[6]M.J. Ablowitz and J.F. Ladik, J. Math. Phys. **16**, 598 (1975).
[7]R.G. Newton, *Scattering Theory of Waves and Particles* (McGraw-Hill, New York, 1966).
[8]G. Baxter, J. Math. Anal. Appl. **2**, 223 (1961).
[9]N.I. Akhiezer, *The Classical Moment Problem* (Olive and Boyd, Edsinbrugh, 1965).
[10]G. Baxter, Trans Amer. Math. Soc. **99**, 471 (1961).
[11]This is also true of $\phi^*(Z,n)$.[3,4]
[12]A. Zygmond, *Trignometric Series*—1 (Cambridge U.P., New York, 1959).
[13]G. Baxter, Ill. J. Math. 7, 97 (1963).
[14]J.S. Geronimo, *Scattering Theory and Orthogonal Polynomials*, Doctoral dissertation.
[15]K.M. Case, J. Math. Phys. **17**, 1703 (1976).
[16]U. Grenander and G. Szegö, *Toeplitz Forms and Their Applications* (University of California, Berkeley, 1959).
[17]M. Kac, Duke Math. J. **21**, 501 (1954).
[18]I.I. Hirshman, Jr., J. Anal. Math. **14**, 225 (1965).
[19]R.E. Hartwig and M.E. Fisher, Arch. Rat. Mech. Anal. **32**, 190 (1969).
[20]H. Widom, Adv. Math. **21**, 1 (1976).
[21]F.J. Dyson, Commun. Math. Phys. **47**, 171 (1976).

Covariant perturbed wave equations in arbitrary type-D backgrounds

Alan L. Dudley and J. D. Finley, III

The University of New Mexico, Albuquerque, New Mexico 87131
(Received 25 April 1978)

We present an approach to the fundamental tensorial quantities of general relativity which is inherently covariant and based on the irreducible representations of the Lorentz group, O(3,1). Using this technique, the wave equations appropriate to perturbing, massless, $D(0,s)$, fields in an arbitrary curved background are studied and a relationship between the decoupling of (at least) one of the equations and the algebraic degeneracy of the spacetime is shown. It is then found that sufficient conditions for decoupling the equations determining both of the radiative components (the extremal helicities) are that the space be of type D. Using Plebański-Demiański coordinates to describe such an arbitrary vacuum spacetime (of type D), we separate the (decoupled) perturbation equations for the radiative components corresponding to spin $s = 0$, 1/2, 1, and 2.

1. INTRODUCTION

In recent years there has been considerable interest in (linear) perturbations to gravitational fields, as well as test electromagnetic and neutrino fields imposed on a gravitational (or electrogravitational) background. The study of such perturbations on a flat background (i.e., with no initial gravitational field) was begun by Einstein.[1] However, of considerably more interest for astrophysical problems are perturbations away from some model which one expects to represent a reasonable idealization of a particular astrophysical system.[2] Much time has been devoted to the study of perturbations away from a vacuum Schwarzchild background.[3] A new level of perturbation capability was achieved by Teukolsky[4] who successfully attacked the problem for the background of a rotating black hole, the Kerr geometry,[5] reducing the acquisition of desired results to the solution of ordinary differential equations, thereby allowing an extensive numerical and analytic study of perturbations in this case. This result was actually quite important not only for the examination of physically reasonable approximations to particular astrophysical problems,[6] but also for a better understanding of the general structure of solutions of Einstein's equations—in particular those solutions "near" the Kerr solution.[7]

It is this second benefit that is of more interest to us here. However, from that point of view, the Kerr geometry is only a special case of the much broader class of Petrov type D metrics. In fact Carter[8] has already studied the separability of various scalar equations in such spaces, while Stewart and Walker[9] have made a general study of gravitational perturbations to type D spacetimes. In this study the determination of arbitrary perturbations to a known type D solution of Einstein's equations will be reduced to the solution of uncoupled ordinary differential equations.

The problem of gravitational perturbations to the Kerr geometry had been resistant to many efforts until Teukolsky generated the desired decoupled solution in the Newman–Penrose[10] formalism. Therefore, more recent studies have been done almost exclusively in this formalism, or the compact revision of Geroch, Held, and Penrose.[11] Because of its fundamental use of null tetrads which can be aligned with the radiation,[12] this formalism is well suited to examine the radiative processes which are of most interest. In addition, all type D background geometries (such as Kerr or Schwarzchild) have a pair of doubly-degenerate Debever–Penrose null directions which may be used to align the null tetrads in physically relevant directions. However, the procedures used are not covariant and are algebraically burdensome. This study develops, instead, a completely covariant approach to massless field equations of arbitrary half-integral spin using the irreducible representations of the Lorentz group most appropriate to each case. This formalism, being based in large part on the group isomorphism between O(3,1), the Lorentz group, and $O(3,C)$, the group of rotations in three *complex* dimensions, is very similar to that used by Debever[13] and others working with him,[14] but is extended to the more general use of other representations as well and to the utilization of an extensive algebraic superstructure which reduces many of the complicated manipulations to an algorithmic level.

The equations governing a perturbing massless field of spin s (with $s = 2$ for the gravitational case) can be written as a set of $2s + 1$ wavelike equations in which the various different helicity components of the perturbing field are coupled not only with each other but also with the curvature of the background space, all with four independent variables as coordinates over the manifold. The problem is to decouple the $2s + 1$ equations, or some physically important subset of them, and to then separate the decoupled equations so as to obtain *ordinary* differential equations, which can be handled numerically if necessary. The fact that all Petrov type D vacuum metrics possess at least two Killing vectors suggests the possibility of at least partial separation. However, Debever[15] has also shown that a conformal Killing tensor is admitted by all of the type D solutions of the Einstein–Maxwell equations found by Plebański and Demiański,[16] referred to hereafter as PD solutions. Since Weir[17] has shown that the PD solutions include all vacuum, type D solutions, the existence of this conformal Killing tensor suggests the possibility, in the PD background geometry, of complete separation of the massless equations for arbitrary spin.

 0022-2488/79/020311-18$01.00

In Sec. 2 we first discuss our approach to the usual determinants of a spacetime (affine connections, Riemann tensor, etc.) via the appropriate irreducible representations of the Lorentz group, defining the projection operators to the carrier spaces of these representations and establishing a generalized covariant derivative defined over all these carrier spaces. In Sec. 3 we write the usual first order field equations (the Bianchi equations for gravitation, the Maxwell equations for electromagnetism) in such a way as to reflect their inherent helicity structure. These equations are then iterated to acquire the coupled wavelike equations for each helicity component. Next general conditions under which these equations may be decoupled are discussed and this decoupling is performed for an arbitrary type D geometry, in certain specific families of allowed gauges. Our method is not particularly difficult, algebraically, and allows one to more or less see the physical meanings of each step of the derivations. All of this is done without the introduction of a specific coordinate system. Therefore, if perturbations to some type D metric other than Kerr are desired, it is not necessary to start over at the beginning for that particular case. In Sec. 4, as an addition to the main problem of perturbations to the gravitational field, the method of Sec. 3 is extended to obtain wavelike equations for massless fields of arbitrary spin s, corresponding to the representation D(0,s). It is then shown that the equations for helicity $h=\pm s$ completely decouple, as before, if the background geometry is of type D.

In Sec. 5 we introduce PD coordinates, still allowing general vacuum, type D solutions, with a possible cosmological constant and some electromagnetic parameters. These include, for example, the accelerating metrics of Levi–Civita[18] (more recently, Kinnersley[19]). However, only perturbations of one spin value at a time are considered. Therefore, in the case of nonzero electric and magnetic charge allowed by the PD solutions we either keep the geometry fixed and perturb the electric field, or, of more interest, keep the electric field fixed and perturb the geometry (gravitational field), as would be appropriate, say, for small (test) values of these charges.[20] The much more difficult problem of mixed perturbations involving more than one nonzero background field and simultaneous perturbations away from these values is still being studied.[21] We show then that the decoupled equations for the radiative helicity components (helicity values of $\pm s$) separate only for the spins 0, $\frac{1}{2}$, 1, and 2, for an arbitrary vacuum type D background. Lastly, we note that the results dervied here for these separated equations have already been announced without derivations in Ref. 22.

2. DEVELOPMENT OF A COVARIANT FORMALISM

The space–time under consideration may be viewed as a four-dimensional hyperbolic Riemannian manifold endowed with a line element, ds^2 represented by a (complex) null tetrad basis of 1-forms,

$$g=ds^2=2e^1 \underset{s}{\otimes} e^2+2e^3 \underset{s}{\otimes} e^4=g_{\alpha\beta}e^{\alpha}\otimes e^{\beta}, \qquad (2.1)$$

where $e^1=\bar{e}^2$, while e^3 and e^4 are real. (The bar is used to denote complex conjugation.) The tetradial indices are manipulated by

$$g_{\alpha\beta}=\begin{pmatrix}0&1&0&0\\1&0&0&0\\0&0&0&1\\0&0&1&0\end{pmatrix}, \qquad (2.2)$$

and its inverse, $g^{\alpha\beta}$, which is numerically the same. (See Ref. 23 for a statement of our sign and index conventions.) We will investigate the structure of the space–time manifold through the language of differential forms.[24] The first structure equations

$$de^{\alpha}=e^{\beta}\wedge\omega^{\alpha}{}_{\beta} \qquad (2.3)$$

serve to define the connection 1-forms (without torsion), $\omega_{\alpha\beta}$, which are skew in α,β, because of the constancy of $g_{\alpha\beta}$, by

$$0=dg_{\alpha\beta}=2\omega_{(\alpha\beta)}. \qquad (2.4)$$

Cartan's second structure equations then determine the curvature 2-forms

$$\Omega_{\alpha\beta}=\omega_{\alpha\beta}+\omega_{\alpha\gamma}\wedge\omega^{\gamma}{}_{\beta}, \qquad (2.5a)$$

whose components determine the Riemann tensor by

$$\Omega_{\alpha\beta}=\tfrac{1}{2}R_{\alpha\beta\gamma\delta}e^{\gamma}\wedge e^{\delta}. \qquad (2.5b)$$

A specific set of basis forms are defined by the metric only to within the Lorentz transformation, $N^{\sigma}{}_{\alpha}$, such that

$$e'^{\sigma}=N^{\sigma}{}_{\alpha}e^{\alpha}, \quad g_{\sigma\tau}=g'_{\sigma\tau}\equiv N_{\sigma}{}^{\alpha}N_{\tau}{}^{\beta}g_{\alpha\beta}. \qquad (2.6)$$

In general, more complicated tensors will transform, under a redefinition of the tetrad, with tensor products of these matrices. However, it is well known that valuable simplifications occur if one breaks any particular tensor up into its parts which transform under particular *irreducible* representations of the fundamental group—in this instance the Lorentz group, O(3,1). Since we are already using *complex* quantities in our tetrad so as to better study null quantities, we look for all representations of O(3,1) irreducible over the complex numbers. [It is actually only the component of O(3,1) connected to the identity that is under consideration here.] All such finite-dimensional representations are well known.[25] They may be specified by a pair of half-integers, j, $j'=0,\frac{1}{2},1,\frac{3}{2},\cdots$, and are denoted by $D(j,j')$. (Each index separately behaves as a usual angular momentum "quantum number" of the type associated with representations of the group of rotations in three dimensions.)

Associated with each representation $D(j,j')$ is the carrier space on which it acts, denoted by $V(j,j')$, which is a vector space of dimension $(2j+1)\times(2j'+1)$. These vector spaces are of course isomorphic to subspaces of particular tensor spaces over the manifold. Therefore, it is convenient to define projection operators, acting on the (graded) tensor algebra, which map tensors into particular carrier spaces. We use $Z(n;j,j')$ to denote the mapping into $V(j,j')$ restricted to act on tensors of order n, when an abstract symbol is needed.

Various specific such projections will be needed often, such as the ones to $V(0,1)$, $V(1,0)$, $V(1,1)$, $V(0,2)$, and $V(2,0)$. Of the representations, the most important are surely $D(0,1)$ and $D(1,0)$, which exist because of the group *isomorphism* of O(3,1) and O(3,C). We start constructing these projections by looking at the space of 2-forms, splitting it into two subspaces via the usual Hodge duality. That is, a basis of (six) 2-forms is found, half of which are self-dual, and half of which are anti-self-dual.[26]

Taking the range of lower case Latin indices to be $+$, 0, $-$, a basis for self-dual 2-forms is the set

$$\mathscr{Z}^a = \frac{i}{\sqrt{2}}\left(e^4\wedge e^2, \frac{e^1\wedge e^2 + e^3\wedge e^4}{\sqrt{2}}, e^3\wedge e^1\right), \tag{2.7a}$$

and a basis for anti-self-dual 2-forms is the set

$$\mathscr{Z}^{\dot a} = \frac{-i}{\sqrt{2}}\left(e^4\wedge e^1, \frac{-e^1\wedge e^2 + e^3\wedge e^4}{\sqrt{2}}, e^3\wedge e^2\right), \tag{2.7b}$$

where the dot over the index indicates their anti-self-dual nature. Under the transformation of the basis tetrad to another equivalent one via the Lorentz transformations [discussed at (2.6)], the $\mathscr{Z}^a$ transform among themselves according to the representation $D(0,1)$. In each of these carrier spaces a metric g_{ab} (and $g_{\dot a\dot b}$ which is numerically equal to g_{ab}), suitable for the raising and lowering of indices, is induced by the original metric tensor on the manifold:

$$g_{ab} = \mathscr{Z}_a{}^{\alpha\beta}\mathscr{Z}_b{}^{\gamma\delta}g_{\alpha\gamma}g_{\beta\delta} = \begin{pmatrix} 0 & 0 & -1\\ 0 & 1 & 0\\ -1 & 0 & 0\end{pmatrix}, \tag{2.8}$$

which is just the usual metric known from the similar representations of the rotation group, for spin-1 particles.

These basis forms satisfy the following useful identities (see the Appendix as well):

$$*\mathscr{Z}^a = \mathscr{Z}^a, \quad *\mathscr{Z}^{\dot a} = -\mathscr{Z}^{\dot a},$$

$$2\mathscr{Z}^a\lrcorner\mathscr{Z}^b\,\mathscr{Z}^a{}_{\alpha\beta}\mathscr{Z}^{b\alpha\beta} = g^{ab}, \quad \mathscr{Z}^a\lrcorner\mathscr{Z}^{\dot b} = 0,$$

$$\mathscr{Z}^a\cdot\mathscr{Z}^b = \mathscr{Z}^a{}_{\mu\nu}\mathscr{Z}^{b\nu}{}_\lambda e^\mu\otimes e^\lambda$$

$$= -\tfrac{1}{8}g^{ab}g - \tfrac{1}{2}\eta^{abc}\mathscr{Z}_c, \tag{2.9}$$

where $g = g_{\alpha\beta}e^\alpha\otimes e^\beta$ is just the metric and η^{abc} is the totally skew tensor in the carrier space of $D(0,1)$. [This tensor is constructed from the three-dimensional Levi–Civita symbol with $s=0$ (see Ref. 26); this implies, e.g., that $\eta^{+0-} = +i = -\eta_{+0-}$.] Several different products are used in this work, which results in considerable convenience. In addition to the tensor and exterior products, we use the right interior product $B\lrcorner A$ and a specific contraction of adjacent tensor indices, denoted by a dot and defined over the tensor space.[27]

Since, together, the $\mathscr{Z}^a$ and $\mathscr{Z}^{\dot a}$ form a basis for all 2-forms, an arbitrary 2-form τ may be decomposed into 2 sets of components, τ^a and $\tau^{\dot a}$, which may then be thought of as being elements of $V(0,1)$ and $V(1,0)$, respectively,

$$\tau = \tfrac{1}{2}\tau_{\alpha\beta}e^\alpha\wedge e^\beta = \tau_a\mathscr{Z}^a + \tau_{\dot a}\mathscr{Z}^{\dot a}, \tag{2.10}$$

$$\tau^a = 2\tau\lrcorner\mathscr{Z}^a = \tau_{\alpha\beta}\mathscr{Z}^{a\alpha\beta}, \quad \tau^{\dot a} = 2\tau\lrcorner\mathscr{Z}^{\dot a}. \tag{2.11}$$

Therefore $\mathscr{Z}^a{}_{\alpha\beta}$ ($\mathscr{Z}^{\dot a}{}_{\alpha\beta}$) constitutes a specific realization of the projection map $Z(2;0,1)$ [$Z(2;1,0)$].

It is now most convenient to introduce a generalization of the operator, D, which Cartan[28] called the tensorial exterior derivative. This operator takes any indexed set of p-forms, such as e^α or $\Omega_{\alpha\beta}$, and generalizes the exterior derivative so that it treats the set as the components of a tensor and gives the components of the covariant derivatives of that tensor. As an example for some set $T^\alpha{}_\beta$ of p-forms, the $(p+1)$-forms $DT^\alpha{}_\beta$ are given by

$$DT^\alpha{}_\beta = dT^\alpha{}_\beta + \omega^\alpha{}_\gamma\wedge T^\gamma{}_\beta - \omega^\gamma{}_\beta\wedge T^\alpha{}_\gamma. \tag{2.12}$$

Therefore, the first structure equations (2.3) merely say that

$$De^\alpha = 0. \tag{2.13}$$

Additionally, if σ is a single p-form,

$$d\sigma = D\sigma = (D\sigma_\alpha)\wedge e^\alpha = \sigma_{\alpha;\beta}e^\beta\wedge e^\alpha. \tag{2.14}$$

The above describes Cartan's original formulation. But our interest in several different $V(j,j')$ requires that the definition be extended to include, as well, the indices denoting components of an element of some such $V(j,j')$. In order to do this the covariant derivative must be extended to these spaces, or, better, to each bundle of such spaces, one at each point of the manifold. We do this in the simplest possible way by requiring that D commute with the projections $Z(n;j,j')$. In particular, for $D(0,1)$, this requires that

$$0 = D\mathscr{Z}^a = d\mathscr{Z}^a + W^a{}_b\wedge\mathscr{Z}^b, \tag{2.15}$$

where the required connection 1-forms in $V(0,1)$ are denoted by W^{ab}. Then, using the properties of $\mathscr{Z}^a$ given in (2.9), it is easily found that

$$W^{ab} = \eta^{abc}\mathscr{Z}_c{}^{\alpha\beta}\omega_{\alpha\beta}, \tag{2.16}$$

with an analogous equation generated by $D\mathscr{Z}^{\dot a} = 0$. Also, thinking about the usual relation between the commutator of covariant derivatives and the Riemann tensor, we find, for the tensorial case,

$$DDV^\alpha = \Omega^\alpha{}_\beta V^\beta, \tag{2.17}$$

while we also calculate, at somewhat greater effort, that

$$DDF^a = \eta^{abc}\mathscr{Z}_c{}^{\alpha\beta}\Omega_{\alpha\beta}F_b. \tag{2.18}$$

From (2.16) and (2.18) it is clear that the irreducible parts of $\omega_{\alpha\beta}$ and $\Omega_{\alpha\beta}$ should be determined since they play the role of connection and curvature in the irreducible subspaces $V(0,1)$ [as well as $V(1,0)$], modulo the factor η^{abc} whose role will be discussed soon. Therefore, we decompose the connection and curvature forms:

$$\omega_{\alpha\beta} = \omega_a\mathscr{Z}^a{}_{\alpha\beta} + \omega_{\dot a}\mathscr{Z}^{\dot a}{}_{\alpha\beta},$$

$$\omega^a = \mathscr{Z}^{a\alpha\beta}\omega_{\alpha\beta}, \quad \omega^{\dot a} = \mathscr{Z}^{\dot a\alpha\beta}\omega_{\alpha\beta},$$

$$\Omega_{\alpha\beta} = \Omega_a \mathscr{Z}^a{}_{\alpha\beta} + \Omega_{\dot a} \mathscr{Z}^{\dot a}{}_{\alpha\beta},$$
$$\Omega^a = \mathscr{Z}^{a\alpha\beta}\Omega_{\alpha\beta}, \quad \Omega^{\dot a} = \mathscr{Z}^{\dot a\alpha\beta}\Omega_{\alpha\beta}. \tag{2.19}$$

The 1-forms ω^a and $\omega^{\dot a}$ are now decomposed as far as is profitable. In fact, the twelve components of the three 1-forms ω^a are just the Newman–Penrose rotation coefficients put together in a form more convenient for covariant manipulations, except for a multiplicative factor. In explicit manipulations, it is convenient to define a set Γ_a by

$$\left(\omega_{24}, \frac{-\omega_{12}-\omega_{34}}{\sqrt{2}}, \omega_{13}\right) = \Gamma_a \equiv \frac{1}{\sqrt{2}i}\,\omega_a,$$

$$-\Gamma_{a\alpha} = \begin{pmatrix} \rho & \sigma & -\tau & \kappa \\ \sqrt{2}\alpha, & \sqrt{2}\beta, & -\sqrt{2}\gamma, & \sqrt{2}\epsilon \\ \lambda & \mu & -\nu & \pi \end{pmatrix}, \tag{2.20}$$

where the assortment of signs reflects the different signature of the metrics chosen by the present authors and that of Newman and Penrose.

Since many calculations are done with these quantities, the $D(0,1)$ forms of several common identities are listed below:

$$\Omega_a = d\omega_a - \tfrac{1}{2}\eta_{abc}\,\omega^b \wedge \omega^c, \tag{2.21}$$

$$D\Omega_a = 0 \text{—(second) Bianchi equations,} \tag{2.22}$$

$$0 = DD\,\mathscr{Z}^a$$
$$= \eta^{abc}\Omega_b \wedge \mathscr{Z}_c \text{— first Bianchi identities.} \tag{2.23}$$

Since the Ω_a and $\Omega_{\dot a}$ are 2-forms, they can be expanded further to give

$$\Omega_a = \Omega_{ab}\mathscr{Z}^b + \Omega_{a\dot b}\mathscr{Z}^{\dot b},$$
$$\Omega_{\dot a} = \Omega_{\dot a b}\mathscr{Z}^b + \Omega_{\dot a\dot b}\mathscr{Z}^{\dot b}. \tag{2.24}$$

From the symmetries of the Riemann tensor it is easily shown that

$$\Omega_{a\dot b} = \mathscr{Z}_a{}^{\alpha\beta}\mathscr{Z}_{\dot b}{}^{\gamma\delta}R_{\alpha\beta\gamma\delta} = \overline{\Omega_{\dot a b}}, \tag{2.25}$$

so that there are nine (real) degrees of freedom. In particular

$$\Omega_{a\dot b} = -2\,\mathscr{Z}_a{}^{\mu\nu}\mathscr{Z}_{\dot b\nu}{}^{\lambda}R_{\mu\lambda} \equiv W_{a\dot b}{}^{\mu\lambda}R_{\mu\lambda}. \tag{2.26}$$

Here $R_{\mu\lambda} = R^\alpha{}_{\mu\alpha\lambda}$ is the Ricci tensor, while $W_{a\dot b}{}^{\mu\lambda}$, which is symmetric and traceless on the indices μ,λ is a realization of the projection mapping $Z(2;1,1)$. The inverse of (2.26) is given by

$$R_{\mu\lambda} - \tfrac{1}{4}g_{\mu\lambda}R = W^{a\dot b}{}_{\mu\lambda}\Omega_{a\dot b}. \tag{2.27}$$

Also $\Omega_{ab} = \overline{\Omega_{\dot a\dot b}}$, is clearly symmetric on its pair of indices, but is not yet irreducible. Instead a scalar portion may be extracted,

$$\Omega^a{}_a = \Omega^{\dot a}{}_{\dot a} = \tfrac{1}{2}R \equiv \tfrac{1}{2}g^{\alpha\beta}R_{\alpha\beta}; \tag{2.28}$$

then we may define the irreducible portion as

$$C_{ab} = \Omega_{ab} - \tfrac{1}{6}g_{ab}R, \tag{2.29}$$

which is symmetric and traceless, and therefore equivalent to an element of $V(0,2)$. To display this more explicitly a renumbering operator is introduced,

$$Z_A{}^{ab} = \begin{pmatrix} \delta_A^{++} & \frac{1}{\sqrt{2}}\delta_A^{+} & \frac{1}{\sqrt{6}}\delta_A^{0} \\ \frac{1}{\sqrt{2}}\delta_A^{+} & \frac{2}{\sqrt{6}}\delta_A^{0} & \frac{1}{\sqrt{2}}\delta_A^{-} \\ \frac{1}{\sqrt{6}}\delta_A^{0} & \frac{1}{\sqrt{2}}\delta_A^{-} & \delta_A^{--} \end{pmatrix}, \tag{2.30}$$

where the capital Latin indices take on the five values $++$, $+$, 0, $-$, $--$, which label the components of an element of $V(0,2)$, and the rather peculiar coefficients have been chosen so that the metric induced in $V(0,2)$ is simply

$$g_{AB} = Z_A{}^{ab}Z_B{}^{cd}g_{ac}g_{bd} = \begin{pmatrix} 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \end{pmatrix}, \tag{2.31a}$$

which is again common in studies of spin-2 systems under the ordinary group of rotations. [The $Z_A{}^{ab}$ are in fact just the usual vector-addition coefficients $C(1,1,2;a,b,A)$ for the rotation group.] Using the renumbering operator we obtain an explicit realization of $Z(4;0,2)$, namely

$$Z_A{}^{\alpha\beta\gamma\delta} = \mathscr{Z}_a{}^{\alpha\beta}\mathscr{Z}_b{}^{\gamma\delta}Z_A{}^{ab}. \tag{2.31b}$$

Therefore, the irreducible content of the curvature consists of

$$C_A = Z_A{}^{ab}\Omega_{ab} = Z_A{}^{\alpha\beta\gamma\delta}R_{\alpha\beta\gamma\delta} \text{ and } C_{\dot A}$$
$$= \overline{C_A} \text{ — the conformal curvature,} \tag{2.32}$$
$$R = 2\Omega^a{}_a = 2\Omega_{\dot a}{}^{\dot a} = R^{\alpha\beta}{}_{\alpha\beta} \text{— the Ricci scalar,}$$
$$\Omega_{a\dot b} = W_{a\dot b}{}^{\alpha\beta}R_{\alpha\beta}$$
$$= \mathscr{Z}_a{}^{\gamma\alpha}\mathscr{Z}_{\dot b}{}^{\delta\beta}R_{\gamma\alpha\delta\beta} \text{— the traceless Ricci tensor,}$$

tensor,

which are elements of $V(0,2)$, $V(2,0)$, $V(0,0)$, and $V(1,1)$, respectively. The C_A are just the usual ψ_0 through ψ_4 introduced by Newman and Penrose, modulo constant numerical factors, which are given in detail in the Appendix.

