Proc. Natl. Acad. Sci. USA
Vol. 75, No. 4, pp. 1609–1611, April 1978
Applied Physical and Mathematical Sciences

Symmetry and causality properties of physical fields

(wave equation/conformal group/fundamental particles)

H. P. JAKOBSEN†, B. ØRSTED, I. E. SEGAL, B. SPEH‡, AND M. VERGNE

Department of Mathematics, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

Contributed by I. E. Segal, February 7, 1978

ABSTRACT **Representations of groups of causality-preserving transformations on locally Minkowskian space-times, by actions on classes of wave functions of designated transformation properties, are analyzed, in extension of the conventional theoretical treatment of free relativistic particles. In particular, the constraints of positivity of the energy and finiteness of propagation velocity are developed, and the concept of mass is explored, within the indicated framework.**

The problems of the modelling of space-time (or the "cosmos"), and of elementary particles, seem to have reached a stage where certain very general considerations with clear-cut mathematical interpretations can be brought to bear effectively. The present work aims to develop systematically some of the implications of these considerations, and to specify and derive relevant properties of the possible physical systems. The motivation and intuitive basis are thus physical, but the formulation and techniques are perforce mathematical.

In the latter terms, we are concerned with harmonic analysis of vector bundles over locally Minkowskian spaces; spaces of cross sections defined by hyperbolic partial differential equations; and holomorphy features dual to the desideratum of "positivity of the energy," which in turn is related to causality—as is also, in a different way, the hyperbolicity of the cited equations.

It seems plausible that stable or quasi-stable free elementary particles should be describable in a familiar theoretical way by "fields" that are cross-section spaces of the type indicated. If one foregoes the sanctity of the Lorentz group and insists only on the more fundamental features of temporal and spatial isotropy and homogeneity, another possibility emerges, which has found observationally cogent application to astronomy (see refs. 1–3). This macroscopic indication supplements a variety of previous microscopic ones in suggesting the potential physical relevance of a study of the transformation properties of all particle models of the type designated, under symmetry groups implementing the requisite isotropy and homogeneity.

The basic representations

We recall that a *causal manifold* is a C^∞ manifold that is endowed with a smooth closed convex proper cone field; the *causal group* of any such manifold m is the group denoted $G(m)$ of all diffeomorphisms that preserve the cone field. Spatial and temporal homogeneity and isotropy are then naturally definable, and it is known (see refs. 1, 4, and 5) that the only causal 4-manifolds with these properties are locally Minkowskian; i.e., locally, have the same cone-field structure as the future cone-field structure in Minkowski space.

There are only three such manifolds that are globally causal in the sense that there exist no closed time-like loops in the manifold: Minkowski space M, the universal covering space $\tilde{M}$ of its conformal compactification $\overline{M}$, and a hyperbolic space M^*. Both M and M^* can be naturally realized as open submanifolds of $\tilde{M}$, as orbits under subgroups G_0 and G_1 of $G(\tilde{M})$ that are locally homogeneous and isotropic both spatially and temporally.

Infinitesimally, $G(\tilde{M})$ is isomorphic to the conformal group of M; G_0 to its Poincaré subgroup; and G_1 to an associated subgroup that is locally isomorphic to $SO(2, 3)$. Globally,§ the connected components of these groups are isomorphic to the corresponding components of $\tilde{SO}(2,4)/Z_2$ in the case of $G(\tilde{M})$, in which Z_2 is the unique central subgroup of order 2 in the indicated group; the universal covering group of the Poincaré group, in the case of G_0, and to $\tilde{SO}(2,3)/Z_2$, with a similar

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U. S. C. §1734 solely to indicate this fact.

† Present address: Department of Mathematics, Brandeis University, Waltham, MA 02172.

‡ Present address: Department of Mathematics, University of Chicago, Chicago, IL 60637.

§ These results follow from the structure of the subgroups and of the discrete center of $\tilde{G}$, which may be computed as follows: $\tilde{G}$ is locally isomorphic to $SO_0(2,4)$, which has the form $SO(2)\cdot SO(4)\cdot F$, in which F is a group that is topologically flat. Hence $\tilde{G}$ may be expressed as $\tilde{SO}(2)\cdot\tilde{SO}(4)\cdot F$, and the corresponding group D' of covering transformations is generated by the kernels of the canonical homomorphisms of $\tilde{SO}(2)$ onto $SO(2)$ and $\tilde{SO}(4)$ onto $SO(4)$. In terms of the usual parametrizations of $\tilde{SO}(2)$ as R^1 and of $SO(2)$ as $[z \in C^1: |z| = 1]$, the first of these homomorphisms takes the form $t \rightarrow e^{it}$, of kernel $D'' = [2\pi n: n = 0, \pm 1, \ldots]$. In terms of the representation of $\tilde{SO}(4)$ as $\tilde{SO}(3) \times \tilde{SO}(3)$, the second of these homomorphisms is the canonical one modulo the Z_2 group generated by the element $a \times a$, in which a is the unique central element of order 2 in $\tilde{SO}(3)$.

