International Journal of Theoretical Physics, Vol. 21, Nos. 10/11, 1982

Covariant Chronogeometry and Extreme Distances. III. Macro–Micro Relations[1]

I. E. Segal

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

Received April 19, 1982

A theory is proposed that provides a coherent, concise, and nonparametric analysis of major features of the fundamental physical structure of the universe, from micro- to macroscopic. In particular it indicates that gravity is essentially the transform of the aggregate of the basic microscopic forces under conformal inversion, and not a specific force in itself. The theory also suggests a natural form for elementary particle structure that implies a nonparametric cosmological effect and indicates an intrinsic hierarchy among the microscopic forces.

The theory replaces Minkowski space M_0 by a larger space-time $\tilde{M}$ that osculates at the point of observation, and in which it is canonically imbedded by a causality-preserving transformation that is a relativistic variant of stereographic projection. The natural energy in $\tilde{M}$ exceeds that in M_0 by an amount that becomes observationally significant in large-scale contexts; an earlier proposal that this energy difference represents the extragalactic redshift has resolved a variety of anomalies and provided a greatly improved nonparametric fit to statistical data in cosmology. The energetic contents of the universe appear to distinguish a global inertial frame in which $\tilde{M}$ takes the metric form $R^1 \times S^3$ proposed by Einstein. It is argued that in the case of massive particles the energy excess in $\tilde{M}$ in this frame over that in M_0 is observed as gravity. Implications include more explicit forms of the Mach and Einstein Equivalence principles.

1. INTRODUCTION

A distinction is made between an empty or reference space-time, called a cosmos, and one additionally structured by a specific separation of space-time into time and space components, called a universe. It is argued that there exists an empirically applicable cosmos, consisting of the maximal

[1]Dedicated to Professor P. A. M. Dirac on the occasion of his 80th Birthday.

four-dimensional manifold endowed with general features of causality and uniformity (cf. Segal et al. 1981). Locally this cosmos $\tilde{M}$ is identical to the Minkowski cosmos M_0 in its causality and associated symmetry structures.

It is further argued that the energetic contents of the physical cosmos determine an approximate decomposition of the associated model $\tilde{M}$ into time and space components that is well represented by the standard presentation of the Einstein universe as $R^1 \times S^3$. It is noted that the natural time, or dually energy, in the Einstein universe is only infinitesimally identical to that in the maximally osculating Minkowski universe, determined by a corresponding decomposition of the Minkowski cosmos into time and space components as $R^1 \times R^3$. The M_0 osculating $\tilde{M}$ at a given point (interpreted as that of observation) is identical in its causal structure and corresponding symmetries (consisting of the Poincaré group extended by scale transformations) with the submanifold of $\tilde{M}$ to which it corresponds by stereographic projection (as adapted to a Lorentzian rather than Riemannian metric). Identifying this submanifold with M_0, the Minkowski energy may then be characterized as the scale covariant component of the Einstein energy, and is indicated by theory to be observed as localized, leaving a positive scale contravariant component, indicated by theory to be observed as diffuse.

It is argued that physically the Minkowski component represents the portion of the Einstein energy that is observed in laboratory particle experiments (apart from inertial effects), and that the excess of the Einstein over the Minkowski energy is observed as gravity, resulting in its action over a cosmic time scale in what is observed as mass and inertia. Among other validating considerations, the effective potentials having the same symmetry properties as these energy components are necessarily proportional to r and to $-r^{-1}$, where r denotes the euclidean distance, as phenomenologically indicated for nuclear and gravitational forces respectively.

The conclusion is reached that there is no gravitational force *per se*, and that gravity represents simply the totality of the fundamental forces exerted by matter and radiation outside the microscopic region around the point in question. The effects of these forces exerted from all parts of the universe over arbitrarily long periods are observed as action at a distance resulting from the attainment of an approximate equilibrium; and the temporal and spatial homogeneity of the forces account for the apparent uniformity of the masses and coupling constants of fundamental particles throughout the universe. In particular the concept of graviton is rendered superfluous, and Mach's Principle is given a concrete form.

Within the framework of the theory it is natural to assume that the fundamental interactions are invariant under the full 15-parameter group $\tilde{G}$ of causal symmetries of $\tilde{M}$, including transformations locally identical to scaling and to special conformal transformations. Macroscopically, invariance may be directly construed as an explicit form of the Einstein equivalence principle; microscopically, Bjorken scaling is indicated as a special case. The observed mass of a particle correspondingly appears as a summary statistic descriptive of its interaction with the universe, to an approximation adequate for microscopic theory. This is in the sense of making possible an empirically applicable few particle theory in which interactions with the rest of the universe are neglected, as e.g., in QED in which the mass of the electron is given its nonvanishing empirical value. The theoretical masses of particles conceived as isolated in the cosmos, i.e., 'bare' masses, correspondingly all vanish; the observed empirical masses may consistently be presumed to represent the results of 'clothing' by interactions with energetic states throughout the universe.