As before, using $DZ_A{}^{ab} = 0$ as a requirement induces a covariant derivative on $V(0,2)$, which provides a connection there, described as follows. Let T_A represent some (covariant) element of $V(0,2)$. Then

$$D\,T_A = d\,T_A - \mathscr{W}^B{}_A T_B, \tag{2.33}$$

where the 1-forms $\mathscr{W}^{BA}$ are defined as

$$\mathscr{W}^{BA} = -\mathscr{J}^{aBA}\omega_a, \tag{2.34}$$

and the $\mathscr{J}^{aBA}$ are the $j=2$ representations of the angular momentum operators which generate the rotation group. Their appearance in the equations follows from

$$\mathscr{J}_a{}^{AB} = -2Z^{Abc}Z^{Bd}{}_c\eta_{bda}. \tag{2.35}$$

In the Appendix the $\mathscr{J}^{(2)}{}_a{}^{AB}$ are displayed along with a few of their relevant properties. This also indicates the reason for the appearance of η^{abc} in (2.16) since the representation of the angular momentum operators for $j=1$ is just $-\eta^{abc}$. This coupling of the appropriate representation matrices for $\mathscr{J}^{(s)}{}_a$ with the projections of the tensorial connection is to be expected since, from a modern pont of view,[29] the connection form takes its values in the Lie algebra of the group generated by the allowed tetrad equivalences—the Lorentz group. It is convenient, as well, to note the usual commutator identity over $V(0,2)$,

$$DD\,T_A = -\mathscr{J}^d{}_{AB}\Omega_d T^B. \tag{2.36}$$

3. PERTURBATIONS OF WAVE EQUATIONS

In this section both the field equations and their perturbations for gravitation and electromagnetism—the Bianchi and Maxwell equations, respectively—are discussed. We include the electromagnetic case *mostly* because it shares many important properties with the gravitational case, while being considerably simpler. It is convenient to study radiation generated by these perturbations in terms of appropriate covariant second-order wave equations, generated from the first-order field equations. Therefore these equations are derived as well, and it is shown that appropriate components of these equations may be decoupled provided we have certain conditions on the background gravitational field. If we require that both equations of extremal helicity decouple, these conditions amount to the requirement that the background conformal field be of type D (or conformally flat).

In terms of the usual electromagnetic field tensor, $F_{\alpha\beta}$, Maxwell's equations are just

$$dF=0, \quad {}^*d{}^*F=4\pi iJ, \tag{3.1}$$

which become, in $V(0,1)$.

$$e^\alpha\mathscr{Z}^b{}_{\alpha\beta}D^\beta F_b = \mathscr{Z}^b\cdot DF_b = 2\pi J. \tag{3.2}$$

Contracting (3.2) with $\mathscr{Z}^a$ and using (2.9) gives

$$DF_a + 2\eta_{abc}\mathscr{Z}^c\cdot DF^b = -8\pi\,\mathscr{Z}_a\cdot J, \tag{3.3}$$

from which we obtain, using (2.18),

$$(i{}^*D{}^*D - R/3)F_a + C_a{}^bF_b = 8\pi\mathscr{Z}_a \rfloor DJ. \tag{3.4}$$

This is the fundamental covariant $D(0,1)$ massless wave equation.

In a similar way, by inserting the full decomposition of Ω_a,

$$\begin{aligned}\Omega_a &= C_{ab}\mathscr{Z}^b + (R/6)\mathscr{Z}_a + \Omega_{ab}\mathscr{Z}^b\\ &= C_{ab}\mathscr{Z}^b + \mathscr{Z}_{a\gamma\delta}M^\delta{}_\epsilon e^\gamma\wedge e^\epsilon,\end{aligned} \tag{3.5}$$

where

$$M_{\alpha\beta} = R_{\alpha\beta} - (R/6)\,g_{\alpha\beta}, \tag{3.6}$$

into the Bianchi equations (2.22), the appropriate decomposition is obtained,

$$B_a \equiv \mathscr{Z}^b\cdot DC_{ab} - e^\alpha\mathscr{Z}_a{}^{\gamma\beta}D_\beta M_{\gamma\alpha} = 0, \tag{3.7}$$

$$D_{[\alpha}M_{\beta]}{}^\alpha = 0. \tag{3.8}$$

This result corresponds to (3.2) for the electromagnetic field. (The one-form B_a is simply a covenient name for the equations for later use.) Contracting (3.7) with $\mathscr{Z}^c$ gives, as before,

$$DC_{ab} + 2\eta_{acd}\mathscr{Z}^d\cdot DC^c{}_b = -4e^\alpha\mathscr{Z}_{a\alpha\delta}\mathscr{Z}_b{}^{\gamma\beta}D_\beta M_\gamma{}^\delta, \tag{3.9}$$

from which we obtain, using (2.18) and (3.5),

$$\begin{aligned}&(i{}^*D{}^*D - \tfrac12 R)\,C_{ab} + 3C_{ad}C^d{}_b - g_{ab}C^{cd}C_{cd}\\ &\quad = -4\,\mathscr{Z}_{a\alpha\delta}\mathscr{Z}_b{}^{\gamma\beta}D^\alpha D_\beta M_\gamma{}^\delta.\end{aligned} \tag{3.10}$$

Projecting this equation into one over $V(0,2)$, via $Z_A{}^{ab}$, requires the introduction of the invariant tensor h_{ABC}, which is considered next.

There are two quantities, invariant under Lorentz transformations, that can be formed from a $V(0,2)$-type quantity, one quadratic and one cubic in the conformal tensor and created using $g^{\alpha\beta}$ and $\eta^{\alpha\beta\gamma\delta}$. Since these are properties of the $D(0,2)$ representation, they are generated by two tensors g^{AB} and h^{ABC}, both totally symmetric, and invariant under Lorentz transformations represented in $V(0,2)$:

$$g_{AB}C^AC^B = \tfrac12\Big(C^{\alpha\beta\gamma\delta}C_{\alpha\beta\gamma\delta} + \frac{i}{2}\eta^{\alpha\beta\gamma\delta}C_{\alpha\beta\epsilon\zeta}C_{\gamma\delta}{}^{\epsilon\zeta}\Big), \tag{3.11a}$$

$$h^{ABC}C_AC_BC_C = \tfrac38\Big(C^{\alpha\beta\gamma\delta}C_{\epsilon\zeta\gamma\delta}C_{\alpha\beta}{}^{\epsilon\zeta} + \frac{i}{2}\eta^{\alpha\beta\gamma\delta}C_{\alpha\beta\epsilon\zeta}C_{\gamma\delta\eta\theta}C^{\epsilon\zeta\eta\theta}\Big), \tag{3.11b}$$

$$\begin{aligned}h^{ABC} &= \frac{1}{\sqrt6}\,[\delta^A_0\delta^B_0\delta^C_0 - \tfrac12\,\delta^A_{(+}\delta^B_0\delta^C_{-)} - \delta^A_{(+}{}_{+}\delta^B_0\delta^C_{-}{}_{-)}]\\ &\quad + \tfrac14[\delta^A_{(+}{}_{+}\delta^B_{-}\delta^C_{-)} + \delta^A_{(-}{}_{-}\delta^B_{+}\delta^C_{+)}]\\ &= \tfrac12\eta^{ace}\eta^{bdf}Z^A_{ab}Z^B_{cd}Z^C_{ef}\end{aligned} \tag{3.11c}$$

These invariant tensors are analogous to the two tensors, $g_{\alpha\beta}$, and $\eta_{\alpha\beta\gamma\delta}$, in the original tensor space, and are invariant under proper Lorentz transformations, but they have different parity. They can also be thought of as generated by the projection operator, $Z_A{}^a{}_b$, considered as a 3×3 matrix,

$$g^{AB} = \mathrm{tr}(Z^AZ^B), \quad h^{ABC} = \mathrm{tr}(Z^AZ^BZ^C). \tag{3.12}$$

The projection of (3.10) may then be written in the final form

$$(i^*D^*D-\tfrac{1}{2}R)\,C_A+3h_{ABC}C^BC^C=-4M_A, \quad (3.13)$$

where M_A is simply an abbreviation for the source terms, $Z_A{}^{\alpha\beta\gamma\delta}D_\alpha D_\delta M_{\beta\gamma}$. The differential operator on the left can also be written in the form

$$\begin{aligned} i^*D^*DC_A &= D^\alpha D_\alpha C_A \\ &= \Box C_A - 2\mathscr{W}^B{}_A{}^\alpha C_{B,\alpha} - (\mathscr{W}^B{}_A{}^\alpha{}_{;\alpha} \\ &\quad - \mathscr{W}^C{}_A{}^\alpha \mathscr{W}^B{}_{C\alpha})C_B, \end{aligned} \quad (3.14)$$

where $\Box$ is the usual covariant d'Alembertian as it acts on a scalar.

These wave equations for the gravitational and electromagnetic cases are special cases of a more general formulation which can be given for arbitrary massless fields of spin s, corresponding to a $D(0,s)$ representation of the Lorentz group, which is given in the next section. However, the main purpose of these equations is to determine linear perturbations of the relevant field. In the gravitational case we consider arbitrary first order perturbations of the background geometry. The most interesting perturbations induce nonzero values of C_{++} and/or C_{--}, which are referred to as the radiative components of the conformal tensor. Situations involving the simultaneous perturbations of multiple fields with different spins are not considered. Hence, for spins other than 2, only test fields, in an unperturbed background geometry, are considered.

The technique used for the gravitational case is to consider a new metric—a new Riemannian manifold—which is only slightly different from the background metric. This new metric is thought of as generated by a null tetrad, as before, which is now perturbed slightly from the background values; however we maintain invariant the form of the *tetradial components* of the metric, $g_{\alpha\beta}$, as given by (2.2). That is, the new tetrad is chosen such that, with respect to the new metric, it maintains the convenient null form in (2.2). We then generate perturbed connections, inverse tetrads, fields, curvature, etc. The perturbations in all variables are assumed to be "first-order small" with respect to their background values, or, specifically, terms of order greater than one in the perturbations are ignored, as is commonly done.

In a manner analogous to the discussion at the beginning of Sec. 2, the new manifold may be described by a null tetrad (basis of 1-forms)

$$e^\alpha=e^{(0)\alpha}+\delta e^\alpha, \quad g=g_{\alpha\beta}e^\alpha\otimes e^\beta, \quad (3.15)$$

where the superscript (0) is, temporarily, used to denote background variables while the perturbation of the tetrad is denoted with a δ. As soon as feasible, the use of this superscript will be dropped, when no confusion should arise. The matrix $g_{\alpha\beta}$ is numerically the same as the one in the background space, while, of course, g is the new metric,

$$g=g^{(0)}+2g_{\alpha\beta}e^{(0)\alpha}\underset{s}{\otimes}\delta e^\beta. \quad (3.16)$$

The inverse tetrad (basis of 1-vectors), which we denote by ∂_α, is similarly perturbed, so that

$$\partial_\alpha=\partial^{(0)}{}_\alpha+\delta\partial_\alpha. \quad (3.17)$$

Since

$$e^\alpha(\partial_\beta)=\delta^\alpha{}_\beta=e^{(0)\alpha}(\partial^{(0)}{}_\beta),$$

a relation between the perturbations is obtained,

$$e^{(0)\alpha}(\delta\partial_\beta)=-\delta e^\alpha(\partial^{(0)}{}_\beta). \quad (3.18)$$

If we conceive of the two tensor algebras as being the same from the point of view of a nonmetric structure, but merely having different metric structures, as is common and convenient, then we may use $e^{(0)\alpha}$ as a basis for 1-forms over the new manifold and write

$$\delta e^\alpha=C^\alpha{}_\beta e^{(0)\beta}, \quad \delta\partial_\alpha=B^\beta{}_\alpha\partial^{(0)}{}_\beta, \quad (3.19a)$$

$$C^\alpha{}_\beta=-B^\alpha{}_\beta, \quad (3.19b)$$

where the last line of the equation is simply the restatement of (3.18). In general $C_{\alpha\beta}$ has no symmetry properties, so that there are 16 independent quantities which may be thought of as generating the perturbation. However, (3.16) tells us

$$g-g^{(0)}=2C_{(\alpha\beta)}e^{(0)\alpha}\otimes e^{(0)\beta} \quad (3.20)$$

so that the symmetric part of $C_{\alpha\beta}$ corresponds to the quantity $h_{\mu\nu}\equiv g_{\mu\nu}-g^{(0)}{}_{\mu\nu}$ in the usual coordinate-basis versions of gravitational perturbation theory. It is therefore useful to split $C_{\alpha\beta}$ into two distinct parts,

$$H_{\alpha\beta}\equiv 2C_{(\alpha\beta)}, \quad L_{\alpha\beta}\equiv 2C_{[\alpha\beta]}. \quad (3.21)$$

The $L_{\alpha\beta}$ are affected by the particular choice of infinitesimal gauge. This may be seen by considering a gauge transformation near the identity. Setting $N^\sigma{}_\alpha=\delta^\sigma{}_\alpha+\delta N^\sigma{}_\alpha$, (2.6) requires

$$\delta N_{(\sigma\alpha)}=0. \quad (3.22a)$$

Performing such a transformation on the structure defined by (3.15) and (3.19) gives

$$\delta e'^\sigma=(C^\sigma{}_\alpha+\delta N^\sigma{}_\alpha)\,e^{(0)\alpha}. \quad (3.22b)$$

This indicates a choice of $\delta N^\sigma{}_\alpha$ can be made which either cancels $L^\sigma{}_\alpha$ altogether, or some particular portion. Chrzanowski[30] and Demiański[31] have both made the same choice of $\delta N^\sigma{}_\alpha$ for fairly clear physical reasons. However, in the later discussion we will consider some reasons for other choices as well. In all cases, however, the six degrees of freedom of $C_{\alpha\beta}$ corresponding to $L_{\alpha\beta}$ can be completely determined solely by choice of the infinitesimal gauge transformation, reducing any question of the determination of $C_{\alpha\beta}$ to just determining $H_{\alpha\beta}$.

The restrictions on $N^\sigma{}_\alpha$ imposed by (3.22a) allow six (real) degrees of freedom to the Lorentz transformations. Since specific forms of various allowed gauges will be of considerable use in the later discussions, it is convenient, at this point, to present the details of the allowable gauge transformations from the point of view of null tetrads. The degrees of freedom may be specified by three complex parameters σ, ρ, η, which generate three independent sets of gauge transformations[32]:

σ-gauge: $e'^1 = (e^{2i\mathrm{Im}\sigma})e^1, \quad e'^2 = (e^{-2i\mathrm{Im}\sigma})e^2,$

$$e'^3 = (e^{2\mathrm{Re}\sigma})e^3, \quad e'^4 = (e^{2\mathrm{Re}\sigma})e^4, \tag{3.23a}$$

or $N = [\exp 2i(\sigma\mathscr{L}_0 - \bar{\sigma}\mathscr{L}_{\bar{0}})]$,

ρ-gauge: $e'^1 = e^1 + \bar{\rho}e^3, \quad e'^2 = e^2 + \rho e^3,$

$$e'^3 = e^3, \quad e'^4 = e^4 - \rho e^1 - \bar{\rho}e^2 - \rho\bar{\rho}e^3 \tag{3.23b}$$

or $N = \exp[-\sqrt{2}i(\rho\mathscr{L}_+ - \bar{\rho}\mathscr{L}_{\bar{+}})]$,

η-gauge: $e'^1 = e^1 - \eta e^4, \quad e'^2 = e^2 - \bar{\eta}e^4,$

$$e'^3 = e^3 + \bar{\eta}e^1 + \eta e^2 - \eta\bar{\eta}e^4, \quad e'^4 = e^4, \tag{3.23c}$$

or $N = \exp[+\sqrt{2}i(\eta\mathscr{L}_- - \bar{\eta}\mathscr{L}_{\bar{-}})]$.

Note that, in the infinitesimal case, the associated quantities $\delta N_a = \mathscr{L}_a{}^{\alpha\beta}\delta N_{\alpha\beta}$ may easily be read off from the exponential forms of the corresponding matrices,

$$\delta N_a = i\sqrt{2}(-\delta\eta, \sqrt{2}\delta\sigma, \delta\rho). \tag{3.23d}$$

It is also useful here to list the behavior of the connections and conformal tensor under these transformations. These are most efficiently specified by noting that

$$C'_T = N_T{}^B C_B, \tag{3.24a}$$

where $N_T{}^B$ is a $D(0,2)$ representation of the original transformation, while

$$\Gamma'_{t\tau}e'^\tau = \Gamma'_t = N_t{}^a\Gamma_a + X_t = (N_t{}^a N_\tau{}^\alpha\Gamma_{a\alpha} + X_{t\tau})e^\tau, \tag{3.24b}$$

with $N_t{}^a$ corresponding to the $D(0,1)$ representation. These matrices are given by:

$$\sigma\text{-gauge: } N_t{}^a = \begin{pmatrix} e^{2\sigma} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & e^{-2\sigma} \end{pmatrix}, \quad X_t = \begin{pmatrix} 0 \\ -\sqrt{2}d\sigma \\ 0 \end{pmatrix}, \tag{3.25a}$$

$$N_T{}^A = \mathrm{diag}\{e^{4\sigma}, e^{2\sigma}, 1, e^{-2\sigma}, e^{-4\sigma}\},$$

$$\rho\text{-gauge: } N_t{}^a = \begin{pmatrix} 1 & 0 & 0 \\ \sqrt{2}\rho, & 1 & 0 \\ \rho^2, & \sqrt{2}\rho, & 1 \end{pmatrix}, \quad X_t = \begin{pmatrix} 0 \\ 0 \\ -d\rho \end{pmatrix}, \tag{3.25b}$$

$$N_T{}^A = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 2\rho, & 1 & 0 & 0 & 0 \\ \sqrt{6}\rho^2, & \sqrt{6}\rho, & 1 & 0 & 0 \\ 2\rho^3, & 3\rho^2, & \sqrt{6}\rho, & 1 & 0 \\ \rho^4, & 2\rho^3, & \sqrt{6}\rho^2, & 2\rho, & 1 \end{pmatrix},$$

$$\eta\text{-gauge: } N_t{}^a = \begin{pmatrix} 1, & \sqrt{2}\eta, & \eta^2 \\ 0 & 1, & \sqrt{2}\eta \\ 0 & 0 & 1 \end{pmatrix}, \quad X_t = \begin{pmatrix} d\eta \\ 0 \\ 0 \end{pmatrix}, \tag{3.25c}$$

$$N_T{}^A = \begin{pmatrix} 1, & 2\eta, & \sqrt{6}\eta^2, & 2\eta^3, & \eta^4 \\ 0 & 1, & \sqrt{6}\eta, & 3\eta^2, & 2\eta^3 \\ 0 & 0 & 1, & \sqrt{6}\eta, & \sqrt{6}\eta^2 \\ 0 & 0 & 0 & 1, & 2\eta \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

The $H_{\alpha\beta}$ are also subject to gauge transformations, but of a considerably different sort. These correspond to the usual infinitesimal coordinate transformations made in the standard theory of linearized gravitational waves. That is, the coordinates x^μ are replaced by a new set $x'^\mu = x^\mu + \xi^\mu$, where ξ is a first-order small quantity. This allows up to eight constraints to be placed on the nature of $H_{\alpha\beta}$.[33] For the time being, we will not specify any particular choice of this coordinate gauge.

The perturbations of the connections and curvature may now be determined, as functions of $C_{\alpha\beta}$. As usual, the connections may be calculated via 1-forms through the use of the first structure equations, (2.3), or via the basis of 1-vectors through the use of commutation coefficients given by

$$\omega_{\alpha\beta} = \tfrac{1}{2}(C_{\alpha\beta\gamma} + C_{\alpha\gamma\beta} + C_{\gamma\beta\alpha})e^\gamma, \quad [\partial_\alpha, \partial_\beta] = C_{\alpha\beta}{}^\gamma\partial_\gamma. \tag{3.26}$$

The $\delta\omega_{\alpha\beta}$ may be calculated directly, using the commutation coefficients. In order to proceed from the first structure equations, it is simplest to note that the exterior derivative is a nonmetric structure and therefore is unchanged from its background version. Some understanding of this observation can be motivated by looking at the exterior derivative acting on a scalar,

$$\begin{aligned} d\phi &= (\partial_\alpha\phi)e^\alpha \\ &= (\partial^{(0)}{}_\alpha\phi)e^{(0)\alpha} + (B^\beta{}_\alpha \\ &\quad + C^\beta{}_\alpha)(\partial^{(0)}{}_\beta\phi)e^{(0)\alpha} \\ &= d^{(0)}\phi, \end{aligned}$$

where the last equality is obtained by use of (3.19b). Either approach leads to

$$\delta \omega_{\alpha\beta} = H_{\gamma[\alpha;\beta]} e^{(0)\gamma} - \tfrac{1}{2} D^{(0)} L_{\alpha\beta}, \tag{3.27a}$$

as well as

$$D^{(0)} \delta e^{\alpha} = e^{(0)\beta} \wedge \delta \omega^{\alpha}{}_{\beta}, \tag{3.27b}$$

where the covariant derivatives are naturally with respect to the background metric. It is to be noted that the perturbed components are expressed by

$$\delta \omega_{\alpha\beta} = \delta \omega_{\alpha\beta\gamma} e^{(0)\gamma} + \omega^{(0)}{}_{\alpha\beta\gamma} \delta e^{\gamma}. \tag{3.28}$$

Continuing, one easily finds that

$$\begin{aligned} \delta \Omega_{\alpha\beta} &= D^{(0)} \delta \omega_{\alpha\beta} \\ &= H_{\delta[\alpha;\beta]\gamma} e^{(0)\gamma} \wedge e^{(0)\delta} + 2\Omega^{(0)}{}_{[\alpha}{}^{\gamma} L_{\beta]\gamma}, \end{aligned} \tag{3.29}$$

and, since

$$\delta \Omega_{\alpha\beta} = \tfrac{1}{2}(\delta R_{\alpha\beta\gamma\delta} e^{(0)\gamma} \wedge e^{(0)\delta} + 2R^{(0)}{}_{\alpha\beta\gamma\delta} e^{(0)\gamma} \wedge \delta e^{\delta}), \tag{3.30}$$

it follows that

$$\delta R_{\alpha\beta\gamma\delta} = H_{\beta[\gamma;\delta]\alpha} - H_{\alpha[\gamma;\delta]\beta} + R_{\gamma\delta\epsilon[\alpha} H^{\epsilon}{}_{\beta]} + 2R_{\alpha\beta[\gamma}{}^{\epsilon} L_{\delta]\epsilon} + 2R_{\gamma\delta[\alpha}{}^{\epsilon} L_{\beta]\epsilon}. \tag{3.31}$$

The set of Eqs. (3.28) and (3.31) completely determines the connections and curvature in terms of the perturbations $C_{\alpha\beta}$ of the tetrad, and are thus equivalent, for example, to those given in Ref. 30. Since they are covariant in form and have had no prior gauge conditions built into them, they can be easily adapted to whatever problem is desired. However, when determining the complete perturbed geometry through the use of these equations, one must take many (arbitrary) gauge conditions into account. A more direct determination of the physical degrees of freedom of the perturbation is therefore desirable. This is accomplished, following Teukolsky,[4] by directly calculating the perturbations of the Riemann tensor and then inverting the equations above to evaluate the perturbed tetrad, if desired.[34] We proceed to the alternate approach by writing the first-order perturbation of (3.13), accounting for (3.14), and suppressing, usually, the superscript (0) for background quantities,

$$\begin{aligned} &D^{\alpha} D_{\alpha} \delta C_A + [D^{\alpha} \delta D_{\alpha} + (\delta D^{\alpha}) D_{\alpha}] C_A \\ &\quad - \tfrac{1}{2}\delta(R C_A) + 6 h_{ABC} C^B \delta C^C = -4\delta M_A. \end{aligned} \tag{3.32}$$

Every term contains a perturbed quantity and, therefore, all other entries in that term must be from the background metric. Note that, for example,

$$(\delta D_{\alpha}) C^A = (\delta \partial_{\alpha}) C^A + (\delta \mathscr{W}^A{}_{B\alpha}) C^B, \tag{3.33}$$

and the simplified notation

$$C^A{}_{,\delta\alpha} \quad \text{instead of} \quad (\delta \partial_{\alpha}) C^A \tag{3.34}$$

will occur quite often.

In general, (3.32) constitutes a set of five coupled, second-order, nonlinear partial differential equations for the desired δC_A. There are certain sufficient conditions which allow decoupling of the most interesting of the equations—the ones with extremal values of the helicity, $A = ++$ or $--$. It is believed that these conditions are in fact necessary as well,[9] but we do not have a complete proof. To determine sufficient conditions suppose first that the background metric is *algebraically special*—that there is a multiple Debever–Penrose null vector. It is desirable to give labels to two particular allowed modes of alignment of the tetrad with one or more multiple Debever–Penrose vectors. By choosing our tetrad so that e^3 is aligned with this vector, the Sachs–Goldberg theorem guarantees that

$$\Gamma_{+2} = 0 = \Gamma_{+4}, \quad C_{++} = 0 = C_{+} \text{—positive alignment of tetrad.} \tag{3.35}$$

On the other hand, if we align e^4 with a multiple Debever–Penrose vector, we obtain

$$\Gamma_{-1} = 0 = \Gamma_{-3}, \quad C_{--} = 0 = C_{-} \text{—negative alignment of tetrad.} \tag{3.36}$$

We will show that the former alignment guarantees the decoupling of the δC_{++} equation, while the latter insures that the δC_{--} equation decouples. From now on we continue to assume that the space–time under consideration is algebraically special, thereby assuring that (at least) one of the wave equations can be decoupled. Both will decouple only if there are two distinct multiple Debever–Penrose vectors, which requires a background metric whose conformal tensor is either of type D or conformally flat.