On the other hand, the quotient $\tilde{G}/D$ of $\tilde{G}$ modulo its center D is isomorphic to the adjoint group of $\tilde{G}$, which may also be expressed as $SO_0(2,4)/\{\pm I\}$. The element $\lambda = (\pi\cdot a\cdot e)\cdot e$ of $(SO(2)\cdot SO(3)\cdot SO(3))\cdot F$ has as its square a generator of D''; and modulo D' is identical to the central element $-I$ in $SO_0(2,4)$, because the image of $a \times e$ in $O(4)$ must be a central element of order 2, and hence $-I_4$; while the image of π in $O(2)$ is $-I_2$. It follows from the latter fact that $\lambda \in D$, and because $\tilde{G}/D'$ $[\cong SO_0(2,4)]$ covers the adjoint group $[\cong SO_0(2,4)/\{\pm I\}]$ exactly twice, D consists of $D' \cup \lambda D''$.

Turning now to the subgroup of G generated by the infinitesimal Poincaré group, this is globally isomorphic to either the Poincaré group itself or to its (2-fold) universal covering group. With the same notation as above, the element of $\tilde{G}$, $(O\cdot a\cdot a)\cdot e$ is contained in the center of the $\tilde{SO}(4)$ subgroup of $\tilde{G}$. This subgroup overlaps with the Poincaré subgroup in the common sub-subgroup generated by all infinitesimal spatial rotations, infinitesimally an $O(3)$ group. As a subgroup of the $\tilde{SO}(4)$ group, this locally $O(3)$ group is imbedded in the following fashion (known, e.g., from the theory of angular momentum): on representing $\tilde{SO}(4)$ as $SU(2) \times SU(2)$, it is the diagonal, i.e. $[U \times U: U \in SU(2)]$. Now the center of $\tilde{SO}(4)$ is generated by $(-I) \times (-I)$, which is contained in the $O(3)$ subgroup, and hence in the Poincaré group, which thus occurs in its simply connected form as a subgroup of $\tilde{G}$.

Because $\tilde{M}$ is the homogeneous space of $\tilde{G}$ modulo an extended-Poincaré subgroup G_0' (the extension being by scale transformations), and $Z_2 \subset G_0'$, it follows that the Z_2 central subgroup of $\tilde{G}$ acts trivially on $\tilde{M}$. On the other hand, the effective action of the Z_∞ central subgroup is readily computable from the foregoing, and it results that $G(\tilde{M})_0 \cong \tilde{G}/Z_2$. Similarly, $G(M^*)_0 \cong \tilde{SO}_0(2,3)/Z_2$.

definition of the Z_2 subgroup in the case of G_1. The global groups are generated by these components together with an operation of space reversal, which as an automorphism of the infinitesimal causal group can be considered the same in all cases.

The conventional relativistic particle of given mass and spin has wave function in a bundle over M which may be obtained by induction to $\tilde{G}_0$ from a finite-dimensional representation of the stationary subgroup $\tilde{G}_{0,s}$, which representation is holomorphic relative to the complex structure on $\tilde{G}_{0,s}$ naturally consequent to its isomorphism with $SL(2,C)$. We are concerned here with those representations of $\tilde{G}$ that are similarly obtained by induction from finite-dimensional representations of the stationary subgroup $\tilde{G}_s$ for the canonical action of $\tilde{G}$ on $\tilde{M}$. This stationary subgroup $\tilde{G}_s$ is locally isomorphic to the Poincaré group extended by scale transformations. We consider here only those representations F of $\tilde{G}_s$ that are trivial on a translation subgroup, relative to this isomorphism, and essentially holomorphic on the homogeneous subgroup, in the sense of corresponding to the restriction of a holomorphic representation of $GL(2,C)$, in the isomorphism of the homogeneous subgroup with a subgroup of $GL(2,C)$ of codimension 1. The restriction to¶ $\tilde{G}_0$ of a representation of $\tilde{G}$ of the type just indicated yields, on further restriction to an open orbit in $\tilde{M}$ under the action of $\tilde{G}_0$, a class of relativistic bundles that includes the conventional ones in Minkowski space.