Next write out Eq. (3.32) for the case $A = ++$, utilizing (2.34) to determine the $V(0,2)$ connections in terms of the quantities Γ_a defined in (2.20), and assuming a positive alignment of the tetrad [(3.35)],

$$\begin{aligned} &(\Box - \tfrac{1}{2}R)\delta C_{++} + 4(\sqrt{2}\,\Gamma_0{}^{\alpha} \delta C_{++,\alpha} - \Gamma_{+}{}^{\alpha} \delta C_{+,\alpha}) \\ &\quad + 2(\sqrt{2}\,\Gamma_0{}^{\alpha}{}_{;\alpha} - 2\Gamma_{+}{}^{\alpha}\Gamma_{-\alpha} + 4\Gamma_0{}^{\alpha}\Gamma_{0\alpha} \\ &\quad - \tfrac{1}{2}\sqrt{6}\,C_0)\delta C_{++} - 2(\Gamma_{+}{}^{\alpha}{}_{;\alpha} + 3\sqrt{2}\,\Gamma_{+}{}^{\alpha}\Gamma_{0\alpha})\delta C_{+} \\ &\quad + 4\sqrt{6}\,C_0 \Gamma_{+}{}^{\alpha} \delta\Gamma_{+\alpha} = -4\delta M_{++}. \end{aligned} \tag{3.37}$$

The procedure to decouple this equation—eliminate terms involving δC_{+} or $\delta\Gamma_{+}$—is quite straightforward. We refer back to the Bianchi equations as given by Eqs. (3.7) and perturb the appropriate ones. The combination $\omega_{+}\cdot\delta B_{+} = 0$, evaluated for the positive alignment conditions, gives

$$\begin{aligned} -4\Gamma_{+}{}^{\alpha}\delta C_{+,\alpha} &= 16\Gamma_{+[1}\delta C_{++,3]} + 8(\Gamma_{+}{}^{\alpha}\Gamma_{-\alpha} \\ &\quad + 4\sqrt{2}\,\Gamma_{+[1}\Gamma_{03]})\delta C_{++} + 4\sqrt{2}\,\Gamma_{+}{}^{\alpha}\Gamma_{0\alpha}\delta C_{+} \\ &\quad - 4\sqrt{6}\,C_0\Gamma_{+}{}^{\alpha}\delta\Gamma_{+\alpha} + 16\Gamma_{+}{}^{\alpha}\delta(D_{[4}M_{2]\alpha}). \end{aligned} \tag{3.38}$$

Inserting this into (3.37) eliminates several undesirable terms, leaving[35]

$$(\Box - \tfrac{1}{2}R)\delta C_{++} + 4(\sqrt{2}\Gamma_0{}^\alpha \delta C_{++,\alpha} + 4\Gamma_{+[1}\delta C_{++,3]}) + 2(\sqrt{2}\Gamma_0{}^\alpha{}_{;\alpha} + 2\Gamma_+{}^\alpha\Gamma_{-\alpha} + 16\sqrt{2}\Gamma_{+[1}\Gamma_{03]} + 4\Gamma_0{}^\alpha\Gamma_{0\alpha} - \tfrac{1}{2}\sqrt{6}C_0)\delta C_{++} = -4[\delta M_{++} + 4\Gamma_+{}^\alpha\delta(D_{[4}M_{2]\alpha})]. \quad (3.39)$$

The remaining term in δC_+ has not been written since the coefficient, after inserting (3.38), simply becomes $\Gamma_+{}^\alpha{}_{;\alpha} + \sqrt{2}\,\Gamma_0{}^\alpha\Gamma_{+\alpha}$, which is equal to C_+ in the background and vanishes. Therefore, at least in the absence of the source terms generated by $M_{\alpha\beta}$ and $\delta M_{\alpha\beta}$, the equation is now decoupled, containing only δC_{++} and known quantities.

By writing out (3.32) for the case $A = --$, assuming the negative alignment conditions specified by (3.36) and inserting, from (3.7), the conditions specified by $\omega_-\cdot\delta B_-$, the very similar, decoupled, equation for C_{--} is obtained,

$$(\Box - \tfrac{1}{2}R)\delta C_{--} - 4(\sqrt{2}\Gamma_0{}^\alpha \delta C_{--,\alpha} - 4\Gamma_{-[2}\delta C_{--,4]}) + 2(-\sqrt{2}\Gamma_0{}^\alpha{}_{;\alpha} + 2\Gamma_+{}^\alpha\Gamma_{-\alpha} - 16\sqrt{2}\Gamma_{-[2}\Gamma_{04]} + 4\Gamma_0{}^\alpha\Gamma_{0\alpha} - \tfrac{1}{2}\sqrt{6}C_0)\delta C_{--} = -4[\delta M_{--} + 4\Gamma_-{}^\alpha\delta(D_{[3}M_{1]\alpha})]. \quad (3.40)$$

An exactly analogous procedure will also decouple the electromagnetic wave equations, (3.4). In this instance the geometry is held fixed, but satisfying the conditions (3.35) or (3.36), while δF_a is conceived of as a test field which does not modify the geometry. (That is, the resultant equations are the usual electromagnetic wave equations in a curved space.) Writing out (3.4) for $a = +$ with δF_a for F_a gives the appropriate result. The use of the constraint given by (3.35) allows the coefficient of δF_0 to be made equal to C_+, which vanishes, while the use of $\omega_+\cdot\delta J$, calculated from the perturbed Maxwell equations [(3.2) with δF_b written for F_b] allows the elimination of terms in $\Gamma_+{}^\alpha\delta F_{0,\alpha}$, leaving the decoupled equation for δF_+,

$$(\Box - \tfrac{1}{3}R)\delta F_+ + 2(\sqrt{2}\Gamma_0{}^\alpha F_{+,\alpha} + 4\Gamma_{+[1}\delta F_{+,3]}) + (\sqrt{2}\Gamma_0{}^\alpha{}_{;\alpha} + 2\Gamma_+{}^\alpha\Gamma_{-\alpha} + 2\Gamma_0{}^\alpha\Gamma_{0\alpha} + 8\sqrt{2}\Gamma_{+[1}\Gamma_{03]} - C_0/\sqrt{6})\delta F_+ = -8\sqrt{2}i\pi(\delta J_{[2;4]} + \Gamma_+{}^\alpha\delta J_\alpha). \quad (3.41)$$

Again, using (3.4) with $a = -$, the equation for C_-, the constraint (3.36) and $\omega_-\cdot\delta J$ calculated from (3.2) gives the decoupled equation for δF_-,

$$(\Box - \tfrac{1}{3}R)\delta F_- - 2(\sqrt{2}\Gamma_0{}^\alpha \delta F_{-,\alpha} - 4\Gamma_{-[2}\delta F_{-,4]}) + (-\sqrt{2}\Gamma_0{}^\alpha{}_{;\alpha} + 2\Gamma_0{}^\alpha\Gamma_{0\alpha} + 2\Gamma_+{}^\alpha\Gamma_{-\alpha} - 8\sqrt{2}\Gamma_{-[2}\Gamma_{04]} - C_0/\sqrt{6})\delta F_- = -8\sqrt{2}i\pi(\delta J_{[1;3]} - \Gamma_-{}^\alpha\delta J_\alpha). \quad (3.42)$$

It is often useful to consider gauge transformations as an aid to simplifying these perturbed equations. Such transformations can be considered on both the original manifold and the perturbed manifold. It is easily seen that any ρ-gauge or η-gauge transformation in the background space will interfere with the important conditions for alignment of the tetrad, (3.35) or (3.36). Therefore, only σ-gauge transformations are allowed, on the background manifold. In Sec. 5 it is shown that such (finite) σ-gauge transformations are in fact necessary to reduce the equations to separable form.

The infinitesimal gauge transformations generated by the $N_{\alpha\beta}$ [see (3.23)] affect only the perturbed quantities and therefore are of considerable use, especially when the background geometry is of type D. In that case, an infinitesimal ρ-gauge leaves invariant all the δC_A with the exception of δC_-, while an infinitesimal η-gauge perturbs only δC_+.[36] Therefore, under a combined $\delta\rho$, $\delta\eta$-gauge, the effect on the components of the conformal tensor is

$$\delta C'_{++} = \delta C_{++}, \quad \delta C'_0 = \delta C_0, \quad \delta C'_+ = \delta C_+ + \sqrt{6}C_0\delta\eta,$$
$$\delta C'_{--} = \delta C_{--}, \quad \delta C'_- = \delta C_- + \sqrt{6}C_0\delta\rho. \quad (3.43)$$

Since $\delta\eta$ and $\delta\rho$ are arbitrary first-order quantities at our disposal and C_0 is different from zero, these gauge transformations may be used to cause $\delta C'_+$ and $\delta C'_-$ to vanish, which determines certain of the $L_{\alpha\beta}$. The ability to choose $\delta C'_+$ and $\delta C'_-$ as zero can be used to simplify both the $\delta C'_0$ wave equation and the route backward toward determination of the perturbations to the connections and tetrad. The behavior of the δC_A under (finite) σ-gauge transformations is also needed in Sec. 5. Therefore, consider the gauge-transformation determined by $\sigma^{(0)} + \delta\sigma$. It is easily seen that

$$\delta C'_A = \exp(2A\sigma^{(0)})\delta C_A + 2A\,C_A{}^{(0)}\delta\sigma, \quad (3.44)$$

where there is no sum on A in the equation. In a background geometry of type D, the second term vanishes in every case. As a result of this, the simple transformation equation

$$\delta C'_A = \exp(2A\sigma^{(0)})\delta C_A, \quad A = ++, +, 0, -, -- \quad \text{(background canonical type } D), \quad (3.45)$$

is acquired.

4. SPINORIAL APPROACH [FOR ARBITRARY $D(0,s)$ MASSLESS FIELDS]

The main emphasis of this work is the gravitational (and electromagnetic) wave equations. However, for reasons of completeness and esthetic interest, we would like to include neutrinos as well; there is also some interest in particles of spin 3/2 generated by work on supergravity.[37] If the reader has more interest in the gravitational case he can skip this section and proceed to the question of separability discussed in Sec. 5.

We utilize the usual spinor spaces[38] which correspond to the representation spaces $V(0,\tfrac{1}{2})$, whose elements are denoted ζ^M, and $V(\tfrac{1}{2},0)$, whose elements are denoted $\zeta^{\dot N}$, where these capital Latin indices take on the values 1 and 2. [Be cautious so as not to confuse these indices and the ones denoting components of an element of $V(0,2)$, used in other sections of this work.] As is well known, the metric induced on these spaces can be represented by the two-dimensional Levi-Civita symbol

$$\epsilon^{MN}=\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \tag{4.1}$$

and its complex conjugate $\epsilon^{\dot M \dot N}$. Indices are raised and lowered by the rules

$$\zeta_M=\epsilon_{MN}\zeta^N, \quad \zeta^N=\epsilon^{MN}\zeta_M. \tag{4.2}$$

The homomorphism between $D(0,\frac{1}{2}) \otimes D(\frac{1}{2},0)$ and the usual tensorial representation of O(3,1) is expressed via the Pauli matrices

$$\sigma^{M\dot N}=\sqrt{2}\begin{pmatrix} e^4 & e^2 \\ e^1 & -e^3 \end{pmatrix}. \tag{4.3}$$

The useful relation between the Pauli matrices and the important projection operators $\mathscr{Z}^a$ is

$$\mathscr{Z}^a=\frac{i}{4\sqrt{2}}Q^a{}_{MN}\sigma^{M\dot P}\wedge\sigma^N{}_{\dot P}, \tag{4.4}$$

where the $Q^a{}_{MN}$ are just the vector addition coefficients, $C(\frac{1}{2},\frac{1}{2},1;M,N,a)$,

$$Q^a{}_{MN}=\begin{pmatrix} \delta^a_+ & \frac{1}{\sqrt{2}}\delta^a_0 \\ \frac{1}{\sqrt{2}}\delta^a_0 & \delta^a_- \end{pmatrix}. \tag{4.5}$$

As usual the covariant derivative into these spaces is determined by insisting that $D\sigma^{M\dot N}=0$. This generates the spinorial connections

$$\Gamma_{MN}=\begin{pmatrix} \Gamma_+, & \frac{1}{\sqrt{2}}\Gamma_0 \\ \frac{1}{\sqrt{2}}\Gamma_0, & \Gamma_- \end{pmatrix}, \tag{4.6}$$

as well as the complex conjugate quantities $\Gamma_{\dot M\dot N}$.

The conformal tensor is just determined by a fourth-order spinorial quantity, corresponding to $D(0,\frac{1}{2})^4 \supset D(0,2)$, C_{MNPL}, given by

$$C_{1111}=-\tfrac{1}{2}C_{++}, \quad C_{1222}=-\tfrac{1}{4}C_-,$$

$$C_{1112}=-\tfrac{1}{4}C_+, \quad C_{2222}=-\tfrac{1}{2}C_{--}, \quad C_{1122}=-\frac{1}{2\sqrt{6}}C_0. \tag{4.7}$$

If $\Psi_{M_1\cdots M_{2s}}$ is an element of $V(0,s)$, then the usual massless field equations[39] for any spin $s=0, \frac{1}{2}, 1, \frac{3}{2}, 2,\cdots$, may be written most easily using $\nabla^{M\dot N}\equiv\sigma^{M\dot N}\cdot D$ such that

$$\nabla^P{}_{\dot N}\Psi_{PM_1\cdots M_{2s-1}}=0, \tag{4.8}$$

where the zero on the right hand side indicates an explicit omission of possible source terms. In the case of $s=1$ (Maxwell's equations) or $s=2$ (the gravitational field equations), appropriate sources have already been included in Sec. 3.

Since, for other spins, choices of source are not too easily made, sources are excluded from consideration in this section, as well as most of the next section. As is well known[40] this equation is subject to the constraint

$$(s-\tfrac{1}{2})(s-1)\,C^{MNP}{}_{(L_1}\Psi_{L_2\cdots L_{2s})MNP}=0, \tag{4.9}$$

which couples the massless field in question with the gravitational field of the background space for spins other than $s=0,\frac{1}{2}$, or 1. [It is possible that the field equations (4.8) should be modified in order to avoid this coupling but no satisfactory method is yet known.]

In a manner completely analogous to that carried out in Sec. 2, (4.8) can be used to generate the wave equation

$$\left(D^\alpha D_\alpha-\frac{s+1}{6}R\right)\Psi_{M_1\cdots M_{2s}}$$

$$+2(2s-1)\,C^{NP}{}_{M_1(M_2}\Psi_{M_3\cdots M_{2s})NP}=0. \tag{4.10}$$

Before proceeding to the separation of these equations, two special cases should be noted. If one takes $s=0$, (4.10) becomes simply

$$(D^\alpha D_\alpha-\tfrac{1}{6}R)\Psi=0 \quad s=0. \tag{4.11}$$

This has the additional term, $-R/6$, which is unconventional for a scalar ($s=0$) wave equation. Our approach has resulted in the additional factor because the technique guaranteed, *ab initio*, the conformal invariance of the equations.[41] It will be shown, in the following section, that this additional term is essential for the equation to be separable in PD coordinates.

The neutrino wave equation ($s=\frac{1}{2}$) is also very simple,

$$(D^\alpha D_\alpha-\tfrac{1}{4}R)\Psi_M=0, \tag{4.12}$$

even though all higher spins involve a coupling with the spinorial conformal tensor. Equations (4.12) are not yet decoupled due to the fact that the (covariant) operator D_α mixes components.

The separation of the wave equations (4.10) is initiated by viewing them as the expression appropriate for a perturbing (test) field in a background geometry. This requires only the replacement of $\Psi_{M_1\cdots M_{2s}}$ by $\delta\Psi_{M_1\cdots M_{2s}}$ except in the case of $s=2$. For the gravitational case, where the Ψ_{MNPL} are the C_{MNPL}, a perturbation of the Ψ_{MNPL} generates an additional perturbation of the C_{MNPL} in the coupling term. The perturbed equation may be written as

$$\left(D^\alpha D_\alpha-\frac{s+1}{6}R\right)\delta\Psi_{M_1\cdots M_{2s}}$$

$$+2A_s C^{NP}{}_{M_1(M_2}\delta\Psi_{M_3\cdots M_{2s})NP}=0, \tag{4.13}$$

$$A_s=\begin{cases} 2(2s-1)=6, & s=2\text{(gravity)},\\ 0, & s=0,\tfrac{1}{2},\\ (2s-1), & \text{otherwise}.\end{cases}$$

Denoting $\delta\Psi_{11\cdots1}$ by $\delta\Psi_s$ and $\delta\Psi_{22\cdots2}$ by $\delta\Psi_{-s}$, or the pair by $\delta\Psi_h$, $h=\pm s$, (4.13) may be written out for the case of maximal helicity,

$$\left(\Box-\frac{s+1}{6}R\right)\delta\Psi_s+4s\Gamma_{M1\alpha}\partial^\alpha\delta\Psi^M{}_{1\cdots1}$$

$$+2s(\Gamma_{M1}{}^\alpha{}_{;\alpha}+\Gamma^N{}_M{}^\alpha\Gamma_{N1\alpha})\delta\Psi^M{}_{1\cdots1}$$

$$+2s(2s-1)\Gamma^M{}_1{}^\alpha\Gamma^N{}_{1\alpha}\delta\Psi_{MN1\cdots1}$$

$$+2A_sC^{MN}{}_{11}\delta\Psi_{MN1\cdots1}=0. \quad (4.14)$$

and a very similar expression for $\delta\Psi_{-s}$. Following the technique applied in Sec. 3, the field equations themselves, (4.8), are used to eliminate from (4.14) the term $\Gamma_{11\alpha}\partial^\alpha\delta\Psi_{211\cdots1}$. This can only be accomplished when the condition (3.35) is met since only $\partial_4\delta\Psi_{211\cdots1}$ and $\partial_2\delta\Psi_{211\cdots1}$ are determined by the field equations in usable form. Next, the coefficients of $\delta\Psi_{211\cdots1}$ and $\delta\Psi_{221\cdots1}$ are determined to be exactly C_+ and C_{++} both of which vanish, by condition (3.35) again. This leaves the equation decoupled. In an exactly analogous fashion, with the use of condition (3.36), the equation for $\delta\Psi_{-s}$ is decoupled. The two may then be written as one equation,

$$\left(\Box-\frac{s+1}{6}R+\sqrt{2}h(2\Gamma_0{}^\alpha\partial_\alpha+\Gamma_0{}^\alpha{}_{;\alpha})\right.$$

$$+2s\left[-2\mathscr{L}_{-k}{}^{\alpha\beta}\Gamma_{k\alpha}(\partial_\beta+\sqrt{2}h\Gamma_{0\beta})\right.$$

$$\left.\left.+\Gamma_*^\alpha\Gamma_{-\alpha}+s\Gamma_0{}^\alpha\Gamma_{0\alpha}\right]-\frac{1}{\sqrt{6}}A_sC_0\right)\delta\Psi_h=0, \quad (4.15)$$

where $k\equiv h/s=\pm1$, $h=\pm s$. Equations (3.39), (3.40), (3.41), and (3.42) may be seen to be special cases of this equation for the appropriate values of s and h.

5. SEPARATION OF THE EQUATIONS

In the previous section it was shown that, if the background metric is of Petrov type D, both the extremal-helicity wave equations decouple from other unknown quantities for arbitrary massless fields of type $D\,(0,s)$ as well as for gravitational perturbations. However, Weir[17] has shown that all vacuum type D solutions are realized in the solutions of the Einstein–Maxwell equations found by Plebański and Demiański,[16] hereafter referred to as PD solutions. Further, Debever[15] has shown that, in addition to the obvious two Killing vectors which these solutions possess, all the PD solutions, including the nonvacuum ones, admit a second-rank conformal Killing tensor. Motivated by these facts, we now write these wave equations in PD coordinates and show that their solution can be reduced to the solution of a pair of (uncoupled) ordinary differential equations, only. This therefore generalizes Teukolsky's[4] results with the Kerr metric to an arbitrary PD metric which includes all those of vacuum D, as well as the possibility of a cosmological constant or an electric or magnetic charge. (In the case of an electric or magnetic charge, in principle the gravitational perturbations induce electromagnetic perturbations and vice versa. This coupled perturbation problem is one which we have not yet been able to solve. Here we are restricting our attention to dealing with perturbations of a single spin, only.)

The PD solutions depend on seven arbitrary parameters. In addition to the cosmological constant, the parameters group into three complex ones, $m+in$, $a+ib$, $e+ig$, which correspond to mass, NUT parameter, angular momentum per unit mass, acceleration parameter, electric charge, and magnetic charge. The more usual type D solutions, such as the Kerr metric, are recovered by performing specific limiting transitions which will be described later. The metric is described in coordinates (p,q,σ,τ) by[42]

$$ds^2=\phi^{-2}\left(\frac{p^2+q^2}{\mathscr{P}}dp^2+\frac{\mathscr{P}}{p^2+q^2}(d\tau+q^2d\sigma)^2\right.$$

$$\left.+\frac{p^2+q^2}{\mathscr{Q}}dq^2-\frac{\mathscr{Q}}{p^2+q^2}(d\tau-p^2d\sigma)^2\right), \quad (5.1)$$

with the "conformal factor"

$$\phi\equiv(1-pq). \quad (5.2)$$

The functions $\mathscr{P}$ and $\mathscr{Q}$ are fourth order polynomials in p and q, respectively,

$$\mathscr{P}=-(\lambda/6+e^2+\gamma)p^4+2mp^3-\epsilon p^2$$

$$+2np-(\lambda/6+g^2-\gamma), \quad (5.3a)$$

$$\mathscr{Q}=+(-\lambda/6+g^2-\gamma)q^4-2nq^3+\epsilon q^2$$

$$-2mq+(-\lambda/6+e^2+\gamma), \quad (5.3b)$$

where ϵ and γ are related to a and b by

$$\epsilon=-\frac{1}{ab}\left(\frac{a^2-b^2}{a^2+b^2}\right)\left(1-\frac{\lambda}{3}(a^2+b^2)\right)^{1/2},$$

$$\gamma=(a^2+b^2)^{-1}-\lambda/6. \quad (5.4)$$

In order to proceed further a choice of (null) tetrad is required. More than one such choice will be of considerable interest to us because the maximal and minimal helicity equations do not separate in the same gauge. We therefore present three choices of tetrad which are needed. Define a tetrad $e^{*\alpha}$ such that $-e^{*\alpha}$ is the tetrad given in Ref. 16 by Plebański and Demiański. The tetrad used most often in our work, denoted by e^α, is obtained from $e^{*\alpha}$ by a σ-gauge transformation with σ chosen as Λ defined by

$$e^{2\Lambda}\equiv(q+ip)(2/\mathscr{Q})^{1/2}. \quad (5.5)$$

Explicitly, the tetrad e^α is then

$$e^1=\frac{1}{\sqrt{2}}\phi^{-1}\left(\frac{q+ip}{\sqrt{\mathscr{P}}}dp+i\frac{\sqrt{\mathscr{P}}}{q+ip}(d\tau+q^2d\sigma)\right)=\overline{e^2},$$

$$e^3=\phi^{-1}\left(\frac{p^2+q^2}{\mathscr{Q}}dq+d\tau-p^2d\sigma\right), \quad (5.6)$$

$$e^4=\tfrac{1}{2}\phi^{-1}\left(dq-\frac{\mathscr{Q}}{p^2+q^2}(d\tau-p^2d\sigma)\right),$$

while the inverse tetrad (basis of the tangent space) is

$$\partial_1=\sqrt{\frac{\mathscr{P}}{2}}\,\frac{\phi}{q+ip}\Big(\partial_p-\frac{i}{\mathscr{P}}(\partial_\sigma+p^2\partial_\tau)\Big)=\overline{\partial_2},$$

$$\partial_3=\frac{\phi\,\mathscr{Q}/2}{p^2+q^2}\Big(\partial_q-\frac{1}{\mathscr{Q}}(\partial_\sigma-q^2\partial_\tau)\Big), \tag{5.7}$$

$$\partial_4=\phi\Big(\partial_q+\frac{1}{\mathscr{Q}}(\partial_\sigma-q^2\partial_\tau)\Big),$$

where by ∂_p, etc., we mean $\partial/\partial p$, etc., as usual. This particular choice of gauge is motivated by the fact that, when the appropriate limit is taken, it becomes (modulo sign conventions) the tetrad used by Teukolsky and most others who considered similar problems. It is the factor $q+ip$ in the definition of $e^{2\Lambda}$ given in (5.5) that is important for this separation. The factor $\sqrt{2/\mathscr{Q}}$ is required only if the tetrad is to agree with that of Teukolsky. The connections can then be calculated to yield

$$\Gamma_+=\omega_{24}=\Big(\frac{\phi}{q+ip}-\phi_q\Big)e^1$$
$$+\Big(\frac{\mathscr{P}}{2}\Big)^{1/2}\frac{1}{q-ip}\Big(\phi_p-\frac{i\phi}{q+ip}\Big)e^3,$$
$$\Gamma_-=\omega_{13}=\frac{\mathscr{Q}/2}{p^2+q^2}\Big(\frac{\phi}{q+ip}-\phi_q\Big)e^2$$
$$+\Big(\frac{\mathscr{P}}{2}\Big)^{1/2}\frac{1}{q+ip}\Big(\phi_p-\frac{i\phi}{q+ip}\Big)e^4, \tag{5.8}$$
$$\omega_{12}=\Big(\frac{\mathscr{P}}{2}\Big)^{1/2}\Big[(q+ip)^{-1}\Big(\phi_p+\frac{i\phi}{q+ip}-\frac{1}{2}\phi\frac{\dot{\mathscr{P}}}{\mathscr{P}}\Big)e^1$$
$$-(q-ip)^{-1}\Big(\phi_p-\frac{i\phi}{q-ip}-\frac{1}{2}\phi\frac{\dot{\mathscr{P}}}{\mathscr{P}}\Big)e^2\Big]$$
$$-ip\phi\,\mathscr{Q}(p^2+q^2)^{-2}e^3,$$
$$\omega_{34}=i\Big(\frac{\mathscr{P}}{2}\Big)\phi\,[(q+ip)^{-2}e^1-(q-ip)^{-2}e^2]$$
$$+\tfrac{1}{2}\mathscr{Q}(p^2+q^2)^{-1}\Big(\phi_q-\phi\frac{\dot{\mathscr{Q}}}{\mathscr{Q}}+\frac{2q\phi}{p^2+q^2}\Big)e^3-\phi_q e^4,$$

where ϕ_p and ϕ_q mean $\partial\phi/\partial p$ and $\partial\phi/\partial q$, respectively, while the dot is used to designate differentiation of the function with respect to its argument (e.g., $\dot{\mathscr{P}}=d\mathscr{P}/dp$). As well, we have

$$-\sqrt{6}\,C_0=\frac{\phi^2}{p^2+q^2}\Big(\ddot{\mathscr{P}}-\frac{6i\dot{\mathscr{P}}}{q+ip}-\frac{12\mathscr{P}}{(q+ip)^2}$$
$$+\ddot{\mathscr{Q}}-\frac{6\dot{\mathscr{Q}}}{q+ip}+\frac{12\mathscr{Q}}{(q+ip)^2}\Big) \tag{5.9}$$

and

$$R=\frac{-\phi^2}{p^2+q^2}\Big[\ddot{\mathscr{P}}-6\dot{\mathscr{P}}\frac{\phi_p}{\phi}+12\mathscr{P}\Big(\frac{\phi_p}{\phi}\Big)^2$$
$$+\ddot{\mathscr{Q}}-6\dot{\mathscr{Q}}\frac{\phi_q}{\phi}+12\mathscr{Q}\Big(\frac{\phi_q}{\phi}\Big)^2\Big]. \tag{5.10}$$

The equations for maximal helicity ($A=++$ for C_A, for example) will be found to separate in this tetrad while the equations for minimal helicity are, instead, separable in the tetrad e''^{α} obtained from $e^{*\alpha}$ by the inverse of the previous σ-gauge transformation, i.e., with σ chosen as $\Lambda''=-\Lambda$,

$$e^{-2\Lambda''}=(q+ip)(2/\mathscr{Q})^{1/2}. \tag{5.11}$$

Using (3.45) for the transformation properties of δC_A, it is easily shown that the gravitational wave equations are form invariant under σ-gauge transformations so that the equations need not be recalculated for the new tetrad choice.

Working in the tetrad e^α, the wave equations from Sec. 3 for both the gravitational and the electromagnetic cases could be written out in PD coordinates. However, we prefer here to use the equations from Sec. 4 which are valid for arbitrary massless, $D(0,s)$ fields, including gravitation simply as the case when $s=2$. This allows for maximal generality of the results. It will also permit the demonstration that, in fact, the desired wave equations only separate for certain values of s. Additionally the wave equations from Sec. 4 are sourceless, which is desirable in this section since the discussion involves separability. Separability of a given partial differential equation is a property of the differential operator and the coordinate system used, rather than of any (inhomogeneous) source term. When the source term is also separable one may proceed as before. But, in the common case of a nonseparable source term, a (separated) Green's function may be constructed from the sourceless equation, and then applied to the given source term.

Because the maximal and minimal helicity equations separate in different tetrads, they are treated separately. We begin with the case of helicity $h=+s$ and rewrite (4.15) for that case denoting the desired perturbation by $\delta\Psi_{+s}$ (δC_{++} for $s=2$, δF_+ for $s=1$, etc.):

$$\Big(\Box-\frac{s+1}{6}R+s[2\sqrt{2}\Gamma_0{}^\alpha\partial_\alpha+8\Gamma_{+[1}\partial_{3]}$$
$$+\sqrt{2}\Gamma_0{}^\alpha{}_{;\alpha}+2\Gamma_+{}^\alpha\Gamma_{-\alpha}+2s(4\sqrt{2}\Gamma_{+[1}\Gamma_{03]}$$
$$+\Gamma_0{}^\alpha\Gamma_{0\alpha})]-\frac{1}{\sqrt{6}}A_sC_0\Big)\delta\Psi_{+s}=0, \tag{5.12}$$

where the constant A_s is defined in (4.13). This equation may now be rewritten in PD coordinates with the tetrad e^α, which results in the rather lengthy equation,

$$\frac{\phi^2}{q^2+p^2}\{\partial_p\mathscr{P}\partial_p+\partial_q\mathscr{Q}\partial_q+s\dot{\mathscr{Q}}\partial_q-2(s+1)\phi^{-1}(\phi_p\mathscr{P}$$

$$\times\partial_p+\phi_q\mathscr{Q}\partial_q)+B+(s+1+A_s)\ddot{\mathscr{P}}/6-(s\dot{\mathscr{P}})^2/4\mathscr{P}$$

$$+(4s+1+A_s)\ddot{\mathscr{Q}}/6-(s+1)\phi^{-1}[\phi_p\dot{\mathscr{P}}+(s+1)\phi_q\dot{\mathscr{Q}} \tag{5.13}$$

$$-(s+2)\phi^{-1}(\phi_p{}^2\mathscr{P}+\phi_q{}^2\mathscr{Q})]+[s(2s-1)-A_s]$$

$$\times(q+ip)^{-1}[i\dot{\mathscr{P}}+\dot{\mathscr{Q}}+2(q+ip)^{-1}$$

$$\times(\mathscr{P}-\mathscr{Q})]\}\delta\Psi_{+s}=0,$$

where B is used to denote those terms which have derivatives with respect to σ and τ,

$$B\equiv\mathscr{P}^{-1}(\partial_\sigma+p^2\partial_\tau)^2-\mathscr{Q}^{-1}(\partial_\sigma-q^2\partial_\tau)^2$$

$$+s(\dot{\mathscr{Q}}/\mathscr{Q})(\partial_\sigma-q^2\partial_\tau)+is(\dot{\mathscr{P}}/\mathscr{P})$$

$$\times(\partial_\sigma+p^2\partial_\tau)+4s(q-ip)\partial_\tau. \tag{5.14}$$

None of the coefficients in (5.13) involve σ or τ since ∂_σ and ∂_τ are Killing vectors for the background geometry. Therefore, we may look for solutions having the form of Fourier coefficients in these variables—a reasonable thing to do since the existence of these Killing vectors characterizes the background as axially symmetric and stationary. When acting on such Fourier coefficients, the term B will simply become a separable term, with no differential operators, to be added to the equation. However, since the "conformal factor" $\phi=1-pq$ is inherently nonseparable, as well as the terms with the factors of $(q+ip)^{-1}$, more work must be done.

The known conformal invariance of the equation suggests a "conformal transformation" of some power of ϕ. Therefore, we denote the operator in (5.13) by L and evaluate the new operator $\phi^{-l}L\phi^{+l}$. By looking at the first-order derivatives in this new operator, it is found that a necessary condition for separability (in these coordinates) is that $l=s+1$, which eliminates the (nonseparable) terms $-2(s+1)\phi^{-1}(\phi_p\mathscr{P}\partial_p+\phi_q\mathscr{Q}\partial_q)$. This new operator is then notably simpler,

$$\phi^{-(s+1)}L\phi^{s+1}=\frac{\phi^2}{q^2+p^2}\{\partial_p\mathscr{P}\partial_p+\partial_q\mathscr{Q}\partial_q$$

$$+s\dot{\mathscr{Q}}\partial_q+B+(s+1+A_s)\ddot{\mathscr{P}}/6-(s\dot{\mathscr{P}})^2/4\mathscr{P}$$

$$+(4s+1+A_s)\ddot{\mathscr{Q}}/6+[s(2s-1)-A_s](q+ip)^{-1}$$

$$\times[i\dot{\mathscr{P}}+\dot{\mathscr{Q}}+2(q+ip)^{-1}(\mathscr{P}-\mathscr{Q})]\}. \tag{5.15}$$

The $(q+ip)^{-1}$ terms are still not separable. However, their coefficient—$s(2s-1)-A_s$—vanishes for $s=0$, $\frac{1}{2}$, 1, and 2. We conclude that, at least in PD coordinates, the equation separates only for these values of the spin. It is worth noting that these are just exactly the values of the spin for which the usual massless field equations (4.8) (which we used) are valid without any constraint on the background curvature. That is, (4.9) imposes a coupling between the background geometry and the massless field in question which is usually ignored, but which is in fact satisfied for arbitrary type D background geometry only for spin $s=0$, $\frac{1}{2}$, 1, and the gravitational field itself, with $s=2$. It is therefore a very reasonable presumption that the fact that the constraint (4.9) is not generally satisfied is indeed the source of the nonseparability of the wave equations for the other spins.

From now on, we consider only the spins $s=0$, $\frac{1}{2}$, 1, and 2, and notice that, for all these spins, A_s can be written in the analytic form $s(2s-1)$. Therefore, (5.13) may be rewritten in the simplified form

$$(\phi^{-s-1}L\phi^{s+1})(\phi^{-s-1}\delta\Psi_{+s})$$

$$=\frac{\phi^2}{p^2+q^2}\{\partial_p\mathscr{P}\partial_p+\mathscr{Q}^{-s}\partial_q\mathscr{Q}^{s+1}\partial_q+B$$

$$+[(2s^2+1)\ddot{\mathscr{P}}+(s+1)(2s+1)\ddot{\mathscr{Q}}]/6 \tag{5.16}$$

$$-(s\dot{\mathscr{P}})^2/4\mathscr{P}\}(\phi^{-s-1}\delta\Psi_{+s})=0.$$

Separated solutions may now be obtained[43] by writing

$$\delta\Psi_{+s}=\phi^{s+1}R^+(q)S^+(p)\ e^{i\alpha\sigma}e^{i\omega\tau}. \tag{5.17}$$

This form separates (5.16) into the following two ordinary differential equations,

$$[\partial_p\mathscr{P}\partial_p+(2s^2+1)\ddot{\mathscr{P}}/6+4s\omega p-\mathscr{P}^{-1}$$

$$\times(\alpha+\omega p^2+\tfrac{1}{2}s\dot{\mathscr{P}})^2+A]S^+(p)=0, \tag{5.18a}$$

$$[\mathscr{Q}^{-s}\partial_q\mathscr{Q}^{s+1}\partial_q+(s+1)(2s+1)\ddot{\mathscr{Q}}/6+4is\omega q$$

$$+\mathscr{Q}^{-1}(\alpha-\omega q^2)(\alpha-\omega q^2+is\dot{\mathscr{Q}})-A]R^+(q)=0, \tag{5.18b}$$

where A is a separation constant. Note that this implies that R^+ and S^+ are, in principle, functions, parametrically, of s, α ω, and A, as well as their arguments. The general solution would then be of the form

$$\delta\Psi_{+s}=\phi^{s+1}\int d\alpha e^{i\alpha\sigma}f(\alpha)$$

$$\times\int d\omega e^{i\omega\tau}g(\omega)R^+(q;s,\alpha,\omega)S^+(p;s,\alpha,\omega). \tag{5.19}$$

The solutions of (5.18) depend upon a large number of parameters—$s,\alpha,\omega,A,m,n,\epsilon,\gamma,e,g,\lambda$—as well as the boundary conditions. Consequently, we do not say very much about them here. In the limit to the Kerr geometry, it is the function $S^+(p)$ which becomes the spin-weighted spherical harmonics, while in that limit, the equation for $R(q)$ must be solved numerically,[44] although some studies of its analytic properties have been made.[45] We do note that $\mathscr{P}$ must be positive in order for the signature of the metric to be correct.

However $\mathscr{P}$ goes to $-\infty$ for sufficiently large $|p|$, so that the allowed (physical) values of p are restricted to some *finite* region, in which $\mathscr{P} \geqslant 0$. Therefore, (5.18a) does constitute a well-defined Sturm-Liouville eigenvalue problem for A, and $S^{+}(p)$.

The equation for $h = -s$—for $\delta\Psi_{-s}$—could also be written out explicitly from (4.15), but recall that it does not separate in the tetrad in which we have been working. The tetrad e''^{α} defined by (5.11) has been chosen, however, so that not only is the equation for $\delta\Psi''_{-s}$ separable but it is identical to the one for $\delta\Psi_{+s}$ except for a change of sign of σ and τ. Therefore, defining

$$R^{-}(q;s,\alpha,\omega) = R^{+}(q;s,-\alpha,-\omega),$$
$$S^{-}(p;s,\alpha,\omega) = S^{+}(p;s,-\alpha,-\omega), \tag{5.20}$$

we have

$$\delta\Psi''_{-s} = \phi^{s+1} R^{-}(q) S^{-}(p) e^{i\alpha\sigma} e^{i\omega\tau}. \tag{5.21}$$

But, using (3.45), $\delta\Psi_{-s} = e^{-4sA}\delta\Psi''_{-s}$, so that

$$\delta\Psi_{-s} = e^{-4sA}\phi^{s+1} R^{-}(q) S^{-}(p) e^{i\alpha\sigma} e^{i\omega\tau}$$

$$= (q+ip)^{-2s}\phi^{s+1}\left(\frac{\mathscr{Q}}{2}\right)^{s} R^{-}(q) S^{-}(p) e^{i\alpha\sigma} e^{i\omega\tau}. \tag{5.22}$$

Another choice of tetrad in which the equation for $\delta\Psi_{-s}$ separates is the one defined by the σ-gauge transformation from e^{α} in which σ is chosen to be $\tilde{A}$ such that

$$e^{2\tilde{A}} = (q+ip)^{-2}. \tag{5.23}$$

This transformation is similar to the previous one, but excludes the factor $(2/\mathscr{Q})$, included earlier for reasons for symmetry. [As was already pointed out, it is the correct power of $(q-ip)$ which is important for separability.] In this case the equation determining $\delta\Psi_{-s}$ is the same as the earlier one for $\delta\Psi_{+s}$ except for the change of sign of s. Therefore, setting

$$\tilde{R}^{-}(q;s,\alpha,\omega) = R^{+}(q;-s,\alpha,\omega),$$
$$\tilde{S}^{-}(p;s,\alpha,\omega) = S^{+}(p;-s,\alpha,\omega), \tag{5.24}$$

and transforming back to our primary tetrad, we have[46]

$$\delta\Psi_{-s} = (q+ip)^{-2s}\phi^{s+1}\tilde{R}^{-}(q)\,\tilde{S}^{-}(p)\,e^{i\alpha\sigma}e^{i\omega\tau}. \tag{5.25}$$

These various choices of tetrad can all be shown to correspond[47] to different behavior of the solutions at "radial" infinity or on the event horizon. A particular choice is determined by the type of boundary conditions desired for a particular application.

The limiting procedure from a general PD solution to the Kerr geometry is described in Ref. 16; however, we repeat it here. One must make the following substitutions,

$$n,\lambda,e,g \to 0, \quad m \to c^{3}M, \quad \epsilon \to c^{2}, \quad \gamma \to c^{4}a^{2}, \quad q \to cr,$$
$$p \to -ca\cos\theta, \quad \sigma \to c^{-3}\varphi/a, \quad \tau \to c^{1}(-t + a\varphi), \tag{5.26}$$

and then take the limit as c goes to zero. This procedure generates the following substitutions:

$$\phi \to 1, \quad dq \to dr, \quad dp \to a\sin\theta d\theta, \quad d\sigma \to d\varphi/a,$$
$$d\tau \to -dt + ad\varphi, \partial_{q} \to \partial_{r}, \quad \partial_{p} \to (a\sin\theta)^{-1}\partial_{\theta},$$
$$\partial_{\sigma} \to a\partial_{\varphi} + a^{2}\partial_{t}, \partial_{\tau} \to -\partial_{t}, \mathscr{P} \to a^{2}\sin^{2}\theta,$$
$$\mathscr{Q} \to \Delta = (r^{2}-2Mr+a^{2}), p^{2}+q^{2} \to \Sigma$$
$$= (r^{2}+a^{2}\cos^{2}\theta), -(q+ip)^{-1} \to -(r-ia\cos\theta)^{-1} = \rho \tag{5.27}$$
$$\alpha \to a(m-a\omega), \quad A \to \lambda - a\omega(a\omega-2m),$$

which transform all of the equations in this section into the usual equations for the Kerr geometry.[4] Similar limits may be taken to acquire, for example, the Kerr–NUT metric, the accelerating C metric, or even the Kerr–Newman metric, provided the perturbations considered are purely gravitational or purely electromagnetic, as would be appropriate when the background electric charge is taken to be of the same order as the (gravitational) perturbation.

Let us now return to the pure gravitational, sourceless case and discuss the rest of the geometrical problem. As was indicated at (3.43), a choice of infinitesimal gauge can always be made so that $\delta C_{+} = 0 = \delta C_{-}$. Assuming that this has been done, all gauge freedom has been accounted for except that of σ-gauge transformations. Given then a determination of δC_{++} and δC_{--} from (5.17) and (5.22) above, the Bianchi equations may be used to determine $\delta\Gamma_{+2}$, $\delta\Gamma_{+4}$, $\delta\Gamma_{-1}$, and $\delta\Gamma_{-3}$—quantities which vanish in the background. In particular the equations δB_{+2}, δB_{-1} tell us that

$$\delta\Gamma_{+2} = (\tfrac{1}{2}\sqrt{6}C_{0})^{-1}(\partial_{3} + \Gamma_{-2} + 2\sqrt{2}\Gamma_{03})\delta C_{++},$$
$$\delta\Gamma_{-1} = (\tfrac{1}{2}\sqrt{6}C_{0})^{-1}(\partial_{4} + \Gamma_{+1} - 2\sqrt{2}\Gamma_{04})\delta C_{--}, \tag{5.28}$$

while the other pair may be obtained by the index substitutions $1 \leftrightarrow 3$, $2 \leftrightarrow -4$ (the minus sign indicates that every term with an index of 2 or 4 should have its sign changed as well).

The other connections are somewhat more difficult to determine. Wald[48] has given the procedure for showing that all perturbations for which either δC_{++} or δC_{--} vanish have the property that only δC_{0} is nonzero *and* that these perturbations amount only to perturbations of the seven parameters of the PD solution; i.e., that one ends up with only another PD solution. Since we have made no restraint on the generality of the PD solution in question, assuming $\delta C_{++} \neq 0$, for example, there is no loss of generality by assuming that δC_{0} is zero. That this is so may be seen by simply noting that any nonzero value of δC_{0} could always be transformed away by performing an additional perturbation with only δC_{0} nonzero, which would maintain our assumption of an arbitrary PD solution. The extra δC_{0}-perturbations can always be recovered later by allowing each of the seven parameters of the PD solution to acquire an extra infinitesimal portion. Notice therefore that the perturbation is now determined only by δC_{++}, δC_{--} and the perturbation of the seven background parameters.

Following this reasoning assume that δC_{0} vanishes. The other four perturbed Bianchi equations may then be solved for $\delta\Gamma_{+1}$, $\delta\Gamma_{+3}$, $\delta\Gamma_{-2}$, and $\delta\Gamma_{-4}$ in terms of the sixteen $C_{\alpha\beta}$. An example is $\delta\Gamma_{+1} = -(3C_0)^{-1}C_{0,\delta 4}$. However, (3.19) permits rewriting this equation in terms of the (unknown) coefficients $C_{\alpha\beta}$ since, e.g., $C_{0,\delta 4} = -B^{\beta}{}_{4}C_{0,\beta}$. The results are then

$$
\begin{aligned}
\tfrac{1}{2}\delta\Gamma_{+1} &= \Gamma_{+\,[3}C_{1]4} + \Gamma_{-\,[4}C_{2]4},\\
\tfrac{1}{2}\delta\Gamma_{+3} &= \Gamma_{+\,[1}C_{3]2} + \Gamma_{-\,[2}C_{4]2},\\
\tfrac{1}{2}\delta\Gamma_{-2} &= \Gamma_{+\,[3}C_{1]3} + \Gamma_{-\,[4}C_{2]3},\\
\tfrac{1}{2}\delta\Gamma_{-4} &= \Gamma_{+\,[1}C_{3]1} + \Gamma_{-\,[2}C_{4]1}.
\end{aligned}
\tag{5.29}
$$

The $\delta\Gamma_0$ may be determined in terms of the $\delta\Gamma_+$, $\delta\Gamma_-$, and $C_{\alpha\beta}$ by writing out the equations defining certain components of the perturbed Ricci tensor, which vanishes. In particular those $\delta\Omega_{a\dot{b}}$ and $\delta\Omega_{\dot{a}b}$ for which neither index is 0 involve the $\delta\Gamma_0$ algebraically only and may be solved explicitly. We do not write them as the forms are messy and complicated while the procedure is quite straightforward. Lastly one needs to determine the $C_{\alpha\beta}$ themselves. The degrees of freedom in L_{24} and L_{13}, are already determined by our choice of gauge such that $\delta C_+ = 0 = \delta C_-$. (This also requires that $\delta C_{\dot{+}}$ and $\delta C_{\dot{-}}$ vanish, determining L_{14} and L_{23}.) Therefore, there are only 12 components of $C_{\alpha\beta}$ to be determined.

The σ and τ dependence of $C_{\alpha\beta}$ must surely be of the form of Fourier coefficients, inherited from δC_{++} and δC_{--}, which implies that they are effectively functions of only two independent variables. Therefore, it would seem that the 24 first structure equations, (2.3), which are first order partial differential equations for the $C_{\alpha\beta}$ in terms of $\delta\Gamma_a$ and $\delta\Gamma_{\dot{a}}$, would suffice to determine the remaining $C_{\alpha\beta}$. In fact, however, the equations cannot be solved for the $\partial C_{\alpha\beta}/\partial q$ and $\partial C_{\alpha\beta}/\partial p$, since the associated matrix is singular. This is simply an indication of the fact that the $C_{\alpha\beta}$ are still subject to some gauge conditions so that they cannot be determined uniquely by just the δC_A. The work of Demiański[31] is an example showing a particular method by which this solution may be obtained, having imposed sufficiently many gauge conditions. Much more general, however, are the complete solutions to this problem for the Kerr metric (with a specific choice of gauge) recently obtained by Chandrasekhar.[49] We are attempting a general reduction of the problem to specify the separation of the gauge-dependent details from the essential ones, but have not completed this procedure.

6. CONCLUSIONS

The present work has the aim both of generalizing the work of Teukolsky on the separability of the wave equations for the conformal curvature of a perturbed space–time, and of putting the structure associated with this separability into a covariant format. The latter aim necessitated the extension of the usual covariant (tensorial) structures over a manifold into the bundle of representation spaces of the Lorentz group. This extension provided a compact, algebraic approach to the relevant equations, allowing them to be manipulated with relative ease. The result has been to show the decoupling, for an arbitrary massless (spinor) $D(0,s)$ field [or $D(s,0)$], of those wave equations corresponding to extremal helicities in an arbitrary type D background geometry. As well, it has been shown that, in such a background, the solution of these wave equations can be reduced to the solution of ordinary differential equations, in PD coordinates, only for the spins 0, $\frac{1}{2}$, 1, and 2.

In the important gravitational case of spin 2 we have shown that gauge conditions can be so chosen that the entire perturbed conformal tensor, δC_A, can be determined. Then we have indicated how these might be used to determine a general algorithmic path to the lower-order parts of the perturbed metric structure, which depend on yet more gauge choices, although this program has not yet been carried out in all detail. Also this technique shows some promise when applied to the Einstein–Maxwell perturbation system in which both spins 1 and 2 are simultaneously perturbed, although any actual solution is as yet unavailable.

We believe that the approach to tensor quantities via the higher-order representations of the Lorentz group (pioneered by Debever[13]), can be used effectively to determine the general structure of other problems as well. An example of this is given by our demonstration that the only requirement for the decoupling of *one* of the wave equations (such as for δC_{++}) is that the space–time be algebraically special. This fact should motivate study of perturbations to background spaces other than type D. A particular case of clear interest would be those of type N, corresponding to perturbations over sourceless gravitational waves. We also point out the work of Cohen and Kegeles,[50] who have been looking at Deybe potentials for massless (sourceless) fields. They find that the condition that a space–time be vacuum algebraically special permits for a description of such perturbing massless fields (for spin $s = 0, \frac{1}{2}, 1$, and 2) in terms of a single Deybe potential, which must satisfy a wavelike equation. Since this potential is a single (complex) scalar quantity the equation is "already decoupled" and one may proceed to determine the components of the perturbing field itself using only differential operations.

Additionally, another relevant problem concerns the solutions of equations (5.18), which determine $R(q)$ and $S(p)$, presumably by numerical techniques. There are two essential prerequisites for such a program. The first is a better understanding of the general PD solutions, especially those which have nonvanishing acceleration parameter. It appears this would be a logical additional step in applications to binary stellar systems with condensed objects. A second need is to properly organize the dependence of the solutions on the many parameters so that the essential physics of the problem can be clearly understood.

APPENDIX

For convenience of the reader we display explicitly the components of the $D(0,1)$ [and $D(1,0)$] projection operators defined in (2.8):

$$\mathscr{L}^{a}{}_{\alpha\beta}=\frac{i}{2}\begin{pmatrix} 0 & \delta^a_0 & -\sqrt{2}\,\delta^a_- & 0 \\ -\delta^a_0 & 0 & 0 & -\sqrt{2}\delta^a_+ \\ \sqrt{2}\,\delta^a_- & 0 & 0 & \delta^a_0 \\ 0 & \sqrt{2}\,\delta^a_+ & -\delta^a_0 & 0 \end{pmatrix}, \tag{A1}$$

$$\mathscr{L}^{\dot a}{}_{\alpha\beta}=\frac{i}{2}\begin{pmatrix} 0 & \delta^{\dot a}_0 & 0 & \sqrt{2}\delta^{\dot a}_+ \\ -\delta^{\dot a}_0 & 0 & \sqrt{2}\,\delta^{\dot a}_- & 0 \\ 0 & -\sqrt{2}\delta^{\dot a}_- & 0 & -\delta^{\dot a}_0 \\ -\sqrt{2}\,\delta^{\dot a}_+ & 0 & \delta^{\dot a}_0 & 0 \end{pmatrix}. \tag{A2}$$

In addition to the properties listed in (2.9), other useful relations satisfied by the $\mathscr{L}^{a}{}_{\alpha\beta}$ are

$$4\,\mathscr{L}_{a}{}^{\alpha\beta}\mathscr{L}^{a}{}_{\gamma\delta}=\delta^{\alpha\beta}_{\gamma\delta}+i\,g^{\alpha\epsilon}g^{\beta\zeta}\eta_{\epsilon\zeta\gamma\delta}, \tag{A3}$$

$$\eta_{abc}\mathscr{L}^{b\alpha\beta}\mathscr{L}^{c}{}_{\gamma\delta}=2\mathscr{L}_{a}{}^{[\alpha}{}_{[\gamma}\delta_{\delta]}{}^{\beta]}. \tag{A4}$$

We also note a few of the properties of the $Z(2;1,1)$ projections $W^{a\dot b}{}_{\alpha\beta}\equiv-2\mathscr{L}^{a}{}_{\alpha}{}^{\gamma}\mathscr{L}^{\dot b}{}_{\gamma\beta}$:

$$W^{a\dot b}{}_{\alpha\beta}W_{c\dot d}{}^{\alpha\beta}=\delta^{a}_{c}\delta^{\dot b}_{\dot d}, \tag{A5}$$

$$W^{a\dot b}{}_{\alpha\beta}W_{a\dot b}{}^{\gamma\delta}=\delta^{\gamma}{}_{(\alpha}\delta_{\beta)}{}^{\delta}-\tfrac{1}{4}g_{\alpha\beta}g^{\gamma\delta}, \tag{A6}$$

$$-W^{a\dot b}{}_{\alpha\beta}=\begin{pmatrix} \delta^a_-\delta^{\dot b}_+ & \tfrac12\delta^a_0\delta^{\dot b}_0 & -\frac{1}{\sqrt{2}}\delta^a_-\delta^{\dot b}_0 & \frac{1}{\sqrt{2}}\delta^a_0\delta^{\dot b}_+ \\ \text{sym} & \delta^a_+\delta^{\dot b}_- & -\frac{1}{\sqrt{2}}\delta^a_0\delta^{\dot b}_- & \frac{1}{\sqrt{2}}\delta^a_+\delta^{\dot b}_0 \\ \text{sym} & \text{sym} & \delta^a_-\delta^{\dot b}_- & -\tfrac12\delta^a_0\delta^{\dot b}_0 \\ \text{sym} & \text{sym} & \text{sym} & \delta^a_+\delta^{\dot b}_+ \end{pmatrix}, \tag{A7}$$

where the notation sym designates the symmetric entry.