For λ a character of D and F the given finite-dimensional representation of $\tilde{G}_s$, we denote as $\mathbf{H}^\infty(F,\lambda)$ the space of all C^∞ sections f of the bundle on $\tilde{M}$ induced from the given representation F of $\tilde{G}_s$, such that $f(dp) = \lambda(d)f(p)$, for all $d \in D$ and $p \in \tilde{M}$; and we denote as $U(F,\lambda)$ the corresponding natural representation of $\tilde{G}$ in $\mathbf{H}^\infty(F,\lambda)$. The present note treats only the case $\lambda = \pi$, in which π is the character that is the parity on the Z_∞ component of D and trivial on the Z_2 component; this case is distinguished by its close connection with the representations associated with the wave, neutrino, and other relativistic wave equations for massless particles.

The concept of a *positive-energy representation* of a given Lie group G of Lie algebra $\mathbf{G}$ is relative to a given invariant closed convex cone $\mathbf{C}$ in $\mathbf{G}$. A unitary representation V of G in a Hilbert space is said to be of positive energy with respect to $\mathbf{C}$ in case the self-adjoint generator of the one-parameter unitary group $[V(e^{tX}): t \in R^1]$ is nonnegative for all $X \in \mathbf{C}$. In the case of nonunitary representations, which arise naturally in the present connection, it is convenient to employ a definition applicable to arbitrary semi-simple groups satisfying the condition:

[1] If K is the maximal essentially compact subgroup of G, there exists an element $H \in \mathbf{K}$, which, together with its transforms under the adjoint representation, has $\mathbf{C}$ as its convex closure.

Specifically, a representation of such a group in a locally convex quasi-complete topological linear space is said to be of positive energy if the only K-types that occur are such that the one-dimensional representations of the group generated by H within these K-types are periodic and positive-frequency in the usual sense. It can be shown that the two notions of positivity of the energy agree when both are applicable.

In the case of the Poincaré group, positivity of the energy will refer to the usual Lorentz-invariant cone of infinitesimal forward temporal displacements. In the case of the groups $O(n,2)$, the term will refer to the unique minimal invariant cone, with H taken as the generator of the center of K.

If V is an arbitrary continuous periodic one-parameter group of linear transformations on a locally convex linear topological space $\mathbf{L}$, its positive-energy subspace, denoted as $\mathbf{L}^+(V)$, is defined as the closed linear space of the set of all vectors $z \in \mathbf{L}$ such that $V(t)z = e^{it\lambda}z$ for all $t \in R^1$ and some $\lambda \geq 0$. If V' is a continuous representation in $\mathbf{L}$ of a group G satisfying condition 1, and such that the one-parameter group V given by the equation $V(t) = V'(e^{tH})$ is periodic, its positive-energy subspace, denoted as $\mathbf{L}^+(V')$, is defined as the common part of the $\mathbf{L}^+(V_a)$ as a varies over G, in which V_a denotes the one-parameter group: $V_a(t) = V'(a^{-1}e^{tH}a)$.

Principal results

A redefinition of the energy, as in the chronometric theory (1), leads naturally to a redefinition of the mass. In conventional relativistic theory, the mass-square operator M^2 is given essentially by the D'Alembertian. More precisely, if V is the representation of the Poincaré group G_0 corresponding to the field in question, a vector state ψ is "of mass-square M^2" if $dV(\square_0)\psi = -M^2\psi$, in which $\square_0$ is the element of the enveloping algebra of the Lie algebra of G_0 given by the equation $\square_0 = Y_0^2 - Y_1^2 - Y_2^2 - Y_3^2$, in which Y_j is the infinitesimal generator of translation in the x_j direction. On the other hand, from the standpoint of the chronometric principle, $-dV(\square_0)$ represents an apparent, rather than true physical, mass-square; for localized states it should differ negligibly from the exact physical mass-square, but may differ greatly for delocalized states; it is not invariant under *physical* (chronometric) temporal evolution, as the physical mass-square must be.

There are in fact two operators that enjoy the latter property, as well as other appropriate invariance features, differ negligibly from $-dV(\square_0)$ in localized states (by terms at most of order R^{-1}, R being the "radius of the universe" in units of the local distance scale), and closely resemble $-dV(\square_0)$ analytically. One of these operators is $-dV(\square_1)$, in which $\square_1$ denotes the second-order Casimir operator of G_1; it is invariant under G_1, as $-dV(\square_0)$ is invariant under G_0, and deforms into $-dV(\square_0)$ as $R \to \infty$; it may be called the "chronometric mass-square," and is a natural definition from a group- or particle-theoretic standpoint.