The generators of SL$(2,C)$—a basis for its Lie algebra—are very important to the details of all the discussion in this article. However, for the purpose considered we only need the generators restricted to the representations $D(0,j)$ (and their complex conjugates), which are just the usual angular momentum matrices $\mathscr{J}^{(j)}=-iJ^{(j)}$, which, for every (j), satisfy the commutation relations

$$[\mathscr{J}^{a},\mathscr{J}^{b}]=\eta^{abc}\mathscr{J}_{c}. \tag{A8}$$

We merely note the well known fact that for $j=1$, the 3×3 matrices $(\mathscr{J}^{a})^{bc}$ can be represented by

$$(\mathscr{J}^{(1)a})^{bc}=-\eta^{abc} \tag{A9}$$

as can be verified by explicit insertion into (A8). The generators for $[D(0,1)]^2$ would be $I\otimes\mathscr{J}^{(1)a}+\mathscr{J}^{(1)a}\otimes I$. Therefore, using the renumbering operator $Z_{A}{}^{ab}$, one obtains

$$\mathscr{J}^{(2)a}{}_{AB}=Z_{Acd}Z_{B}{}^{ef}(\delta^{c}_{e}\mathscr{J}^{(1)ad}{}_{f}+\delta^{d}_{f}\mathscr{J}^{(1)ac}{}_{e})=-2\eta^{abc}Z_{Ab}{}^{d}Z_{Bcd}, \tag{A10}$$

where it is probably useful to actually write out the matrix

$$
i\,\mathscr{J}^{(2)a}{}_{AB} = \begin{pmatrix} 0 & 0 & 0 & -\sqrt{2}\,\delta^a_+ & -2\delta^a_0 \\ 0 & 0 & \sqrt{3}\,\delta^a_+ & \delta^a_0 & -\sqrt{2}\,\delta^a_- \\ 0 & -\sqrt{3}\,\delta^a_+ & 0 & \sqrt{3}\,\delta^a_- & 0 \\ \sqrt{2}\,\delta^a_+ & -\delta^a_0 & -\sqrt{3}\,\delta^a_-, & 0 & 0 \\ 2\delta^a_0 & \sqrt{2}\,\delta^a_- & 0 & 0 & 0 \end{pmatrix}. \tag{A11}
$$

The useful relation for the product of two such $\mathscr{J}^{(2)}$'s is given by

$$
\mathscr{J}^{(2)aAB}\mathscr{J}^{(2)b}{}_A{}^C = \tfrac{4}{3}g^{ab}g^{BC} - 2h^{BCD}Z_D{}^{ab} - \tfrac{1}{2}\eta^{abc}\mathscr{J}^{(2)}{}_c{}^{BC} - 2\,\eta^{acd}\eta^{bef}Z^B{}_{ce}Z^C{}_{df}. \tag{A12}
$$

The proof of this relation is performed by starting with (A10) and repeatedly using the product relations for $Z_A{}^{ab}$,

$$
Z^{Aab}Z_{Acd} = \delta^a{}_{(c}\delta_{d)}{}^b - \frac{1}{3}g^{ab}g_{cd}/3, \tag{A13}
$$

$$
Z^A{}_{ac}Z^B{}_b{}^c = \frac{1}{3}g_{ab}g^{AB} + h^{ABC}Z_{Cab} - \tfrac{1}{4}\eta_{abd}\mathscr{J}^{(2)dAB}. \tag{A14}
$$

A comparison of the connection 1-forms with the rotation coefficients of Newman and Penrose has already been given in (2.20). Here we give as well a comparison of the components of the self-dual part of the conformal tensor. The comparison is given in the form of a table in which we list, in order, our notation, that of Newman and Penrose,[10] that of Plebański,[23] and that of Bardeen and Press,[3] and, in the last columns, the tensorial components:

$$
\begin{aligned}
C_{++} &= 2\Psi_0 = -C^{(5)} = 2\Psi_{+2} = -2C_{2424},\\
C_{+} &= 4\Psi_1 = -2C^{(4)} = 4\Psi_{+1} = 4C_{1224} = 4C_{2434},\\
C_0 &= 2\sqrt{6}\Psi_2 = -\sqrt{6}C^{(3)} = 2\sqrt{6}\Psi_0 = -2\sqrt{6}C_{1324},\\
C_{-} &= 4\Psi_3 = -2C^{(2)} = 4\Psi_{-1} = 4C_{1312} = 4C_{1334},\\
C_{--} &= 2\Psi_4 = -C^{(1)} = 2\Psi_{-2} = -2C_{1313}.
\end{aligned} \tag{A15}
$$

A comparison with the spinor components is given in (4.7). It is also worthwhile to recommend the work of Ernst[51] in which a more detailed comparison of many useful notations is made.

[1] A. Einstein, Sitzber. Preuss. Akad. Wiss., 688 (1916); 154 (1918).

[2] C. Lanczos, Z. Phys. **31**, 112 (1925); E. Lifshitz, J. Phys. USSR **10**, 116 (1946); T. Regge and J.A. Wheeler, Phys. Rev. **108**, 1063 (1957).

[3] P.C. Peters, Phys. Rev. **146**, 938 (1966); F.J. Zerilli, Phys. Rev. D **2**, 2141 (1970); W.H. Press, Astrophys. J. **170**, L105 (1971); D.M. Chitre and R.H. Price, Phys. Rev. Lett. **29**, 185 (1972); R.H. Price, Phys. Rev. D **5**, 2419 (1972); J.M. Bardeen and W.H. Press, J. Math. Phys. **14**, 7 (1973); S. Chrandrasekhar, Proc. R. Soc. London, Ser. A **343**, 280 (1975).

[4] S.A. Teukolsky, Phys. Rev. Lett. **29**, 1114 (1972); Astrophys. J. **185**, 635 (1973).

[5] R. Kerr, Phys. Rev. Lett. **11**, 237 (1963).

[6] Examples include J.B. Hartle, Phys. Rev. D **9**, 2749 (1974) and S. Chandrasekhar, Proc. R. Soc. London, Ser. A **352**, 325 (1977).

[7] S. Chandrasekhar, Proc. R. Soc. London, Ser. A **345**, 145 (1975); see also Refs. 44 and 48.

[8] B. Carter, Commun. Math. Phys. **10**, 280 (1968). The connection between separability and second-rank Killing tensors may be found in M. Walker and R. Penrose, Commun. Math. Phys. **18**, 265 (1970).

[9] J.M. Stewart and M. Walker, Proc. R. Soc. London, Ser. A **341**, 49 (1974).

[10] E.T. Newman and R. Penrose, J. Math. Phys. **3**, 566 (1962).

[11] R. Geroch, A. Held, and R. Penrose, J. Math. Phys. **14**, 874 (1973).

[12] See the pioneering article by R. Sachs, Proc. R. Soc. London, Ser. A **264**, 309 (1961).

[13] M. Cahen, R. Debever, and L. Defrise, J. Math. Mech. **16**, 761 (1967). The result was announced earlier by R. Debever, Cah. Phys. **18**, 1 (1964). See also M. Carmeli, Nuovo Cimento A **7**, 9 (1972).

[14] R.G. McLenaghan and N. Tariq, J. Math. Phys. **16**, 2306 (1975).

[15] R. Debever, Bull. Cl. Sci. Acad. R. Belg. **62**, 662 (1976); L.P. Hughston, R. Penrose, P. Sommers, and M. Walker, Commun. Math. Phys. **27**, 303 (1972).

[16] J.F. Plebański and M. Demiański, Ann. Phys. (N.Y.) **98**, 98 (1976).

[17] G.J. Weir, "Type D Spaces and Quasidiagonalizability," Ph.D. thesis, University of Canterbury, Christchurch, New Zealand, 1976 (unpublished). One may also check case by case through the table of W. Kinnersley, J. Math. Phys. **10**, 1195 (1969).

[18] T. Levi-Civita, Atti dei Acc. Lincei Rendiconti **27**, 343 (1918).

[19] W. Kinnersley and M. Walker, Phys. Rev. D **2**, 1359 (1970).

[20] D.M. Chitre, R.H. Price, and V.D. Sandberg, Phys. Rev. D **11**, 747 (1975).

[21] For a discussion of some of the difficulties associated with this problem see, e.g., D.M. Chitre, Phys. Rev. D **13**, 2713 (1976). Also see V. Moncrief, Phys. Rev. D **10**, 1057 (1974); F.J. Zerilli, Phys. Rev. D **9**, 860 (1973).

[22] A.L. Dudley and J.D. Finley, III, Phys. Rev. Lett. **38**, 1505 (1977). [Note errata, Phys. Rev. Lett. **39**, 367 (1977).]

[23] We use a metric of signature +2 and generally follow the conventions of J.F. Plebański, "Spinors, Tetrads and Forms," Centro de Investigaciones y de Estudios Avanzados del Instituto Politécnico Nacional, Ap. Postal 14-740, México 14, D.F. México (unpublished). See as well J.F. Plebański, J. Math. Phys. **16**, 2395 (1975); G.C. Debney, J. Math. Phys. **12**, 1088 (1971). Greek indices run from 1 to 4, while lower case Latin indices taken on the values +, 0, − and upper case Latin indices take on the values ++, +, 0, −, −−, except in Sec. 4 where they are used for spinors and take the values 1, 2. A comma is used to denote partial differentiation, either in a coordinate direction or a tetradial direction, while a semicolon denotes the usual covariant derivative (acting on Greek indices only). Square brackets denote antisymmetrization over those indices of the same kind which are contained between them while round brackets denote a similar symmetri-

zation. The symbol $\otimes$ is the usual tensor product, while $\underset{s}{\otimes}$ is the symmetric part: $\alpha \underset{s}{\otimes} \beta = \frac{1}{2}(\alpha \otimes \beta + \beta \otimes \alpha)$.

[24]In addition to the information in Ref. 23, a good background on differential forms may be obtained from C.W. Misner, K.S. Thorne, and J.A. Wheeler, *Gravitation*, parts III and IV (Freeman, San Francisco, 1973).

[25]E.P. Wigner, Ann. Math. **40**, 149 (1939); S. Weinberg, *Lectures on Particle and Field Theory*, Brandeis Summer Institute in Theoretical Physics (Prentice-Hall, Englewood Cliffs, N.J., 1965), Vol. 2.

[26]We use a definition of duality which is arranged so that the dual of the dual of any p-form is exactly that p-form again. If

$$\omega = \frac{1}{p!}\,\omega_{\mu_1\cdots\mu_p} e^{\mu_1} \wedge \cdots \wedge e^{\mu_p}$$

is an arbitrary p-form, then

$$^*\omega = \frac{i^{pp'+s}}{p!p'!}\,\omega^{\nu_1\cdots\nu_p} \eta_{\nu_1\cdots\nu_p\mu_1\cdots\mu_{p'}} e^{\mu_1} \wedge \cdots \wedge e^{\mu_{p'}}$$

is a p'-form, where $p' \equiv n - p$, n is the dimension of the manifold in question (four in the present discussion), and the nature of the signature is given by s which is either 0 or 1 depending on whether the metric is positive definite or of the Minkowski type, respectively, while $\eta_{\alpha_1\cdots\alpha_n}$ is the tensor made from the Levi-Civita alternating symbol,

$$\eta_{\alpha_1\cdots\alpha_n} \equiv [(-1)^s \det(g_{\gamma\delta})]^{1/2} \epsilon_{\alpha_1\cdots\alpha_n}.$$

(Note that in our null tetrad basis, $\eta^{1234} = +i = \eta_{1234}$.)

[27]For A a p-form and B a q-form, $p \geqslant q$, the right interior product is defined by $B \lrcorner A = {}^*(B \wedge {}^*A)$: W. Slebodziński, *Exterior Forms and Their Applications* (Polish Scientific Publishers, Warsaw, 1970), p. 396. Also the contraction of the two p-forms is denoted by

$$B\cdot A = \left(\frac{1}{q!} B_{\mu_1\cdots\mu_q} e^{\mu_1} \wedge \cdots e^{\mu_q}\right)\cdot\left(\frac{1}{p!} A_{\nu_1\cdots\nu_p} e^{\nu_1} \wedge \cdots \wedge e^{\nu_p}\right)$$

$$= (B_{\mu_1\cdots\mu_q} e^{\mu_1} \otimes \cdots \otimes e^{\mu_q})\cdot(A_{\nu_1\cdots\nu_p} e^{\nu_1} \otimes \cdots \otimes e^{\nu_p})$$

$$= B_{\mu_1\cdots\mu_{q-1}\lambda} A^{\lambda}{}_{\nu_2\cdots\nu_p} e^{\mu_1} \otimes \cdots \otimes e^{\mu_{q-1}} \otimes e^{\nu_2} \otimes \cdots \otimes e^{\nu_p}.$$

[28]E. Cartan, *Lecons sur la Géometrie des Espaces de Riemann* (Gauthier-Villars, Paris, 1946), Secs. 187–92.

[29]See S. Kobayashi and K. Nomizu, *Foundations of Differential Geometry* (Interscience-Wiley New York, 1963), Vol I. Also see Ref. 13.

[30]P.L. Chrzanowski, Phys. Rev. D **13**, 806 (1976).

[31]M. Demiański, Gen Rel. Grav. **7**, 551 (1976).

[32]For instance, see the nice discussion, in a somewhat different notation, of A.I. Janis and E.T. Newman, J. Math. Phys. **6**, 902 (1965). It is clear from the exponential formulas given in (3.23) that this parametrization (and products thereof) covers all elements in the component of O(3,1) which contains the identity.

[33]See, for example, the discussion in Chap. 20 of Misner, Thorne, and Wheeler, cited in Ref. 24.

[34]An early indication that this could be done was the procedure given in P.L. Chrzanowski, Phys. Rev. D **11**, 2042 (1975). A particular example where the complete solution is determined in a special case is given in Ref. 30.

[35]M.P. Ryan, Phys. Rev. D **10**, 1736 (1974).

[36]Although we have not seen any related printed material, the inspiration for this approach and some consequences that it generates in Sec. 5 were generated by a talk by S. Chandrasekhar given at the Eighth International Conference on General Relativity and Gravitation, Waterloo, Canada, Aug. 1977.

[37]D.Z. Freedman and P. van Nieuwenhuizen, Phys. Rev. D **14**, 912 (1976); A. Das and D.Z. Freedman, Nucl. Phys. B **114**, 271 (1976).

[38]For a discussion of the spinor notation see F. Pirani, *Lecutres on General Relativity*, Brandeis Summer Institute in Theoretical Physics (Prentice-Hall, Englewood Cliff, N.J. 1965), Vol. 1; and also Ref. 23.

[39]P.A.M. Dirac, Proc. R. Soc. London, Ser. A **155**, 447 (1936); J.F. Plebański, Acta. Phys. Polonica **27**, 361 (1965).

[40]H.A. Buchdahl, Nuovo Cimento **10**, 96 (1958); and see the second paper of Ref. 30, Eq. (9.13).

[41]See as well the first paper of Ref. 3, Eq. (4.10) and the footnote following.

[42]The coordinate q used here is introduced in Sec. 3 of Ref. 16, and is the negative inverse of the symbol q used in earlier sections of that reference.

[43]This version of separability, where there is an additional known (multiplicative) factor, is known as R-separability. See C.P. Boyer, SIAM J. Math. Anal. **7**, 230 (1976).

[44]S.A. Teukolsky and W.H. Press, Astrophys. J. **185**, 649 (1973).

[45]J.B. Hartle and D.C. Wilkins, Commun. Math. Phys. **38**, 47 (1974).

[46]This required σ-gauge transformation is then the explanation of the rather strange factors ρ^{2s} which occur in Teukolsky's table of separable forms of $\delta\psi_{-s}$, over those of $\delta\psi_{+s}$. (The factor $q + ip$ becomes Teukolsky's factor ρ in the limit to the Kerr geometry.)

[47]S.A. Teukolsky and W.H. Press, Astrophys. J. **193**, 443 (1974).

[48]R.M. Wald, J. Math. Phys. **14**, 1453 (1973).

[49]S. Chandrasekhar, Proc. R. Soc. London, Ser. A **358**, 421, 441 (1978).

[50]J.M. Cohen and L.S. Kegeles, Phys. Lett. A **54**, 5 (1975) is a summary of the results without proof. J.M. Cohen and L.S. Kegeles, Phys. Rev. D **10**, 1070 (1974) gives the details of the construction for spin 1, while a current preprint of theirs gives details for arbitrary spin s.

[51]F.J. Ernst, J. Math. Phys. **19**, 489 (1978).

On the completeness of the natural modes for quantum mechanical potential scattering

B. J. Hoenders

Technical Physical Laboratories, State University at Groningen, Nijenborgh 18, 9747 AG Groningen, The Netherlands
(Received 22 May 1978)

The set of natural modes, associated with quantum mechanical scattering from a central potential of finite-range is shown to be complete. The natural modes satisfy a non-Hermitian homogeneous integral equation, or alternatively, are solutions of the time independent Schrödinger equation subject to a recently formulated nonlocal boundary condition (the quantum mechanical extinction theorem). An expansion theorem similar to that of Hilbert–Schmidt is formulated, valid for values of the solution of the scattering integral equation inside the range of the potential. The boundary conditions generated by the quantum mechanical extinction theorem are shown to be closely connected with the Jost function.

1. INTRODUCTION

For a long time attempts have been made in the theory of quantum-mechanical potential scattering to define the so-called natural modes of the scatterer. The first ones who tried to define the natural modes were Kapur and Peierls.[1] However, as it appeared to have been first pointed out by Siegert[2] their theory suffers from several unphysical phenomena like the dependence of the resonances upon the energy of the incoming wave.

Considering central symmetrical scatterers, Siegert[2] formulated another definition for the natural modes which leads to physically much more satisfactorially results. His theory was completed by Humblet and Rosenfeld.[3] An extensive survey of the literature on this subject can be found in the review article by More and Gerjoy.[4]

The definition of the natural modes given by Humblet and Rosenfeld[3] is essentially one-dimensional, because they restrict themselve to central symmetric potentials. A general definition for natural modes for quantum-mechanical potential scattering, as well as for electromagnetic scattering, has been formulated by Pattanayak and Wolf[4,5] (see also Wolf[6]). Their definition applies to genuine three-dimensional scattering problems and reduces to the Siegert–Humblet–Rosenfeld definition if the potential is central symmetric. A review of their theory, from which the basic relations of this paper are derived, is given in Sec. 2.

The natural modes can be shown to be the solutions of the radial Schrödinger equation

$$\left[-\frac{1}{r^2}\frac{\partial}{\partial r}r^2\frac{\partial}{\partial r}+\frac{l(l+1)}{r^2}+U(r)+k^2\right]\chi_l(r;k)=0\,, \tag{1.1}$$

subject to the conditions that χ_l is regular in the interval $0\leqslant r\leqslant a$, and

$$k\left[B(k)\,\chi_l(a;k)+C(k)\frac{\partial}{\partial a}\chi_l(a,k)\right]=0, \tag{1.2}$$

where

$$B(k)=\frac{\partial}{\partial a}h_l^{(1)}(ka)\,, \tag{1.3a}$$

$$C(k)=-h_l^{(1)}(ka)\,, \tag{1.3b}$$

$$U(r)=\frac{2m}{\hbar^2}V(r), \tag{1.3c}$$

and a denotes the range of the central symmetrical potential $V(r)$. [Siegert,[2] Humblet and Rosenfeld,[3] this paper, Sec. 2. The condition (1.2) is usually obtained from the requirement that both the field and its normal derivative are continuous across the surface of the sphere with radius a]. The purpose of this paper is to construct a Sturm–Liouville type of theory for the set of functions (natural modes) satisfying Eqs. (1.1) and (1.2) and especially to show the completeness of the natural modes inside and not *on*! the sphere with radius a. Once we have shown the completeness of the natural modes, we can solve the following initial value problem: Calculate the field inside the sphere with radius a if at $t=0$ the part of a wavepacket inside the sphere is known. To be more specific, this field can be approximated arbitrarily closely by a series

$$\sum_{l=0}^{\infty}\sum_{m=-l}^{+l}\sum_{n}^{N}a_n(N,l,m)\;\chi_l(r,k_{ln})\,Y_l^m(\theta,\phi)\times\exp\left(i\frac{\hbar}{2m}k_{ln}^2t\right), \tag{1.4}$$

where the numbers k_{ln} are the roots of (1.2). The series $\Sigma_n^N a_n(N,l,m)\;\chi_l(r,k_{ln})$ approximates the l,mth Fourier coefficient of the initial field with respect to the set of functions $Y_l^m(\theta,\phi)$ arbitrarily closely for sufficient large N. The series only determines the field for values of $r<a$. If the boundary $r=a$ is to be included, the set of functions $\chi_l(r,k_{ln})$ is no longer complete in the interval $0\leqslant r\leqslant a$. This point will be discussed in a future paper with Dr. D.N. Pattanayak and is connected with a background scattering term.

It is unfortunately not possible to use ordinary Sturm–Liouville theory to prove the completeness of the natural modes defined by Eqs. (1.1) and (1.2) because the eigenvalue k *explicitly* shows up in the boundary condition. However, the completeness of the natural modes can be shown on using the calculus of residues. It seems that Cauchy[7] was the first one who used this method, which essentially leads to an interpolation formula, (Eq. (3.20), to prove the completeness

 0022-2488/79/020329-07$01.00

of sets of functions. For similar methods and a survey of the literature we refer to Hoenders.[8]

The explicit occurrence of the eigenvalues in the boundary condition spoils the hermiticity of the problem and leads usually to nonreal eigenvalues and nonorthogonal eigenfunctions (see Morse and Feshbach[9] and Nussenzveig[10]).

It has been extensively shown in a previous publication, Hoenders,[8] that this type of problem, connected with continuity conditions on a surface, rather than boundary conditions, arises in many branches of physics. As an example, we mention the solution of an initial value problem connected with a sphere, characterized by a scalar constant complex index of refraction n_1, embedded in an infinite medium characterized by a scalar constant index of refraction n_2, in terms of the natural modes of the sphere.

The frequencies of the natural modes are determined by the continuity requirement on the tangential components of the electromagnetic field vectors which leads to an infinite set of equations similar to Eq. (1.2). Another example of a non-Hermitian problem is constructed by Morse and Feshbach.[9] They considered a string of length l which is under tension T and supported by a rigid support at $x=0$ and a nonrigid support $x=l$. This latter support has enough longitudinal strength to support the tension T, but it yields a little to transverse force imparted to it by the string. Suppose this yielding involves both friction and stiffness of the support for sidewise motion, so that the relation between the transverse force transmitted by the string, which is $-T(\partial y/\partial x)_l$, is equal to R_s times the transverse velocity of the support, $(\partial y/\partial t)_l$, plus K_s times the displacement of the support $y(l)$:

$$-T\frac{\partial y}{\partial x}=R_s\frac{\partial y}{\partial t}+K_s y, \quad \text{at } x=l, \qquad y=0, \quad \text{at } x=0. \tag{1.5}$$

If we assume that $y(x,t)=v(x)\exp(-i\omega t)$ and that $y(x,t)$ is a solution of the wave equation

$$\left(\frac{\partial^2}{\partial x^2}-\frac{1}{c^2}\frac{\partial^2}{\partial t^2}\right)y(x,t)=0 \tag{1.6}$$

we derive from Eqs. (1.5) and (1.6) that the functions $v_n(x)$ are the solutions of the second order linear differential equation $v_n''(x)+k_n^2 v_n(x)=0$, subject to the "boundary" conditions

$$-T\frac{\partial v}{\partial x}=-ikc\,R_s v+K_s v, \quad \text{if } x=l, \tag{1.7a}$$

$$v(0)=0, \quad \text{and} \quad \omega=ck. \tag{1.7b}$$

We cannot use the results of ordinary Sturm-Liouville theory to prove the completeness of the set of functions $\{v_n(x)\}$ because the "boundary" condition (1.7a) depends explicitly on the eigenvalue. The terminology "boundary condition" is even misleading because condition (1.7a) is not generated by a true boundary condition but arises from the condition that the force at the point 0^- is equal to the force at the point 0^+. The eigenvalues even have a nonvanishing imaginary part which accounts for the damping of the natural modes (Morse and Feshbach[9]).

It has been pointed out previously that the conditions (1.2) have been derived from a general definition for the natural modes of quantum mechanical as well as electromagnetic scattering by Wolf and Pattanayak.[5,6,11] At short survey of their theory will be given in the following Section 2, whereas the condition (1.2) will be derived in Section 3.