Our considerations thus far have ignored the presence in the universe of a heterogeneous ensemble of masses, and the consequent inertial mass proposed by Mach's principle. The inertial mass may in fact be given a mathematical form comparable to the chronometric via the circumstance that if ψ represents the state of the universe, there will exist, at any given point of observation, a Lorentz frame—unique except for possible (but implausible) accidental degeneracy—in which the physical energy $\langle dV(H)\psi,\psi\rangle$, in which $dV(H)$ represents the chronometric (conserved) total hamiltonian, is minimal. Corresponding to this frame there is an essentially unique decomposition of the chronometric cosmos into time and space components, $\tilde{M} \cong R^1 \times S^3$. The analogue to $\square_0$ on such a structured cosmos is $\partial^2/\partial\tau^2 - \Delta'$, in which τ is the chronometric time parameter and Δ' differs by a constant (fixed by the validity of Huyghens' principle) from the Laplace–Beltrami operator on S^3, scaled so as to agree infinitesimally with the corresponding Minkowski-space operator; this operator, say $\square_2$, may likewise be regarded as a quadratic element of the enveloping algebra of $\tilde{\mathbf{G}}$. This operator reflects the actual disposition of masses throughout the universe, is invariant under chronometric temporal evolution, and in a localized state differs by unobservably little from $\square_0$ and $\square_1$.

While all three mass operators derive primarily from sym-

¶ *A priori*, the subgroup of $\tilde{G}$ whose Lie algebra is that of $\tilde{G}_0$ (or $\tilde{G}_1$) is not necessarily simply connected, but because it is so in the present case, we use the same notation for two conceptually slightly different subgroups.

metry principles, it is interesting to note the mathematical fact that they are derivable in a natural way from Lorentzian metrics. More specifically, each of the three groups involved—G_0, G_1, and $G_2 = K$—admits invariant Lorentzian metrics on their open orbits in $\tilde{M}$, whose corresponding D'Alembertians are identical with the respective mass-square operators just indicated. These three operators may also be characterized as quadratic elements of the enveloping algebra of G that are invariant under the respective subgroups via the adjoint-representation action, by virtue of work of Cooper (6), an observation for which we are indebted to B. Kostant.

Thus there are three operators that have properties usually associated with the mass-square operator, and which differ negligibly in localized states, but are quite different from a fundamental physical position. Now the quasi-stable elementary particles that are observed have real physical masses, and so should be represented mathematically as vectors in the positive mass-square subspace, but the possibility that these subspaces are distinct, for the various mass-square operators, is *a priori* quite tenable. It is therefore a considerable conceptual simplification, as well as a source of mathematical reassurance, that in fact these subspaces are essentially the same. More specifically,

THEOREM 1. *The representation* $U(F,\pi)$ *of* $\tilde{G}$ *contains a nonempty positive-energy subspace, the restriction to which, followed by further restrictions to the subgroups* $\tilde{G}_0$, $\tilde{G}_1$, *and* $\tilde{G}_2$, *are unitarizable, modulo possible finite-dimensional invariant subspaces; and the respective nonnegative mass-square subspaces for* $\Box_0$, $\Box_1$, *and* $\Box_2$ *in these representations contain a common dense subspace.*

While this result displays the conformal invariance of the concept of the real-mass (as opposed to that "tachyon") particle, it brings into question the character of the restriction $U^+(F,\pi)$ of $U(F,\pi)$ to the positive-energy subspace. This representation has a particularly interesting structure, which has led to a model for decay processes (7), and is mathematically unprecedented.

THEOREM 2. $U^+(F,\pi)$ *is essentially nonunitarizable, but has a composition series whose infinite-dimensional factors are unitarizable. In particular it admits an uncomplemented infinite-dimensional invariant subspace, defined by a quasi-invariant*‖ *differential system that is hyperbolic with respect to the given causal structure.*

In physical terms, the invariant subspace in question is that for the massless fields of the spin type designated by F, and the defining differential equation is an appropriate variant of the wave equation. Such massless fields thus appear as discrete isolated constituents of the general real-mass subspace, and so differ from massive fields by much more than the vanishing of the mass parameter m, in contrast with the familiar situation in the case of the corresponding representations of the Poincaré group.

On the other hand, there is a closer connection between wave functions on M and on $\tilde{M}$ than might have been expected, in that all normalizable solutions of the corresponding differential equations on Minkowski space (i.e., wave, etc., equations) extend uniquely to distributions defined and satisfying the extended differential equations everywhere on $\tilde{M}$; and conversely, all sufficiently regular such distributions arise in this manner.

The massive fields are contained in the quotient "massive" representation, say V, modulo the massless subspace. The restriction $V|\tilde{G}_0$ is essentially the familiar direct integral over all real masses m of the infinitesimal $\tilde{G}_0$-invariant subspaces defined by the corresponding $\tilde{G}_0$-invariant wave equations of mass m. From the chronometric standpoint, these equations represent only a local approximation to the physical behavior of free particles of mass m; their proper description in the large entails their formulation as eigenvectors of $\Box_1$, rather than of $\Box_0$, of eigenvalue m^2. This partially heuristic physical initiative is substantiated mathematically by

THEOREM 3. *On restriction to either* $\tilde{G}_0$ *or* $\tilde{G}_1$, *the massive representation splits into a multiplicity-free direct integral (in the case of* $\tilde{G}_1$, *direct sum) of irreducibly invariant subspaces, modulo possible finite-dimensional invariant subspaces.*

Each of these subspaces is specified by the eigenvalues of the corresponding mass-square operator, and defined by an associated partial differential equation that is hyperbolic with respect to the given causal structure.

Thus the chronometric massive wave functions may be associated with particles of specific spin and mass essentially along conventional lines. On the other hand, the Casimir operators of the successive pseudo-orthogonal subgroups of $\tilde{G}_1 \cong \tilde{SO}_0(2,3)$ [i.e., those having the same Lie algebras as $O(2)$, $O(2,1)$, $O(2,2)$, and $O(2,3)$ itself] provide a system of quantum numbers, which together with the discrete center of $\tilde{G}_1$ is generically complete, and can agree exactly with conventional ones only in the limit $R \to \infty$. In consequence, *Theorem 3* establishes models for massive particles covariant with respect to $\tilde{G}_1$, satisfying the basic physical desiderata of positivity of the energy and finiteness of the propagation velocity; which models have the potential for greater precision than the conventional ones similarly associated with $\tilde{G}_0$.

The proofs of the foregoing theorems involve results and methods given in refs. 10–15, and additional considerations to be detailed elsewhere. It is possible that similar results are valid for arbitrary characters λ of D, except for the existence of the massless invariant subspace.

We thank B. Kostant for his participation in some of our discussions, particularly at an early stage in this work. This research was supported in part by the National Science Foundation under Grant MCS 76-02762.

‖ In the sense of refs. 8 and 9, as extended to systems.

1. Segal, I. E. (1976) *Mathematical Cosmology and Extragalactic Astronomy* (Academic, New York).
2. Nicoll, J. F. & Segal, I. E. (1978) *Proc. Natl. Acad. Sci. USA* **75,** 535–539.
3. Nicoll, J. F. & Segal, I. E. (1978) *Ann. Phys. (N.Y.)* in press.
4. Tits, J. L. (1960) *Colloque sur la théorie de la relativité 1959* (Centre Belge Recherches Mathématique, Brussels), pp. 107–119.
5. Vinberg, E. B. (1963) *Tr. Moscow Math. Soc.* **12,** 303–358.
6. Cooper, A. (1975) Dissertation (Massachusetts Institute of Technology, Cambridge, MA).
7. Segal, I. E. (1976) *Proc. Natl. Acad. Sci. USA* **73,** 3355–3359.
8. Kostant, B. (1975) *Non-Commutative Harmonic Analysis, 1974,* Lecture Notes in Mathematics, no. 466 (Springer, Berlin), pp. 101–128.
9. Kashiwara, M. & Vergne, M. (1977) *Non-Commutative Harmonic Analysis, 1976,* Lecture Notes in Mathematics, no. 587 (Springer, Berlin), pp. 119–137.
10. Jakobsen, H. P. (1976) Dissertation (Massachusetts Institute of Technology, Cambridge, MA).
11. Jakobsen, H. P. & Vergne, M. (1977) *J. Funct. Anal.* **24,** 52–106.
12. Ørsted, B. (1976) Dissertation (Massachusetts Institute of Technology, Cambridge, MA).
13. Ørsted, B. (1978) *J. Funct. Anal.,* in press.
14. Speh, B. E. M. (1977) Dissertation (Massachusetts Institute of Technology, Cambridge, MA).
15. Speh, B. E. M. (1978) *J. Funct. Anal.,* in press.