2. DERIVATION OF THE BASIC EQUATIONS

From the time independent Schrödinger equation

$$(\nabla^2+k^2+U(r))\psi(\mathbf{r};k)=0, \tag{2.1}$$

where

$$k^2=\frac{2mE}{\hbar^2}, \tag{2.2}$$

$$U=\frac{2m}{\hbar^2}V, \tag{2.3}$$

and with the use of Green's theorem, the following three identities can be derived:

$$\psi(\mathbf{r};k)=\int_\tau G(|\mathbf{r}-\mathbf{r}'|;k)U(\mathbf{r}')\psi(\mathbf{r}';k)d\mathbf{r}+\Sigma(\mathbf{r}), \tag{2.4}$$

$$0=\frac{1}{4\pi}[\Sigma(\mathbf{r}_<)-\Sigma^\infty(\mathbf{r}_<)], \tag{2.5}$$

$$\psi(\mathbf{r}_>)=\frac{1}{4\pi}[\Sigma^\infty(\mathbf{r}_>)-\Sigma(\mathbf{r}_>)]. \tag{2.6}$$

Here

$$\Sigma(\mathbf{r})=\iint_\sigma\Big[\psi(\mathbf{r}';k)\frac{\partial}{\partial n}G(|\mathbf{r}-\mathbf{r}'|;k) -G(|\mathbf{r}-\mathbf{r}'|;k)\frac{\partial}{\partial n}\psi(\mathbf{r}';k)\Big]d\sigma, \tag{2.7}$$

if

$$G(|\mathbf{r}-\mathbf{r}'|;k)=\frac{\exp(ik|\mathbf{r}-\mathbf{r}'|)}{|\mathbf{r}-\mathbf{r}'|} \tag{2.8}$$

and τ denotes a finite domain delimited by a surface σ and σ_∞ a sphere ∞ with infinite radius. All points lying inside the sphere are denoted by $\mathbf{r}_<$ and all points lying outside the sphere are denoted by $\mathbf{r}_>$.

The total wavefunction $\psi(\mathbf{r};k)$ is a superposition of the incoming wave $\psi^{(i)}(\mathbf{r};k)$ and the scattered wave $\psi^{(s)}(\mathbf{r};k)$. The latter is required to satisfy Sommerfeld's radiation condition at infinity, and therefore

$$\iint_\sigma\Big[\psi^{(s)}(\mathbf{r}';k)\frac{\partial}{\partial n}G(|\mathbf{r}-\mathbf{r}'|;k) -G(|\mathbf{r}-\mathbf{r}'|;k)\frac{\partial}{\partial n}\psi^{(s)}(\mathbf{r}';k)\Big]d\sigma=0. \tag{2.9}$$

Because the incoming wave satisfies Helmholtz's equation, Green's theorem yields

$$\iint_\sigma \left[\psi^{(i)}(\mathbf{r}';k)\frac{\partial}{\partial n}G(|\mathbf{r}-\mathbf{r}'|;k)\right.$$

$$\left. -G(|\mathbf{r}-\mathbf{r}'|;k)\frac{\partial}{\partial n}\psi^{(i)}(\mathbf{r};k)\right]d\sigma=\psi^{(i)}(\mathbf{r};k). \tag{2.10}$$

Combining of (2.5), (2.9), and (2.10) gives the important relation:

$$\Sigma(\mathbf{r}_< = \psi^{(i)}(\mathbf{r}_<;k), \tag{2.11}$$

which has to be satisfied for all values of $\mathbf{r}_<$ inside the sphere. Equation (2.11) is the quantum mechanical analog of the electromagnetic extinction theorem: The incoming wave is extinguished by the values of ψ and $\partial\psi/\partial n$ at the boundary. Moreover, combination of (2.4) and (2.11) leads to

$$\psi(\mathbf{r}_<)=\int_\tau G(|\mathbf{r}_<-\mathbf{r}'|,k)U(\mathbf{r}')\psi(\mathbf{r}')\,d\mathbf{r}'+\psi^{(i)}(\mathbf{r}_<), \tag{2.12}$$

which is the usual integral formulation for potential scattering for values of $\mathbf{r}$ situated inside σ. Equation (2.12) can also be shown to be valid for values of $\mathbf{r}$ situated outside σ on using the techniques of this section: Let $\mathbf{r}$ be an arbitrarily chosen point, situated outside σ and suppose that σ encloses both $\mathbf{r}$ and the scatterer. Combination of Eqs. (2.4)–(2.10) then shows the validity of (2.12) for values of $\mathbf{r}$ situated outside σ.

The natural modes for quantum mechanical scattering are defined by Wolf and Pattanayak,[4,5] as those solutions of the time independent Schrödinger equation.

$$(\nabla^2+k^2+U)\psi(\mathbf{r},k)=0, \tag{2.13}$$

satisfying the nonlocal boundary condition

$$\iint_\sigma \left[\psi(\mathbf{r}';k)\frac{\partial}{\partial n}G(|\mathbf{r}_<-\mathbf{r}'|;k)\right.$$

$$\left. -G(|\mathbf{r}_<-\mathbf{r}'|;k)\frac{\partial}{\partial n}\psi(\mathbf{r}';k)\right]d\sigma=0, \tag{2.14}$$

to be valid for all values of $\mathbf{r}_<$ lying inside σ. Hence, alternatively, Eqs. (2.4) and (2.7) show that these modes are the solutions of the homogeneus part of Eq. (2.12):

$$\psi_n(\mathbf{r}_<;k_n)=\int_\tau G(|\mathbf{r}_<-\mathbf{r}'|;k_n)\,U(\mathbf{r}')\psi(\mathbf{r}';k_n)d\mathbf{r}'. \tag{2.15}$$

It is to be stressed that the ordinary Hilbert–Schmidt theory for linear integral equations with symmetrical polar kernels cannot be used because the integral of (2.15) depends *nonlinearly* on k. The completeness of the modes (2.15) will be shown in the next section.

3. CALCULATIONAL PROCEDURE

Theorem 1: Consider the time independent Schrödinger equation

$$[\nabla^2+k^2+U(\mathbf{r})]\psi(\mathbf{r},k)=0 \tag{3.1}$$

in a spherical region of radius a bounded by a surface σ, and assume that $U(\mathbf{r})=U(r)$ is of bounded variation. Suppose that (3.1) is to be solved subject to the nonlocal boundary condition

$$\iint_\sigma \left[\psi(\mathbf{r}';k)\frac{\partial}{\partial n}G(|\mathbf{r}-\mathbf{r}'|;k)\right.$$

$$\left. -G(|\mathbf{r}-\mathbf{r}'|;k)\frac{\partial}{\partial n}\psi(\mathbf{r}';k)\right]d\sigma=0, \tag{3.2}$$

which has to be valid for *all* values of $\mathbf{r}$ lying inside the spherical region with radius a, with

$$G(|\mathbf{r}-\mathbf{r}'|;k)=\exp(ik|\mathbf{r}-\mathbf{r}'|)/|\mathbf{r}-\mathbf{r}'|. \tag{3.3}$$

Then

(1) There exists an infinite set of eigenvalues k_n and a set of eigenfunctions (natural modes) $\psi(r,\theta,\phi;k_n)$

(2) The set of natural modes is complete within the sphere of radius a.

Proof: Following the analysis given by Pattanayak and Wolf,[4] we expand the wavefunctions $\psi(\mathbf{r}_<)$ into a series of partial waves (cf. Ref. 10)

$$\psi(\mathbf{r}_<,k)=\sum_{l=0}^{\infty}\chi_l(\mathbf{r}_<,k)P_l(\cos\theta), \tag{3.4}$$

where θ is the angle between the momentum of the incoming plane wave and the direction of the vector $\mathbf{r}_<$ and the functions χ_l are the regular solutions of the radial Schrödinger equation for the lth partial wave. The expansion (3.4) and the expansion

$$G(\mathbf{r},\mathbf{r}')=k\sum_{l=0}^{\infty}(2l+1)\,j_l(kr_<)\,h_l^{(1)}(kr_>)\,P_l(\cos\theta) \tag{3.5}$$

for the Green's function (3.5), valid with $r_<=\min(|\mathbf{r}|,|\mathbf{r}'|)$ and $r_>=\max(|\mathbf{r}|,|\mathbf{r}'|)$,

where $h_l^{(1)}$ is the spherical Hankel function of the first kind and order l and Φ the angle between the directions $\mathbf{r}_<$ and $\mathbf{r}'$, are then substituted in the boundary condition (2.2), which leads to

$$\sum_{l=0}^{\infty}\alpha_l(k)j_l(kr_<)P_l(\cos\Phi)=0, \tag{3.6}$$

where

$$\alpha_l(k)=ka^2[\chi_l(a,k)h'^{(1)}_l(ka)-\chi'_l(a,k)h_l^{(1)}(ka)] \tag{3.7}$$

the prime denoting differentiation with respect to a. Because of the linear independence of the Legendre polynomials $P_l(\cos\Phi)$ in the interval $0\leqslant\Phi\leqslant\pi$ it follows that we must have $\alpha_l=0$ for all l. Equations (3.7) are a set of *local* boundary conditions imposed on the radial wavefunctions $\chi_l(r_<,k)$. From now on we will write r instead of $r_<$. For bound states and resonances states the $\alpha_l(k)$ vanish. [Pattanayak[5,6]; thus this is also true for the Jost function $L_l(k)$]. We therefore expect that both functions are closely related to each other and we will show that in accordance with this expectation both functions are proportional to each other.

The Jost function $L_l(k)$ is defined by (Newton,[12] Eq. 12.142)

$$L_l(k)=[(2l+1)!!]^{-1}k^l$$

$$\exp(-\tfrac{1}{2}i\pi l)W\{f_{l+}(k,r),\phi_l(r,k)\}, \tag{3.8}$$

where W denotes the Wronskian,

$$\phi_l(r,k)=r\chi_l(r,k), \tag{3.9}$$

f_{l+} is the solution of the Volterra equation

$$f_{l+}(k,r)=i\exp(i\pi(l-\tfrac{1}{2}))(kr)h_l^{(1)}(kr)$$

$$+\int_r^\infty G_l(r,r',k)U(r')f_{l+}(k,r')dr' \tag{3.10}$$

and

$$G_l(r,r',k)=(\cos\pi l)^{-1}\tfrac{1}{2}\pi(rr')^{1/2}$$
$$\times\{J_{l+1/2}(kr)J_{l-1/2}(kr')$$
$$-J_{l+1/2}(kr')J_{l+1/2}(kr)\}. \tag{3.11}$$

Combination of Eqs. (3.7), (3.8), and (3.9) shows that, taking $r=a$,

$$L_l(k)=(i^l)[(2l+1)!!]^{-1}k^l\alpha_l(k). \tag{3.12}$$

We will need the asymptotic expansion of $L_l(k)$ for large values of $|k|$. This asymptotic expansion is obtained on inserting the zeroth order approximation $(-1)^l(kr)h_l^{(1)}(kr)$ of $f_{l+}(k,r)$ into the integral representation, Newton,[12] §12.1,

$$L_l(k)=1+(-i)^lk^{-1}\int_0^\infty U(\tau)(k\tau)j_l(k\tau)f_{l+}(k,\tau)d\tau, \tag{3.13}$$

and replacing the spherical Bessel and Hankel functions by their asymptotic expansions. Integration by parts leads to

$$L_l(k)\sim 1+\frac{(-1)^l\exp(2ika)\,U^{(m)}(a-)}{(2ik)^{m+2}},\quad \text{if}$$
$$|k|\to\infty,\quad \pi\leqslant\arg k\leqslant 2\pi, \tag{3.14}$$

where $U^{(m)}(a-)$ denotes the first nonvanishing derivative of $U(r)$ at $r=a$, and $U^{(0)}(a-)\equiv U(a-)$, whereas the Riemann–Lebesgue theorem leads to

$$L_l(k)=1+O(k^{-1}),\quad \text{if}\ |k|\to\infty,\quad \pi\leqslant\arg k\leqslant 0. \tag{3.15}$$

Let the numbers λ_j be an infinite bounded set of arbitrarily chosen complex numbers. It can be shown (Hoenders,[8] Lewin[13]) that every function which is analytic inside a bounded simply connected domain D can be approximated arbitrarily closely and uniformly for all values of $k\in D$ by a suitable linear combination of a sufficiently large number of functions $\cos(\lambda_j\sqrt{k})$; i.e.,

$$(k-\delta)^{m+2}=\sum_j^n a_j(n)\cos(\lambda_jk^{1/2})+o(1), \tag{3.16}$$

if δ denotes an arbitrary complex number. Consider the contour integral

$$I_l(r,b,n)=\frac{1}{2\pi i}\int_{|k|=c_n}H(k,b)dr, \tag{3.17}$$

where

$$H(k,b)=\frac{\phi_l(r,k)\exp(ika)C(k)}{(k-\delta)^{m+2}L_l(k)(k-b)} \tag{3.18}$$

and

$$C(k)=\sum_j^n a_j(n)\cos(\lambda_j\sqrt{k}). \tag{3.19}$$

The numbers c_n are chosen in such a way that the contour passes between two successive zeros of the denominator of (3.18), and b denotes an arbitrary fixed complex number not equal to any of the zeros of $L_l(k)$. From Newton,[12] Eq. 12.137

$$\phi_l(r,k)=(2l+1)!!k^{-l-1}\sin(kr-\tfrac{1}{2}\pi l)$$
$$+o(|k|^{-l-1}\exp|\mathrm{Im}k|r),$$
$$0\leqslant\arg k\leqslant 2\pi, \tag{3.20}$$

and Eqs. (3.14), (3.15), and (3.18) we derive

$$|H(k,b)|=O\{c_n^{-1}\exp[-\tfrac{1}{2}c_n|\sin(\arg k)|(r-a)]\},$$
$$\text{if}\ c_n\to\infty. \tag{3.21}$$

Equations (3.17) and (3.21) lead to

$$\lim_{n\to\infty}I_l(r,b,n)=0,\quad \text{if}\ r<a. \tag{3.22}$$

Suppose that ρ denotes a positive number such that the domain bounded by the circle $|k-\delta|=\rho\in D$ and does not contain a zero of $L_l(k)$. These requirements can always be fulfilled by a suitable choice of the numbers ρ and δ. Calculating the integral (3.17) with the theorem of residues if $n\to\infty$ and Eq. (3.22) leads to

$$\frac{\chi_l(r,b)\exp(iba)C(b)}{(b-\delta)^{m+2}L_l(b)}$$
$$=\sum_n\frac{\chi_l(r,k_{ln})\exp(ik_{ln}a)C(k_{ln})}{L'(k_{ln})(k_{ln}-b)(k_{ln}-\delta)^{m+2}}$$
$$+\int_{|k-\delta|=\rho}H(k,b)dk, \tag{3.23}$$

if $b\neq k_{ln}$, $|b-\delta|>\rho$, and the summation has to be extended over all the zeros of $L_l(k)$. Recalling that the domain bounded by the circle $|k-\delta|=\rho$ does not contain a zero of $L_l(k)$, we derive from (3.16)

$$\left|\int_{|k-\delta|=\rho}H(k,b)dk\right|=o(1). \tag{3.24}$$

Combination of (3.23) and (3.24) yields

$$\chi_l(r,b)=L_l(b)\exp(-iba)$$

$$\times\sum_n \frac{\chi_l(r,k_{ln})\exp(ik_{ln}a)C(k_{ln})}{L'(k_{ln})(k_{ln}-b)(k_{ln}-\delta)^{m+2}}$$

$$+o(1), \quad \text{if } b\in D \text{ and } b\neq k_{ln}. \tag{3.25}$$

While calculating the residues of the integral (3.17), we assumed that the zeros of the functions $L_l(k)$ are simple. This assumption is commonly made (Rosenfeld and Humblet,[3] Nussenzveig[10]) and is certainly true for large values of $|k|$. [Newton,[12] Eq. 12.108, gives an estimate for $L'_l(k_{ln})$ if $|k_{ln}|$ is large.] We will not analyze this difficult question but will conform with the other authors, mentioned above.

The function $\chi_l(r,k)$ is the regular solution of the radial equation

$$\left[-\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right)+\frac{l(l+1)}{r^2}+U(r)+k^2\right]\chi_l(r,k)=0. \tag{3.26}$$

Ordinary Sturm–Liouville theory shows the existence of an infinite denumerable set of eigenvalues $k_{lp}^{(1)}$ and an infinite set of eigenfunctions $\chi_l(r,k_{lp}^{(1)})$, which are regular at the origin, complete on the interval $o\leqslant r\leqslant a$, and zero if $r=a$. For every eigenvalue $k_{lp}^{(1)}\neq k_{ln}$, $n=1,2,\cdots$, we choose a simple connected domain $D\in k_{lp}^{(1)}$. Choose b to be equal to $k_{lp}^{(1)}$ in Eq. (3.20). This equation then leads to the following conclusion: Every eigenfunction $\chi_l(r,k_{lp}^{(1)})$ with $k_{lp}^{(1)}\neq k_{ln}$, $n=1,2,\cdots$, can be approximated arbitrarily closely in the interval $o\leqslant r\leqslant a$ by a suitable linear combination of functions $\chi_l(r,k_{ln})$. (If $k_{lp}^{(1)}$ would coincide with one of the numbers k_{ln}, this conclusion would be trivial!)

This conclusion proves the completeness of the set of functions $\{\chi_l(r,k_{ln})\}$ because the set of functions $\{\chi_l(r,k_{lp}^{(1)})\}$ is complete in the interval $o\leqslant r\leqslant a$. The completeness of the set of natural modes $\{\chi_l(r,k_{ln})Y_l^m(\theta,\phi)\}$ is now easily established for any function $f(r,\theta,\phi)$ which is of bounded variation in the domain $0\leqslant r\leqslant a$, $0\leqslant\phi\leqslant 2\pi$, $0<\theta<\pi$ can be expanded into the set of spherical harmonics $Y_l^m(\theta,\phi)$:

$$f(r,\theta,\phi)=\sum_{l=0}^{\infty}\sum_{m=-l}^{+l} a_{lm}(r)\,Y_l^m(\theta,\phi). \tag{3.27}$$

where

$$a_{lm}(r)=\int_\Omega f(r,\theta,\phi)Y_l^m(\theta,\phi)d\Omega. \tag{3.28}$$

Because every function $a_{lm}(r)$ can be approximated arbitrarily closely by a suitable linear combination of the functions $\{\chi_l(r,k_{ln})\}$ Eq. (3.27) shows the completeness of the natural modes $\{\chi_l(r,k_{ln})Y_l^m(\theta,\phi)\}$.

4. ON THE GENERALIZATION OF THE HILBERT–SCHMIDT EXPANSION FORMULA TO THE CASE OF KERNELS DEPENDING NONLINEARLY ON THE EIGENVALUE

In the preceeding sections we established the completeness of the natural modes $\{\chi_l(r_<,k_{ln})Y_l^m(\theta,\phi)\}$, which are the solutions of the time independent Schrödinger equation subject to the nonlocal boundary condition (2.14). The problem, which arises immediately, in considering the linear integral equation (2.12) is the derivation of a Hilbert–Schmidt type of expansion (well-known in the theory of linear integral equations with kernels depending *linearly* upon the eigenvalue) for the unknown function. Naturally we expect that the eigensolutions of (2.12) are the most appropriate set of functions to formulate such an expansion.

By heuristical reasoning we will "derive" the desired expansion. This generalized Hilbert–Schmidt expansion explicitly shows the dependence of the expansion coefficients on k, which might be very useful for the calculation of the scattering cross section near resonances, and so provide a generalization of the Breit–Wigner formula. For recent developments connected with this expansion we refer to Hoenders[14] and Pattanayak.[15]

The expansion for the kind of problems we are analyzing was given without proof by Miranda[16] and derived by heuristical reasoning by Pattanayak.[17] We will first formulate Miranda's theorem, and then present a derivation that closely resembles the one due to Pattanayak (I am obliged to Dr. Pattanayak for making available to me his unpublished notes on this subject).

Theorem: Let the kernel function $G(x,y;\lambda)$ be defined in the square $-a\leqslant x\leqslant +a$, $-a\leqslant y\leqslant a$, symmetrical in the variables x and y, and analytic in λ. Let the function $\phi(x;\lambda)$ be the (supposedly) unique solution of the integral equation

$$\phi(x;\lambda)=f(x)+\lambda\int_{-a}^{+a}G(x,y;\lambda)\phi(y;\lambda)dy, \tag{4.1}$$

where the function $f(x)$ is defined and integrable on the interval $-a\leqslant x\leqslant +a$. If $\{\phi_n(x;\lambda_n)\}$ is the set of eigenfunctions of (4.1), satisfying the equation

$$\phi_n(x;\lambda_n)=\lambda_n\int_{-a}^{+a}G(x,\lambda;\lambda_n)\phi(y\lambda_n)\,dy, \tag{4.2}$$

then

$$\phi(x,\lambda)=f(x)+\lambda\sum_n\Big\{\phi_n(x)\int\int_{-a}^{+a}f(y)\phi_n(y)dy$$

$$\times\Big[(\lambda_n-\lambda)\Big(1+\lambda_n^2\int_{-a}^{+a}\frac{\partial}{\partial\lambda_n}$$

$$G(s,t)\phi_n(s)\phi_n(t)ds\,dt)\Big]^{-1}\Big)+\omega(x,\lambda), \tag{4.3}$$

if $\omega(x,\lambda)$ denotes a function defined on the interval $0\leqslant x\leqslant a$ and regular in λ for all $x\in 0\leqslant x\leqslant a$.

The formula (4.3) clearly degenerates into the well-known Hilbert–Schmidt expansion formula in case $(\partial/\partial\lambda)G(x,y,\lambda)=0$ and $\omega(x,\lambda)\equiv 0$.

Heuristic "proof": It is *assumed* that the function $\lambda^{-1}\{\phi(x;\lambda)-f(x)\}$ can be expanded into a series of partial fractions:

$$\frac{1}{\lambda}\{\phi(x;\lambda)-f(x)\}=\sum_n A_n\frac{\psi_n(x)}{\lambda-\lambda_n}+\omega(x,\lambda), \tag{4.4}$$

where $\omega(x,\lambda)$ is a regular function of λ for all values of $x \in 0 \leqslant x \leqslant a$ and $\psi_n(x)$ are functions yet to be determined. Then expanding the kernel $G(x,y,\lambda)$ into a Taylor series around the point $\lambda = \lambda_n$ up to the first order:

$$G(x,y,\lambda)$$
$$G(x,y,\lambda_n) + (\lambda - \lambda_n)\frac{\partial}{\partial \lambda_n} G(x,y,\lambda_n) + O\{(\lambda - \lambda_n)^2\}, \tag{4.5}$$

and using the identity

$$\lambda \equiv (\lambda - \lambda_n) + \lambda_n, \tag{4.6}$$

substitution of (4.4), (4.5), (4.6) into Eq. (4.1) and equating the coefficients of $(\lambda - \lambda_n)^{-1}$ and $(\lambda - \lambda_n)^0$ leads to

$$\psi_n(x) = \lambda_n \int_{-a}^{+a} G(x,y;\lambda_n)\psi_n(y)dy \tag{4.7}$$

and

$$\omega(x,\lambda_n)$$
$$=\lambda_n \int_{-a}^{+a} G(x,y\lambda_n)\,\omega(y,\lambda_n)dy + \lambda_n$$
$$\times \int_{-a}^{+a} G(x,y\lambda_n) f(y)\,dy$$
$$+A_n \int_{-a}^{+a} G(x,y\lambda_n)\,\psi_n(y)dy$$
$$+\lambda_n A_n \int_{-a}^{+a} \frac{\partial}{\partial \lambda_n} G(x,y,\lambda)\psi_n(y)dy. \tag{4.8}$$

Equation (4.7) shows that the functions ψ_n are identical with the eigenfunctions $\phi_n(x)$, then, multiplying both sides of (4.8) with $\lambda_n \phi_n(x)$ and integrating over x between $-a \leqslant x \leqslant +a$, yields

$$A_n = -\lambda_n \int_a^b f(y)\,\phi_n(y)\,dy$$
$$\times \Bigl(\int_a^b \phi_n^2(y)dy + \lambda_n^2$$
$$\times \int_{-a}^{+a} \int \frac{\partial}{\partial \lambda_n} G(s,t,\lambda)\phi_n(s)\,\phi_n(t)ds\,dt\Bigr)^{-1}. \tag{4.9}$$

On using (4.4) and (4.9) we see that (4.9) are exactly the expansion coefficients of Eq. (4.3).

It is conjectured that, as in the case of ordinary Hilbert–Schmidt theory, $\omega(x,\lambda) \equiv 0$. If this conjecture is true, combination of (3.4) and (4.3) leads to the generalized Hilbert–Schmidt expansion of (2.12):

$$\psi(\mathbf{r}_<,k) = \psi^{(i)}(\mathbf{r}_<,k) + \sum_n \sum_{l=0}^{\infty} \sum_{m=-l}^{+l} Y_l^m(\theta,\phi) k\,(2l+1)$$
$$\times \frac{\chi_l(r_<,k_{ln}) \int_0^a \psi_{lm}^{(i)}(\tau) U(\tau) \chi_l(\tau,k_{ln}) d\tau}{(k_{ln} - k)(1 + k_{ln}^2 \int_0^a \int W(s,t,k_{ln}) \chi_l(s,k_{ln})}$$
$$\times \chi_l(t,k_{ln}) U(s) U(t)\,ds\,dt \tag{4.10a}$$

if

$$\psi_{lm}^{(i)}(\tau,k) = \int_\Omega d\Omega \psi^{(i)}(\tau,k) Y_l^m(\theta,\phi) \tag{4.10b}$$

and

$$W(s,t,k_{ln}) = (2l+1)\, j_l(k_{ln}s)\, h_l^{(1)}(k_{ln}t), \quad \text{if } s<t,$$
$$= (2l+1)\, j_l(k_{ln}t) h_l^{(1)}(k_{ln}s), \quad \text{if } s>t. \tag{4.10c}$$

For recent developments concerning the conjecture $\omega(x,\lambda)=0$ we refer to Hoenders[14] and Pattanayak.[15]

DISCUSSION

The basic equations of this paper [(1.1) and (1.2)] are derived from the so-called quantum mechanical extinction theorem. This theorem is obtained by means of a procedure with which recently a *macroscopical* electromagnetic extinction theorem has been derived (Wolf[7]).

According to the Ewald–Oseen extinction theorem of molecular optics, the electromagnetic field due to an incoming wave inside a medium whose response is expressible as due to a set of dipoles can be thought of as the sum of two terms. One of these terms exactly cancels the incoming wave at every point inside the medium, and the other then gives rise to the actual macroscopical field.

The cancellation of the incoming wave is mathematicaly expressed by the extinction theorem (Born and Wolf[18]), the fundamental role of which for the foundations of crystal optics was already known for about 60 years from Ewald's pioneering researches, but the true meaning of which was not fully understood until very recently. During the last few years the connection between electromagnetic theory and the extinction theorem was thoroughly investigated by several authors, (Sein,[19] Wolf,[7] Pattanayak[5,6] de Goede and Mazur[20]). They all reached the conclusion that the commonly made assumption relating to the validity of this theorem for the microscopical Maxwell equations is too restrictive and that similar theorems can be derived for the macroscopic Maxwell equations as well. Wolf and Pattanayak then conjectured that the extinction theorem is to be understood as a nonlocal boundary condition to which every solution of Maxwell's equations is subjected.

In this way they completely changed the status of the extinction theorem from a *theorem* applicable only to special problems into a *principle* to be satisfied by every solution of Maxwell's equations.

The basic equations of this paper are derived from this principle. Because the complete set of functions considered in this paper is not generated by a Sturm–Liouville problem, we might expect that this set is perhaps overcomplete. This conjecture is true, as has been indicated by Humblet and Rosenfeld,[3] and a proof of this statement is given by Hoenders.[8,14]

The potential considered in this paper is rotationally symmetric, and we are therefore lead to the question if the natural modes connected with an "arbitrary" cutoff potential are also complete within the range of the potential. The completeness of such sets of natural modes has been proven by Hoenders,[14] using the inhomogeneous integral equation (2.12), with $\psi^{\rm inc}$ replaced by an "arbitrary" function $f(\mathbf{r})$, instead of using the Schrödinger equation (2.13) and the boundary condition (2.14).

The reason for the construction of the proof contained in this paper is that this particular technique [Eqs. (3.17), (3.22), and (3.25)] is rather simple and can be applied to similar problems which are not easily analyzed by the methods of the other proof. As an example we mention the problem of the string, discussed in the Introduction.

ACKNOWLEDGMENT

The author wishes to thank Dr. H.A. Ferwerda and Professor E. Wolf for a critical reading of the manuscript. He is obliged to Professor Wolf for drawing his attention to the problems encountered in formulating a physically satisfactory definition of natural modes in quantum mechanical potential scattering theory. The author also wishes to thank Dr. A.M.J. Huiser for stimulating discussions.

[1]P.L. Kapur and R. Peierls, Proc. R. Soc. London A **166**, 277 (1938).
[2]A.J. Siegert, Phys. Rev. **56**, 750 (1939).
[3]J. Humblet and L. Rosenfeld, Nucl. Phys. **26**, 529 (1961); J. Humblet, Mém. Soc. R. Sci. Liège **12**, No. 4 (1952).
[4]R.M. More and E. Gerjuoy, Phys. Rev. A **7**, 1288 (1973).
[5]D.N. Pattanayak and E. Wolf, Phys. Rev. D **13**, 913 (1976).
[6]D.N. Pattanayak and E. Wolf, Phys. Rev. D **13**, 2287 (1976).
[7]A.L. Cauchy, *Oeuvres de Cauchy*, (Gauthier-Villars, Paris, 1908), Ser. 1, Vol. 2, p. 20.
[8]B.J. Hoenders, J. Phys. A, **11**, 1815 (1978).
[9]P.M. Morse and H. Feshbach, *Methods of Theoretical Physics* (McGraw-Hill, New York, 1953), Vol. II, §11.1.
[10]H.M. Nussenzveig, *Causality and Dispersion Relations* (Academic, New York and London, 1972), §5.7.
[11]E. Wolf, in *Coherence and Quantum Optics*, edited by L. Mandel and E. Wolf (Plenum, New York, 1973), p. 339. See also D.N. Pattanayak and E. Wolf, Opt. Comm. **3**, 217 (1972).
[12]R.G. Newton, *Scattering Theory of Waves and Particles* (McGraw-Hill, New York, 1966).
[13]B.J. Lewin, *Nullstellenverteilung Ganzer Funktionen* (Akademie-Verlag, Berlin, 1962), Chap. 4, §7, Theorem 18.
[14]B.J. Hoenders, in *Proceedings of the Fourth Rochester Conference on Coherence and Quantum Optics*, to be published.
[15]D.N. Pattanayak, in Ref. 14.
[16]C. Miranda, Acta Pontificia Scientiarum **III**, 1 (1939).
[17]D.N. Pattanayak, private communication.
[18]M. Born and E. Wolf, *Principles of Optics* (Pergamon, New York, 1975), 5th ed., Sec. 2.4.2. See also L. Rosenfeld, *Theory of Electrons* (North-Holland, Amsterdam, 1951), Chap. VI, Sec. 4.
[19]J.J. Sein, Ph.D. thesis (New York University, 1969) (unpublished).
[20]J. De Goede and P. Mazur, Physica **58**, 568 (1972).

Canonical Schwartz spaces and generalized operators

S. N. Peck

Mathematics Department, University of Nottingham, University Park, Nottingham, England
(Received 24 March 1978)

For a C^∞ multiplier ω, on R^n we define the concepts of differentiability and codifferentiability in the Von Neumann algebra generated by the regular ω representation of R^n. Analogs of the classical Schwartz space and its dual are formulated and the case where ω is fully antisymmetric is studied. Connections with the canonical Fourier transform of an earlier paper are investigated.

INTRODUCTION

Let p,q be a canonical pair of self-adjoint operators on a Hilbert space, H, i.e., self-adjoint operators such that $\exp(isp)\exp(itq) = \exp(ist)\exp(itq)\exp(isp)$ for all real s and t. It follows that for real a and b, $p + a, q + b$ are another canonical pair, and so, by the von Neumann uniqueness theorem, there is an automorphism of the von Neumann algebra generated by $\{\exp(isp),\exp(itq):s,t\in R\}$ mapping $\exp(isp)$ to $\exp[is(p + a)]$ and $\exp(itq)$ to $\exp[it(q + b)]$. This automorphism is called translation through (a,b). We define a concept of cotranslation through (a,b) in this von Neumann algebra, which will be shown to be complementary to translation through (a,b). An element A of the algebra is said to be differentiable (respectively codifferentiable) with respect to p if the action on A of the infinitesimal generator of translation (respectively cotranslation) through $(a,0)$ gives an element of the algebra. Differentiation and codifferentiation with respect to q are defined analgously. The canonical Schwarz space is defined as the space of elements of the algebra which are infinitely differentiable and codifferentiable.

The Weyl transform is an isometry from $L^2(R^2)$ into the space of Hilbert–Schmidt operators on H which for $f\in L^1(R^2)\cap L^2(R^2)$ is given by

$$f\rightarrow\int \exp[i(xp+yq)]f(x,y)dx\,dy.$$

In Ref. 1, the image under the Weyl transform of $\mathscr{S}(R^2)$, the Schwarz space of infinitely differentiable functions of rapid decrease, is studied. We show that this image is precisely the canonical Schwarz space.

The whole situation may be generalized to the von Neumann algebras generated by ω-representations of R^{2n} where ω is an infintiely differentiable multipler. The preceding theory is the special case when $n=1$ and ω is totally antisymmetric. For trivial ω, the theory reduces to the classical theory of Schwarz spaces.

1. The spaces $\mathscr{S}(\omega,R^n)$

Let R^n denote the additive group of n-dimensional real space, which together with its usual topology has the structure of a locally compact Abelian topological group. By a multiplier on R^n, we shall mean a Borel measurable function

$$\omega:R^n\times R^n\rightarrow T$$

(where T denotes the multiplicative group of unimodular complex numbers) satisfying:

(i) $\omega(x:0)=\omega(0:y)=1$,

(ii) $\omega(x:y)\omega(x+y:z)=\omega(x:y+z)\omega(y:z)$.

Two multipliers ω, $\tilde{\omega}$ are said to be *similar* if there exists a Borel measurable function

$$h:R^n\rightarrow T,$$

such that

$$\tilde{\omega}(x:y)=\frac{h(x)h(y)}{h(x+y)}\omega(x:y).$$

For any multiplier ω, the regular ω-representation R_ω is a map from R^n into the space of bounded operators on the Hilbert space $L^2(R^n)$, given by

$$R_\omega(s)f(x)=\omega(x,s)f(x+s) \text{ for all } f\in L^2(R^n).$$

$R_\omega(s)$ is clearly unitary and

$$R_\omega(s)R_\omega(t)=\omega(s,t)R_\omega(s+t),$$

$$R_\omega(0)=1,$$

showing that R_ω is indeed an ω-representation. It is easy to see that R_ω is strongly continuous if ω is jointly continuous.

Let $\mathscr{V}(\omega,R^n)$ be the von Neumann algebra generated by $\{R_\omega(s):s\in R^n\}$.

Lemma: Let ω, $\tilde{\omega}$ be similar multipliers with

$$\tilde{\omega}(x,y)=\frac{h(x)h(y)}{h(x+y)}\omega(x,y).$$

Then $\mathscr{V}(\omega,R^n)$, $\mathscr{V}(\tilde{\omega},R^n)$ are spatially isomorphic.

Proof: We have that

$$R_{\tilde\omega}(s)=h(s)M_hR_\omega(s)M_h^{-1},$$

where $M_hf(x)=h(x)f(x)$ for all $f\in L^2(R^n)$.

It follows immediately that $\mathscr{V}(\omega,R^n)$, $\mathscr{V}(\tilde{\omega},R^n)$ are spatially isomorphic under the map

$$A\rightarrow M_hAM_h^{-1}.$$

For any element $a\in R^n$ there is a natural automorphism of $\mathscr{V}(\omega,R^n)$ induced by a, known as *translation through a*. Let $T(a)$ be the operator in $L^2(R^n)$ given by

$$T(a)f(x)=\exp[i(a,x)]f(x) \text{ for all } f\in L^2(R^n).$$

Then we define translation through a, t_a by

 0022-2488/79/020336-08$01.00

$t_a A = T(a)^{-1} A T(a), \quad A \in \mathcal{N}(\omega, R^n).$

In particular we have $t_a R_\omega(s) = \exp[i(a,s)] R_\omega(s)$ showing that t_a maps the generators $R_\omega(s)$, and hence the whole of $\mathcal{N}(\omega,R^n)$, into $\mathcal{N}(\omega,R^n)$.

Definition: Let $A \in \mathcal{N}(\omega,R^n)$ and let $\{e_i\}$ be an orthonormal basis for R^n. Then we say that A is differentiable in the jth direction if the weak operator limit

$$(\mathrm{w}) \lim_{s \to 0} \frac{1}{s} \{ t_{se_j} A - A \},$$

exists in $\mathcal{N}(\omega,R^n)$. The limit, when it exists, is denoted by $\Delta_j A$.

We now introduce an operation in $\mathcal{N}(\omega,R^n)$ which, we shall see later, is in some sense conjugate to translation.

Definition: Let $a \in R^n$, then for $A \in \mathcal{N}(\omega,R^n)$ the *cotranslation of A through a*, $t^a A$ is defined by

$$t^a A = R_\omega(a) A R_\omega(a).$$

Observe that $t^a A t^a B \neq t^a(AB)$ in general, and that $t^b t_a A = \exp[-2i(a,b)] t_a t^b A$.

Definition: Let $A \in \mathcal{N}(\omega,R^n)$ we say that A is codifferentiable in the jth direction if the weak operator limit

$$(\mathrm{w}) \lim_{s \to 0} \frac{1}{s} \{ t^{se_j} A - A \}$$

exists in $\mathcal{N}(\omega,R^n)$. If the limit exists it is denoted by $\Delta^j A$.

Let a,b be the positive integral n-tuples $(a_1,\ldots,a_n)$, $(b_1,\ldots,b_n)$. Then let Δ_a, Δ^b denote $(\Delta_1)^{a_1}(\Delta_2)^{a_2}\cdots(\Delta_n)^{a_n}$ and $(\Delta^1)^{b_1}(\Delta^2)^{b_2}\cdots(\Delta^n)^{b_n}$, respectively.

Definition: The space $\mathcal{S}(\omega,R^n)$ is defined as the space of elements A of $\mathcal{N}(\omega,R^n)$ for which the mixed derivatives $\Delta^b \Delta_a A$ exist for all positive, integral n-tuples a and b. We define a locally convex topology on $\mathcal{S}(\omega,R^n)$ by defining the semi norms $\|\cdot\|_a^b$ by

$$\|A\|_a^b = \|\Delta^b \Delta_a A\|,$$

where $\|\cdot\|$ is the operator norm in $L^2(R^n)$. We shall be particularly interested in C^∞ *multipliers*, i.e., those multipliers which are infinitely differentiable. Associated with this definition of a C^∞ multiplier is the concept of two multipliers being C^∞-similar, meaning that the function connecting them can be taken to be a C^∞-function.

Bargmann[2] has shown that each C^∞-multiplier on R^n is similar to a multiplier ω^n_{2m}, for some m, where with a suitable decomposition $R^n \cong R^m \oplus R^m \oplus R^k$ we have

$$\omega^n_{2m}(x,y,z;x',y',z') = \exp\{[(x,y') - (y,x')]\, i/2\}.$$

Theorem: Let $\omega, \tilde{\omega}$ be C^∞-similar C^∞-multipliers. Then $\mathcal{S}(\omega,R^n)$ and $\mathcal{S}(\tilde{\omega},R^n)$ are topologically isomorphic.

Proof: Suppose $\tilde{\omega}(x,y) = [h(x)h(y)\omega(x,y)]/h(x+y)$ where h is C^∞, then we know that $\mathcal{N}(\omega,R^n)$ is naturally isomorphic to $\mathcal{N}(\tilde{\omega},R^n)$ via the map sending $A \in \mathcal{N}(\omega,R^n)$ to $M_h A M_h^{-1} \in \mathcal{N}(\tilde{\omega},R^n)$. Denoting this map by ϕ we have

$$\phi(t_a A) = M_h T(a)^{-1} A T(a) M_h^{-1} = T(a)^{-1} M_h A M_h^{-1} T(a)$$

$$= t_a(\phi A).$$

Hence the structure of translation and differentiation is preserved by a change to a C^∞-similar multiplier. Now,

$$\begin{aligned} \phi\,[t^a R_\omega(S)] f(x) &= M_h R_\omega(2a+s) M_h^{-1} f(x) \\ &= h(x) h^{-1}(x+2a+s)\omega(x,2a+s) f(x+2a+s) \end{aligned}$$

and

$$\begin{aligned} t^a[\phi R_\omega(s)] f(x) &= R_{\tilde{\omega}}(a) M_h R_\omega(s) M_h^{-1} R_{\tilde{\omega}}(a) f(x) \\ &= h(x+a) h^{-1}(x+a+s) \tilde{\omega}(x,a)\omega(x+a,s) \\ &\quad \times \tilde{\omega}(x+a+s,a) f(x+2a+s) \\ &= h(x) h(a)^2 h^{-1}(x+2a+s)\omega(x+2a+s) f(x+2a+s). \end{aligned}$$

So

$$t^a[\phi R_\omega(s)] = h(a)^2 \phi[t^a R_\omega(s)].$$

Thus since h is C^∞ and since $\mathcal{N}(\omega,R^n)$ consists of the strong closure of finite linear combinations of the elements $R_\omega(s)$,

$$\Delta^i(\phi A) = \phi(\Delta^i A) + 2h'(0)\phi(A),$$

wherever $\Delta^i A$ exists.

Hence the image of $\mathcal{S}(\omega,R^n)$ under ϕ is $\mathcal{S}(\tilde{\omega},R^n)$ and it follows from the above that the topologies are equivalent.

2. The case $\omega = 1$

We shall study the case in which $\omega = 1$, i.e., $\omega(x,y) = 1$ for all all x,y and show that in this case, the space $\mathcal{S}(1,R^n)$ is naturally isomorphic to the classical Schwartz space $\mathcal{S}(R^n)$.

Let W denote the classical Fourier transform from $L^2(R^n)$ to $L^2(R^n)$, given by the formula

$$Wf(x) = (2\pi)^{-m/2} \int \exp[-i(x,s)] f(s)\,ds,$$

for $f \in L^1(R^n) \cap L^2(R^n)$.

Then we have

$$W R_1(s) W^{-1} f(s) = \exp(isx) f(x) = M(s) f(x).$$

Hence the von Neumann algebra $\mathcal{M}(R^n)$ generated by $(M(s): s \in R^n\}$ is isomorphic to $\mathcal{N}(1,R^n)$ and consists of multiplications in $L^2(R^n)$ by bounded measurable functions which, as an abstract algebra is isomorphic to the algebra $L^\infty(R^n)$ of bounded measurable functions (with pointwise multiplication). We shall now study the operations induced in $L^\infty(R^n)$ by translation and cotranslation in $\mathcal{N}(1,R^n)$. Since

$$W[t_a R_1(s)] W^{-1} f(x) = \exp[is(x+a)] f(x),$$

translation in $\mathcal{S}(1,R^n)$ induces translation through a, in the classical sense, in $L^\infty(R^n)$.

$$W[t^a R_1(s)] W^{-1} f(x) = \exp(2iax)\exp(isx) f(x)$$

and so contranslation in $\mathscr{N}(1,R^n)$ induces multiplication by $\exp(2iax)$ in $L^\infty(R^n)$.

From the above it is easy to see that the operations induced in $L^\infty(R^n)$ by Δ_j and Δ^j are, respectively, (classical) differentiation with respect to the jth variable, and multiplication by $2i$ times the jth variable. Now, the norm induced in $L^\infty(R^n)$ by the operator norm in $\mathscr{N}(1,R^n)$ is simply the supremum norm. Thus if the element $A\in\mathscr{S}(1,R^n)$ corresponds to the function $g\in L^\infty(R^n)$ we have

$$\|A\|_a^b=\|\Delta^b\Delta_a A\|$$

$$=\sup_{R^n}\left| x_1^{b_1}\cdots x_n^{b_n}\frac{\partial^{|a|}}{\partial x_1^{a_1}\cdots\partial x_n^{a_n}}g(x_1\cdots x_n)\right|,$$

i.e., the norms $\|\cdot\|_a^b$ (a, b positive integral n-tuples), corresponds to the usual norms in (R^n). Thus we have proved

Lemma: $\mathscr{S}(1,R^n)$ is isomorphic, topologically and algebraically, to $\mathscr{S}(R^n)$.

3. The case $\omega=\omega_n$

In the case that $n=2m$ is even, we take ω in its canonical form,

$$\omega(x,y;x',y')=e^{i[(x,y')-(y,x')]}.$$

We can reformulate the preceding theory in the language of quantum mechanics as follows.

Since $R_\omega(s_ie_i)$, $R_\omega(t_je_{m+j})$ are strongly continuous unitary representations of the real line, we have, by Stone's theorem that

$$R_\omega(s_ie_i)=\exp(is_ip_i),\quad R_\omega(t_je_{m+j})=\exp(it_jq_j),$$

where p_i, q_j $i,j=1,\dots,m$ are (unbounded) self-adjoint operators. We have the following relationships known as the Weyl commutation relations,

$$\exp(is_ip_i)\ \exp(it_jq_j)$$
$$=\begin{cases}\exp(is_it_j)\quad\exp(it_jq_j)\quad\exp(is_ip_i), & i=j,\\ \qquad\qquad\quad\ \exp(it_jq_j)\quad\exp(is_ip_i), & i\neq j\end{cases}\tag{A}$$

We find that

$$t_{(a,b)}\exp\left[i\left(\sum s_ip_i+t_iq_i\right)\right]$$
$$=\exp\left\{i\left[\sum s_i(p_i+a_i)+t_i(q_i+b_i)\right]\right\}$$
$$=\exp\left[-i\left(\sum a_iq_i-b_ip_i\right)\right]\exp\left[i\left(\sum s_ip_i+t_iq_i\right)\right]$$
$$\times\exp\left[i\left(\sum a_iq_i-b_ip_i\right)\right]\tag{B}$$

and

$$t^{(a,b)}\exp\left[i\left(\sum s_ip_i+t_iq_i\right)\right]$$
$$=\exp\left[i\left(\sum a_ip_i+b_iq_i\right)\right]\exp\left[i\left(\sum s_ip_i+t_iq_i\right)\right]$$
$$\times\exp\left[i\left(\sum a_ip_i+b_iq_i\right)\right].$$

According to the von Neumann uniqueness theorem, the von Neumann algebra $\mathscr{N}(p_1\cdots p_n q_1\cdots q_n)$ generated by

$$\{\exp\left[i\left(\sum s_ip_i+t_iq_i\right)\right]:s_i,t_i\in R\},$$

where p_i, q_i are any essentially self-adjoint operators on a common dense domain, satisfying (A), is canonically isomorphic to the algebra $B(L^2(R^m))$ consisting of the algebra of all bounded operators on $L^2(R^m)$. The canonical isomorphism maps

$$\exp(is_ip_i)\quad\text{to}\quad 1\otimes\cdots\otimes\exp(is_ip_0)\otimes\cdots\otimes 1$$

and

$$\exp(it_jq_j)\quad\text{to}\quad 1\otimes\cdots\otimes\exp(it_jq_0)\otimes\cdots\otimes 1,$$

where p_0,q_0 are the unique self-adjoint extensions of the operators defined on $\mathscr{S}(R)$ by

$$p_0f(x)=-i\frac{df}{dx}(x),$$

$$q_0f(x)=xf(x).$$

(p_0,q_0) are known as the Schrodinger pair and the corresponding representation of the Weyl commutation relations is known as the Schrodinger representation. We may now define translation and cotranslation in $\mathscr{N}(p_1\cdots p_m,q_1\cdots q_n)$ by using (B). Since the generalization to higher values of n is straightforward we shall study the case $n=2$.

We shall make a change of notation and denote

Δ_1, Δ_2, Δ^1, Δ^2 by Δ_p, Δ_q, Δ^p, Δ^q, respectively.

Lemma: Let $A\in\mathscr{S}(p,q)$. Then A and its adjoint A^* map $L^2(R)$ into $\mathscr{S}(R)$, the Schwarz space of infinitely differentiable complex valued functions which decrease at infinity together with their derivatives, faster than any polynomial.

Proof: Put

$$\Gamma_pA=\frac{1}{2i}(\Delta^pA+\Delta_qA),$$

$$\Gamma_qA=\frac{1}{2i}(\Delta^qA-\Delta_pA).$$

Then

$$\Gamma_pA=(\mathrm{w})\ \lim_{a\to 0}\frac{1}{2ia}[\exp(iap_0)-\exp(-iap_0)]A$$

$$=(\mathrm{w})\ \lim_{a\to 0}[V(a)A]$$

and

$$\Gamma_q A = (\mathrm{w}) \lim_{a\to 0} \frac{1}{2ib}[\exp(ibq_0) - \exp(-ibq_0)]A$$

$$= (\mathrm{w}) \lim_{b\to 0} [W(b)A].$$

Let $f \in \mathscr{S}(R)$ and $g \in L^2(R)$. Then

$\langle Ag, p^m q^n f\rangle$

$$= \lim_{a_1 \cdots a_m, b_1 \cdots b_n \to 0} \langle Ag, V(a_1)\cdots V(a_m)W(b_1)\cdots W(b_n)f\rangle$$

$$= \lim_{a_1 \cdots a_m, b_1 \cdots b_n \to 0} \langle W^*(b_n)\cdots W^*(b_1)V^*(a_n)\cdots V^*(a_1)Ag, f\rangle$$

$= \langle g^*, f\rangle.$

Hence Ag is contained in $D([p^m q^n]^*)$, the domain of $[p^m q^n]^*$ and hence $Ag \in \mathscr{S}(R)$ since it known[3] that

$$\bigcap_{m,n} D([p^m q^n]^*) = \mathscr{S}(R).$$

Since if X is an element of $\mathscr{S}(p,q)$, then so is X^*, and hence X^* maps $L^2(R)$ into $\mathscr{S}(R)$.

Theorem: $\mathscr{S}(p_0,q_0)$ consists of all operators A of the form $Ah(x) = \int d(x,y)h(y)\,dy$ with $\in \mathscr{S}(R^2)$. The map $A \to d$ is a topological isomorphism of $\mathscr{S}(p_0,q_0)$ with $\mathscr{S}(R^2)$.

Proof: Let A be an operator in $\mathscr{N}(p_0,q_0) = B(L^2(R))$ given by

$$Ah(x) = \int d(x,y)h(y)\,dy, \text{ with } d \in \mathscr{S}(R^2).$$

Then $t_{(a,b)}Ah(x) = \int \exp[ia(y-x)]d(x+b,y+b)h(y)\,dy$. Thus

$$\lim_{a\to 0}\left\langle \frac{1}{a}[t_{(a,0)}A - A]h, k\right\rangle$$

$$= \lim_{a\to 0}\int\int \frac{1}{a}\{\exp[ia(y-x)] - 1\}$$

$$\times d(x,y)h(y)\overline{k(x)}\,dx\,dy$$

$$= \int\int i(y-x)d(x,y)h(y)\overline{k(x)}\,dx\,dy,$$

since the convergence of the integrand is uniform since $d \in \mathscr{S}(R^2)$.

Hence

$$\Delta_p Ah(x) = \int i(y-x)d(x,y)h(y)dy.$$

Also,

$$\lim_{b\to 0}\left\langle \frac{1}{b}[t_{(0,b)}A - A]h, k\right\rangle$$

$$= \lim_{b\to 0}\int\int \frac{1}{b}[d(x+b,y+b) - d(x,y)]h(y)\overline{k(x)}\,dx\,dy$$

$$= \int\int\left(\frac{\partial}{\partial x} + \frac{\partial}{\partial y}\right)d(x,y)h(y)\overline{k(x)}\,dx\,dy,$$

since, again, the convergence of the integrand is uniform. Hence

$$\Delta_q Ah(x) = \int\left(\frac{\partial}{\partial x} + \frac{\partial}{\partial y}\right)d(x,y)h(y)\,dy.$$

Similarly we find that

$$\Delta^p Ah(x) = \int\left(\frac{\partial}{\partial x} - \frac{\partial}{\partial y}\right)d(x,y)\cdot h(y)\,dy,$$

$$\Delta^q Ah(x) = \int i(x+y)d(x,y)\cdot h(y)\,dy.$$

Thus $\Delta_p A$, $\Delta_q A$, $\Delta^p A$, $\Delta^q A$ are of the same form as A, i.e., integral operators with kernels in $\mathscr{S}(R^2)$. Proceeding inductively we find that A has all derivatives and coderivatives of every order, i.e., $A \in \mathscr{S}(p_0,q_0)$. Hence the class of operators of this form is contained in $\mathscr{S}(p_0,q_0)$.

We now show that if $A \in (p_0,q_0)$ then A is a Hilbert–Schmidt operator, i.e., $\Sigma\|A\phi_i\|^2 < \infty$ for every complete orthonormal set $\{\phi_i\}$. The operator $(p_0^2 + q_0^2)^{-1}$ satisfies

$$(p_0^2 + q_0^2)^{-1}h_n = \frac{1}{(2n+1)}h_n,$$

where h_n is the nth normalized hermite function,

$$h_n(x) = \frac{(-1)^n}{\sqrt{2^n n!}\pi^{1/2}}\cdot\exp(\tfrac{1}{2}x^2)\frac{d^n}{d^n}[\exp(-x^2)].$$

Hence

$$\sum\|(p_0^2 + q_0^2)^{-1}h_n\|^2 = \sum\frac{1}{(2n+1)^2} < \infty,$$

and so $(p_0^2 + q_0^2)^{-1}$ is a Hilbert–Schmidt operator. Now since A maps $L^2(R)$ into $\mathscr{S}(R)$, $(p_0^2 + q_0^2)A$ is well defined and

$$(p_0^2 + q_0^2)A = (\Gamma_p)^2 A + (\Gamma_q)^2 A.$$

Thus $(p_0^2 + q_0^2)A \in \mathscr{N}(p_0,q_0) = B(L^2(R))$. Since $(p_0^2 + q_0^2)^{-1}(p_0^2 + q_0^2)A = A$, A is Hilbert-Schmidt since the class of Hilbert–Schmidt operators is an ideal in $B(L^2(R))$.

Now since the class of Hilbert–Schmidt operators on $L^2(R)$ is precisely those operators which may be written as integral operators with kernel in $L^2(R^2)$, it follows that the class of operators with kernels in $\mathscr{S}(R^2)$ is dense in the class of Hilbert–Schmidt operators and thus in $\mathscr{S}(p_0,q_0)$.

Now we only need show that the class of operators with kernels in $\mathscr{S}(R^2)$ is closed under the topology of $\mathscr{S}(p_0,q_0)$ and that this topology is equivalent to the standard topology on $\mathscr{S}(R^2)$.

We define a new system of norms on $\mathscr{S}(p_0,q_0)$ by

$$\|A\|_{a,2}^{b} = \|\Delta^b\Delta_a A\|_2,$$

where $\|A\|_2$ is the Hilbert–Schmidt norm given by

$$\|A\|_2 = \sum_i \|A\phi_i\|^2 \quad \text{for some complete orthonormal set}$$

$\{\phi_i\}$.

If A has kernel d in (R^2), then $\|A\|_2 = \|d\|$, the L^2 norm of d. Thus

$$\|A\|^b_{a,2} = \left\| x^{b'} \frac{\partial}{\partial x^{a_1}} d \right\|_2 = \|d\|^{b'}_{a',2},$$

where a', b' are related, linearly, to a and b. But the topology on $\mathscr{S}(R^2)$ induced by the seminorms $\|d\|^{b'}_{a',2}$ is known[4] to be equivalent to the topology induced by the seminorms,

$$\|d\|^{b'}_{a'} = \sup_{R^2} \left| x^{b'} \frac{\partial}{\partial x^{a'}} d(x) \right|.$$

Hence the result is proved if it can be shown that the topologies on $\mathscr{S}(p_0,q_0)$ are equivalent.

Since $\|A\| \leqslant \|A\|_2$, it follows that

$$\|A\|^b_a \leqslant \|A\|^b_{a,2}.$$

Conversely for $A \in \mathscr{S}(p_0,q_0)$ we have

$$A = (p_0^2 + q_0^2)^{-1}(p_0^2 + q_0^2)A$$

and using the inequality

$$\|AB\|_2 \leqslant \|A\|_2 \|B\|,$$

we have

$$\|AB\|_2 \leqslant \|(p_0^2 + q_0^2)^{-1}\|_2 \|(p_0^2 + q_0^2)A\|,$$

so $\|A\|^b_{a,2} \leqslant C\|A\|^{b'}_{a'}$ for some a', b' and hence the result is proved.

It immediatly follows that for an arbitrary canonical pair (p,q) $\mathscr{S}(p,q)$ is toplogically isomorphic to $\mathscr{S}(R^2)$ since $\mathscr{S}(p,q)$ is topologically isomorphic to $\mathscr{S}(p_0,q_0)$ under the map induced by the canonical map from $\mathscr{N}(p,q)$ to $\mathscr{N}(p_0,q_0)$. For $A \in \mathscr{S}(p,q)$ we say A has kernel d if d is the kernel of the operator in $\mathscr{S}(p_0,q_0)$ which is the image of A under the canonical map.

4. Theorem

Theorem: The space $\mathscr{S}(\omega^n_{2m}, R^n)$ is topologically isomorphic to the Schwarz space $\mathscr{S}(R^n)$.

Proof: Let $\mathscr{S}(\omega^{2m}_{2m}, R^{2m}) \otimes \mathscr{S}(1,R^k)$ denote the algebraic tensor product of $\mathscr{S}(\omega^{2m}_{2m}, R^{2m})$ and $\mathscr{S}(1,R^k)$. Since $\mathscr{N}(\omega^n_{2m}, R^n)$ is naturally isomorphic to the norm completion of the algebraic tensor product $\mathscr{N}(\omega^{2m}_{2m}, R^m) \otimes \mathscr{N}(1,R^k)$, and since $\Delta^{b_1}_{a_1} A \otimes A^{b_2}_{a_2} B = \Delta^{(b_1,b_2)}_{(a_1,a_2)} (A \otimes B)$, it follows that $\mathscr{S}(\omega^{2m}_{2m}, R^m) \otimes \mathscr{S}(1,R^k)$ is a dense subset of $\mathscr{S}(\omega^n_{2m}, R^n)$. We know that $\mathscr{S}(\omega^{2m}_{2m}, R^{2m})$ is topologically isomorphic to $\mathscr{S}(R^{2m})$ and that $\mathscr{S}(1,R^k)$ is topologically isomorphic to $\mathscr{S}(R^k)$. Denote these isomorphisms by ϕ, ψ, respectively. But the topology on $\mathscr{S}(R^{2m})$ can be generated by a system of inner products $\langle\cdot,\cdot\rangle_n$, and the topology on $\mathscr{S}(R^k)$ by a system of inner products $\langle\cdot,\cdot\rangle_m$.[4]

Define the inner products $\langle\cdot,\cdot\rangle'_n$, $\langle\cdot,\cdot\rangle'_m$ on $\mathscr{S}(\omega^{2m}_{2m}, R^{2m})$, $\mathscr{S}(1,R^k)$ respectively, by:

$$\langle A,B\rangle'_n = \langle \phi A, \phi B\rangle_n,$$

$$\langle C,D\rangle'_m = \langle \psi C, \psi D\rangle_m.$$

Then the topologies on $\mathscr{S}(\omega^{2m}_{2m}, R^{2m})$ and $\mathscr{S}(1,R^k)$ generated by the inner products $\langle\cdot,\cdot\rangle'_n$, $\langle\cdot,\cdot\rangle'_m$, respectively, are equivalent to their natural topologies.

Define the topology on tensor product $\mathscr{S}(\omega^{2m}_{2m}, R^{2m}) \otimes \mathscr{S}(1,R^k)$ by the inner products

$$\left\langle \sum_{i=1}^n A_i B_i, \sum_{j=1}^m C_j D_j \right\rangle'_{(n,m)} = \sum_{i=1}^n \sum_{j=1}^m \langle A_i, C_j\rangle'_n \langle B_i, D_j\rangle'_m.$$

Then this topology is equivalent to the topology of $\mathscr{S}(\omega^n_{2m}, R^n)$ and so it follows that $\mathscr{S}(\omega^n_{2m}, R^n)$ is topologically isomorphic to $\mathscr{S}(\omega^{2m}_{2m}, R^{2m}) \otimes_\sigma \mathscr{S}(1,R_k)$, the completion of the tensor product in the above topology. But since $\mathscr{S}(\omega^{2m}_{2m}, R^{2m})$, $\mathscr{S}(1,R^k)$ are nuclear[5,6] it follows that $\mathscr{S}(\omega^{2m}_{2m}, R^{2m}) \otimes_\sigma \mathscr{S}(1,R^k)$ is the natural completion and hence it is isomorphic to $\mathscr{S}(R^{2m}) \otimes_\sigma \mathscr{S}(R^k)$ which is isomorphic to $\mathscr{S}(R^n)$.

5. The spaces $\mathscr{S}'(\omega, R^n)$

Definition: The space $\mathscr{S}'(\omega,R^n)$ is the space of all continuous linear functionals on $\mathscr{S}(\omega,R^n)$, i.e., the linear functional X is an element of $\mathscr{S}'(\omega,R^n)$ if

$$|(X,A)| \leqslant C\|A\|^b_a$$

for all $A \in (\omega,R^n)$ and for some constant C and nonnegative integral n-tuples a, b, where (X,A) denotes the value of X on A.

$\mathscr{S}'(\omega,R^n)$ has a natural topology as follows. The sequence $\{X_n\} X_n \in \mathscr{S}'(\omega,R^n)$ converges [to $X \in \mathscr{S}'(\omega,R^n)$] if and only if (X_n,A) converges [to (X,A)] for all $A \in \mathscr{S}(\omega,R^n)$.

Theorem: $\mathscr{S}'(\omega,R^n)$ is topologically isomorphic to $\mathscr{S}'(R^n)$.

Proof: This follows immediately from the fact that $\mathscr{S}(\omega,R^n)$ is topologically isomorphic to $\mathscr{S}(R^n)$. We shall be particularly interested in the case $n=2$ and $\omega(x,y:x',y') = \exp[\frac{1}{2}(xy' - x'y)]$, i.e., the space of continuous linear functionals on $\mathscr{S}(p,q)$ which we denote by $\mathscr{S}'(p,q)$. We have a map $X \to \xi$ from $\mathscr{S}'(p,q)$ to $\mathscr{S}(R^2)$ given by

$$(X,A) = (\xi,a) \text{ for all } A \in \mathscr{S}(p,q)$$

with a the kernel of A. We say that ξ is the kernel of X.

Example 1: Let $B \in \mathscr{N}(p_0,q_0)$. Then B naturally defines an element of $\mathscr{S}'(p_0,q_0)$ (also denoted by B) given by

$$(B,A) = \mathrm{tr}(BA), \quad \mathrm{tr}(A) = \sum \langle A\phi_i, \phi_i\rangle$$

with $\{\phi_i\}$ a complete orthonormal set. Since $|\mathrm{tr}(BA)| \leqslant \|B\| \mathrm{tr}|A|$ and

$$\mathrm{tr}|A| \leqslant \mathrm{tr}[(p_0^2 + q_0^2)^{-2}] \|(p_0^2 + q_0^2)^2 A\|$$

it follows that B is continuous.

Example 2: The unbounded operator $p_0^m q_0^n$ is associated with a member of $\mathscr{S}'(p_0,q_0)$ by

$$(p_0^m q_0^n, A) = \mathrm{tr}(p_0^m q_0^n)A$$

$$\leqslant \mathrm{tr}[(p_0^2 + q_0^2)^{-2}] \|(p_0^2 + q_0^2)^2 p_0^m q_0^n A\|,$$

so $p_0^m q_0^n$ is continuous.

Examples 1 and 2 may be generalized to an arbitrary canonical pair (p,q) since the von Neumann algebra $\mathscr{N}(p,q)$ has a translation invariant trace.

Example 3: Define δ_q by

$$(\delta_q,A)=d(0,0)$$

where d is the kernel of A, i.e.,

$$(\delta_q A)=(\delta\otimes\delta,d),$$

where δ is the delta distribution in $\mathscr{S}'(R)$.

Definition: Let $X\in\mathscr{S}'=(p,q)$. Then we define the derivatives and coderivatives of X by

$$(\Delta_p X,A)=-(X,\Delta_p A):(\Delta^p X,A)=(X,\Delta^p A),$$

$$(\Delta_q X,A)=-(x,\Delta_q A):(\Delta^q X,A)=(X,\Delta^q A),$$

for all $A\in\mathscr{S}(p,q)$.

It is easy to show that if $B\in\mathscr{S}(p,q)$ defines an element of $\mathscr{S}'(p,q)$ by $(B,A)=\tau(BA)$ then the derivatives and coderivatives of B as elements of $\mathscr{S}'(p,q)$ are precisely the derivatives and coderivatives of B as an element of $\mathscr{S}(p,q)$.

Example 4: Define the bounded operator M_H in $L^2(R)$ by

$$M_H f(x)=H(x)f(x),$$

with H the Heaviside function

$$H(x)=\begin{cases}1,x\geqslant 0,\\ 0,x<0.\end{cases}$$

Let H_q be given by

$$(H_q,A)=\mathrm{tr}(M_H A), \text{ for all } A\in\mathscr{S}(p,q).$$

Then if A has kernel d,

$$(H_q,A)=\int_0^\infty\int_0^\infty d(x,y)dx\,dy,$$

$$(\Delta_q H_q,A)=-(H_q,\Delta_q A)$$

$$=\int_0^\infty\int_0^\infty\left(\frac{\partial}{\partial x}+\frac{\partial}{\partial y}\right)d(x,y)dx\,dy$$

$$=d(0,0)$$

$$=(\delta_q,A),$$

i.e.,

$$\Delta_q H_q=\delta_q.$$

Theorem: Let $X\in\mathscr{S}'(p_0,q_0)$: Then X is a polynomial combination of derivatives and coderivatives of a Hilbert–Schmidt operator $\widetilde{X}$.

Proof: Let $\chi\in\mathscr{S}'(R^2)$ be the kernel of X. Let $\alpha_{nm}=(\chi,h_n h_m)$, where h_p is the pth normalized hermite function.

Then by the regularity theorem for $\mathscr{S}'(R^2)$ [4] there exist positive integers r,s and a constant c such that

$$|\alpha_{nm}|\leqslant C(n+1)^r(m+1)^s \text{ for all } n,m.$$

Let $\alpha_{nm}=\beta_{nm}(n+1)^{r+1}(m+1)^{s+1}$, then

$$\sum_{n,m}|\beta_{nm}|^2=\sum_{n,m}|\alpha_{nm}|^2(n+1)^{-2r-2}(m+1)^{-2s-2}$$

$$\leqslant C^2\sum_{n,m}(n+1)^2(m+1)^2<\infty.$$

Define $\widetilde{X}=\Sigma\beta_{nm}(h_n\otimes h_m),\in L^2(R^2)$, then if $\widetilde{X}$ is the operator with kernel $\widetilde{\chi}$, $\widetilde{X}$ is Hilbert–Schmidt. Let

$$\Phi=-\tfrac{1}{4}(\Delta^q-\Delta_p)^2-\tfrac{1}{4}(\Delta_q+\Delta^p)^2,$$

$$\Psi=-\tfrac{1}{4}(\Delta^q+\Delta_p)^2-\tfrac{1}{4}(\Delta_q-\Delta^p)^2.$$

Then

$$([\Phi+\tfrac{1}{2}]^{r+1}[\Psi+\tfrac{1}{2}]^{s+1}\widetilde{X},A)$$
$$=(\widetilde{X},[\Phi+\tfrac{1}{2}]^{r+1}[\Psi+\tfrac{1}{2}]^{s+1}A).$$

If A has kernel $\Sigma\gamma_{nm}h_n\otimes h_m$, then the R.H.S. equals

$$(\widetilde{X},\sum\gamma_{nm}(n+1)^{r+1}(m+1)^{s+1}h_n h_m)$$

$$=\sum\gamma_{nm}(n+1)^{r+1}(m+1)^{s+1}\beta_{nm}$$

$$=\sum\gamma_{nm}\alpha_{nm}=(X,A).$$

Hence

$$X=[\Phi+\tfrac{1}{2}]^{r+1}[\Psi+\tfrac{1}{2}]^{s+1}\widetilde{X}.$$

Each element of $\mathscr{S}'(p,q)$ may be regarded as a map from $\mathscr{S}(R)$ into $\mathscr{S}'(R)$ as follows. Let $f\in\mathscr{S}(R)$ and define $Xf\in\mathscr{S}'(R)$ by

$$(Xf,g)(X,f\otimes g), \text{ for all } g\in\mathscr{S}(R),$$

where $f\otimes g$ is the element of $\mathscr{S}(p,q)$ given by

$$(f\otimes g)h(x)=\int f(x)g(y)h(y)dy.$$

A detailed account has been given in Refs. 7 and 8 of linear mappings from a vector space Φ into some space of linear functionals on Φ, Φ'. The above is a special case of this.

Since $\mathscr{S}(R)$ may be identified with a subspace of $\mathscr{S}(R')$ it seems natural to talk about generalized eigenvectors and eigenvalues of $X\in\mathscr{S}'(p,q)$.

Definition: Let h_p be the pth normalized hermit function, then $f\in\mathscr{S}'(R)$ is said to be a generalized eigenvector of $X\in\mathscr{S}'(p,q)$ with eigenvalue λ if

$$\sum_n(X,g\otimes h_n)(f,h_n)=\lambda(f,g)$$

for all $g\in\mathscr{S}(R)$ where the equation is taken to mean that the L.H.S. converges and equals the R.H.S.

Example: Consider $\tfrac{1}{2}p^2-\delta_q$ as an element of $\mathscr{S}'(p,q)$. Suppose we have

$$\sum_n(\tfrac{1}{2}p^2-\delta_q,g\otimes h_n)(f,h_n)=\lambda(f,g)$$

for some $f\in\mathscr{S}'(R)$ and all $g\in\mathscr{S}(R)$. Then

$$\sum_n[-\tfrac{1}{2}(h_n,g'')(f,h_n)-g(0)\,h_n(0)(f,h_n)]$$

$$=-\tfrac{1}{2}(f,g'')-Ag(0)=\lambda\,(f,g),$$

where $A=\Sigma h_n(0)(f,h_n)$.

The above gives

$$-\tfrac{1}{2}(f'',g)-A\,(\delta,g)=\lambda\,(f,g)$$

i.e.,

$$\tfrac{1}{2}f''+\lambda f=-A\delta,$$

which has the following solutions:

$$\lambda=-\tfrac{1}{2},\quad (f,g)=\int\exp(-|x|)g(x)\,dx,$$

$$\lambda=\tfrac{1}{2}k^2>0,\quad (f,g)=\int_{-\infty}^{\infty}\sin kx\cdot g(x)dx\,,$$

$$\lambda=\tfrac{1}{2}k^2>0,\quad (f,g)=\int_{-\infty}^{0}\cos k\,(x-a)g(x)dx$$

$$+\int_0^{\infty}\cos k\,(x+a)g(x)dx,$$

where $-k=\cot ka$.

6. CONNECTIONS WITH THE FOURIER–WEYL TRANSFORM AND THE CANONICAL FOURIER TRANSFORM

The Fourier–Weyl transform from $L^2(p,q)$, the space of Hilbert-Schmidt operators in $\mathscr{N}(p,q)$, onto $L^2(R^2)$ is defined in Ref. 9. For $A\in\mathscr{S}(p,q)$, the Fourier–Weyl transform $U_\alpha A$ is given by

$$(U_\alpha A)(s,t)=\frac{\alpha}{(2\pi)^{\frac{1}{2}}}\mathrm{tr}\{\exp[-i\alpha(sp+tq)]A\}$$

with α a nonzero real number.

Lemma: U_α maps $\mathscr{S}(p,q)$ continuously onto $\mathscr{S}(R^2)$.

Proof: Let $A\in\mathscr{S}(p,q)$ and let A have kernel $d\in\mathscr{S}(R^2)$. Then $U_\alpha A$ is given by

$$U_\alpha A=(1\otimes W)R_{k(\alpha)}d\,,$$

where W is the Fourier transform in $L^2(R)$, R_M is the operator given by

$$R_M\,d\,(u,v)=d\,([u,v]M),$$

and $k(\alpha)$ is the matrix

$$\begin{pmatrix}-\frac{1}{2}\alpha & \frac{1}{2}\alpha\\ \alpha^{-1} & \alpha^{-1}\end{pmatrix}.$$

But $1\otimes W$ and $R_{k(\alpha)}$ are both isometries of $L^2(R^2)$ mapping $\mathscr{S}(R^2)$ continuously onto itself: Hence U_α maps $\mathscr{S}(p,q)$ continuously onto $\mathscr{S}(R^2)$.

We can now extend U_α to a map (also denoted by U_α) from $\mathscr{S}'(p,q)$ onto $\mathscr{S}'(R^2)$ as follows:

$$(U_\alpha X,f)=(X,\widehat{U}_\alpha f)$$

for all $f\in\mathscr{S}'(R^2)$, where $\widehat{U}_\alpha$ is the map from $\mathscr{S}(R^2)$ onto $\mathscr{S}(p,q)$ given by

$$\widehat{U}_\alpha=(U-\alpha)^{-1}.$$

The extended U_α from $\mathscr{S}'(p,q)$ to $\mathscr{S}'(R^2)$ will be automatically continuous.

Example: We shall compute the Fourier–Weyl transform of $\delta_q\in\mathscr{S}'(p,q)$.

Let $f\in\mathscr{S}(R^2)$, then $\widehat{U}_\alpha f$ has kernel $R_{k(-\alpha)^{-1}}(1\otimes W^{-1})f$.

$$\therefore(\delta_q,\widehat{U}_\alpha f)=R_{k(-\alpha)^{-1}}(1\otimes W^{-1})f(0,0)$$

$$=(1\otimes W^{-1})f(0,0)$$

$$=\frac{1}{(2\pi)^{\frac{1}{2}}}\int f(0,s)ds$$

$$=\left(\frac{1}{(2\pi)^{\frac{1}{2}}}\{\delta\otimes 1\},f\right),$$

i.e.,

$$U_\alpha\delta_q=\frac{1}{(2\pi)^{\frac{1}{2}}}(\delta\otimes 1)\,.$$

The canonical Fourier transform F_α, $0<\alpha<\pi/2$ from $L^2(p,q)$ to $L^2(p',q')$ is defined in Ref. 9. For an element $A\in\mathscr{S}(p,q)$

$$F_\alpha A=\sin\alpha\,\mathrm{tr}_1\{\exp[-i\alpha(p\otimes p'+q\otimes q')](A\otimes 1)\},$$

where tr_1 is the partial trace defined as a function from $\mathscr{S}(p,q)\otimes\mathscr{N}(p',q')$ to $\mathscr{N}(p',q')$ such that

$$\mathrm{tr}[\mathrm{tr}_1(A\otimes B)]=\mathrm{tr}[A\otimes B],\quad \text{for } A\in\mathscr{S}(p,q) \text{ and } B\in\mathscr{S}(p',q').$$

Lemma: F_α maps $\mathscr{S}(p,q)$ continuously onto $\mathscr{S}(p',q')$.

Proof: If $A\in\mathscr{S}(p,q)$ and A has kernel d, then $F_\alpha A$ has kernel b given by

$$b=R_{M(\alpha)}(W\otimes W)d,\quad M_{(\alpha)}=\begin{pmatrix}\operatorname{cosec}\alpha & -\cot\alpha\\ -\cot\alpha & \operatorname{cosec}\alpha\end{pmatrix}.$$

Since $R_{M(\alpha)}$, $W\otimes W$ both map $\mathscr{S}(R^2)$ continuously onto itself, the result follows.

We now extend F_α to a map from $\mathscr{S}'(p,q)$ onto $\mathscr{S}(p',q')$; by defining

$$(F_\alpha X,A)=(X,F_\alpha A)\quad \text{for all } A\in\mathscr{S}(p,q),$$

it follows that the extended F_α is automatically continuous.

Example: We shall compute the canonical Fourier transform of $\delta_q\in\mathscr{S}'(p,q)$. Let $A\in\mathscr{S}(p,q)$ and let A have kernel d.

$$(F_\alpha\delta_q,A)=(\delta_q,F_\alpha A)=\int d\,(u,v)du\,dv,$$

we shall denote $F_\alpha\delta_q$ by δ_p.

Theorem: Let $A\in\mathscr{S}(p,q)$, then

$$F_\alpha \Delta_p = \frac{1}{p_2(\alpha)} \Delta^p F_\alpha A,$$

$$p_1(\alpha) = \operatorname{cosec}\alpha - \cot\alpha,$$
$$p_2(\alpha) = \operatorname{cosec}\alpha + \cot\alpha,$$

$$F_\alpha \Delta^q A = \frac{-1}{p_1(\alpha)} \Delta_q F_\alpha A .$$

Proof: Let A have kernel d. Then $F_\alpha \Delta_p A$ has kernel b given by

$$b(x,y) = \int \exp[-i(x,y)M(\alpha)(u,v)^t i(v-u) d(u,v) du\, dv$$

$$= \frac{1}{p_2(\alpha)} \left\{ \frac{\partial}{\partial x} - \frac{\partial}{\partial y} \right\} \int e^{-i(x,y)M(\alpha)(u,v)^t}$$

$$\times d(u,v) du\, dv.$$

Hence,

$$F_\alpha \Delta_p A = \frac{1}{p_2(\alpha)} \Delta^p F_\alpha A ;$$

the second result is proved similarly.

Corollary: It follows immediately that for $X\in\mathscr{S}'(p,q)$ we have:

$$F_\alpha \Delta_p X = \frac{1}{p_2(\alpha)} \Delta^p F_\alpha X ,$$

$$F_\alpha \Delta^q X = \frac{-1}{p_1(\alpha)} \Delta_q F_\alpha X .$$

[1]R. Lavine, Ph.D. thesis, Massachusetts Institute of Technology (1965).
[2]V. Bargmann, Ann. Math. (N.Y.) **59**, 1–46 (1954).
[3]R. Powers, Commun. Math. Phys. **21**, 85–124 (1971).
[4]M. Reed and B. Simon, *Functional Analysis* (Academic, New York, 1972).
[5]F. Treves, *Topological Vector Spaces, Distributions and Kernels* (Academic, New York, 1967).
[6]I. Gel'fand and N. Vilenkin, *Generalized Functions* (Academic, New York, 1964), Vol. 4.
[7]G. Svetlichny, J. Math. Phys. **11**, 3433 (1970).
[8]G. Svetlichny, Commun. Math. Phys. **33**, 243–51 (1973).
[9]R. Hudson and S. Peck, J. Math. Phys., to appear.

ERRATA

Erratum: "The asymptotic behavior of bound eigenfunctions of Hamiltonians for single variable systems" [J. Math. Phys. 19, 1658 (1978)]

John D. Morgan III

Department of Physics, Princeton University, Princeton, New Jersey 08540
(Received 19 September 1978)

(1) Page 1658, 2nd column, line 6 from bottom should read "(This condition is the "reasonably well-behaved" hypothesis mentioned earlier.) Let $F(x) = W(x)(V(x) - E)$", etc.

(2) Page 1659, below Eq. (12) should read "Φ satisfies $(W\Phi')' = F(r)\Phi$".

CUMULATIVE AUTHOR INDEX

All authors published so far in the current volume are listed alphabetically with the issue and page numbers following the dash. A cumulative author and subject index covering each volume is published annually. An (E) *after the page number indicates an Erratum.*

 0022-2488/79/020344-01$00.00