The symmetry of the theory with respect to conformal inversion is concretely exemplified by consideration of the energy constituents of such $\tilde{G}$ invariant massless particles. Taking for example the case of photons, the Einstein energy has the form $\int_{S^3}(\Sigma_{a,b}F_{ab}^{\;2})\,du$, where the F_{ab} are the components of the electromagnetic field in $R^1 \times S^3$ relative to invariant coordinates in $\tilde{M}$ and du is the element of invariant volume in S^3. Now taking the South pole S as the point of observation and imbedding M_0 into $\tilde{M}$ by relativistic stereographic projection from the North pole N, Minkowski coordinates osculating the invariant ones in $\tilde{M}$ at S are determined, and corresponding components F'_{ab} of the same field. In these terms the Minkowski energy of the photon takes the usual form $\int_{R^3}(\Sigma F'^{\,2}_{ab})\,dx$, where dx denotes the euclidean volume element in R^3. Denoting by F''_{ab} the analogous field components in the Minkowski universe osculating $\tilde{M}$ at N rather than at S, the equation

$$\int_{S^3}\left(\sum F_{ab}^{\;2}\right)du = \int_{R^3}\left(\sum F'^{\,2}_{ab}\right)dx + \int_{R^3}\left(\sum F''^{\,2}_{ab}\right)dx'$$

expresses explicitly the decomposition of the Einstein energy into scale covariant and scale contravariant components. A similar decomposition applies to arbitrary G invariant fields and can be considered an extension of the equation $\partial/\partial\theta = \partial/\partial x - \partial/\partial x'$ relating the angular coordinate θ on a circle, measured in radians, to the infinitesimally synchronous linear coordinates x and x' in the tangents to the circle at antipodal points, each tangent line being regarded as imbedded in the circle by one-dimensional stereographic projection. In the context of $\tilde{M}$, θ, x, and x' are time parameters, and the expressions given are the Hamiltonians for corresponding temporal development of the photon field.

The long standing notion that the mass of the electron is electromagnetic in origin may here be implemented by the hypothesis that the mass is substantially the result of the interaction between the electron and the all pervading cosmic background radiation (CBR). With the CBR modeled as a 3° Planck law radiation field in $\tilde{M}$, and the electron modeled by a quantized spinor field on $\tilde{M}$, a theoretical estimate of the mass as the expectation value of the interaction energy follows. It is noted that on the other hand the observed masses of particles are properly represented by the Minkowski mass, rather than the Einstein mass, just as in the case of the energy. Moreover the former can greatly exceed the latter for particles with sufficiently extended wave functions. Higher electrons may therefore have the same Einstein mass as the normal electron but represent states of an (approximately exact) Minkowski mass considerably greater than their Einstein mass. The paucity of the observed spectrum is then understandable as the exceptionality of the simultaneous attainment of approximate exact values (i.e., values attained with small dispersion) by the Minkowski and Einstein masses in one single particle state.

Although the fundamental interactions are invariant under the full causal group $\tilde{G}$, the actual energetic state of the universe need in principle to have no invariance features whatever; but in the light of empirical circumstances may be expected to be approximately invariant on a large scale under the subgroup $\tilde{K}$ leaving fixed the separation of $\tilde{M}$ into time and space components as $R^1 \times S^3$. $\tilde{K}$ is definable also as the maximal essentially compact subgroup of $\tilde{G}$, and is unique within conjugacy in $\tilde{G}$. Its specification is equivalent to the designation of a particular Einstein metric on $\tilde{M}$; these, or various equivalent specifications, define the Einstein frame of reference; this plays a role analogous to that of a Lorentz frame in M_0, but includes in addition an intrinsic unit of length, definable as the radius of the spatial component S^3 of the universe. This radius R (in laboratory units) provides a natural third fundamental constant, in addition to $\hbar$ and c, which is required for fundamental physical theory and to complete the program suggested by Minkowski (1908) of replacing limiting cases (as the Galilean group is of the Poincaré group, when $c \to \infty$, or classical physics is of quantum physics as $\hbar \to 0$) by less degenerate and mathematically more natural structures. The natural (conformally invariant, 'chronometric'), units in which $R = \hbar = c = 1$ are invariant under the causal group $\tilde{G}$, which is terminal from the Minkowski standpoint, not being representable as a limiting case of any inequivalent group.

From an *a priori* standpoint there are other subgroups of $\tilde{G}$ that are tenable as metric and mass conserving subgroups implementing spatial and temporal homogeneity. These are the causal groups of various submanifolds of $\tilde{M}$ (in keeping with the appellation 'universal' cosmos), which include in addition to Minkowski also the de Sitter and anti- de Sitter space-times, with causal groups locally isomorphic to $SO(1,4)$ and $SO(2,3)$. However, only for $\tilde{K}$ is there theoretical assurance of the existence of equilibrium and vacuum states in interacting quantized field actions, and only $\tilde{K}$ is compatible with the present physical interpretation, founded in part on extensive empirical indications for $\tilde{K}$ in extragalactic astronomy. The Einstein frame representing the inertial frame of the universe may be defined theoretically as that in which the Einstein energy is minimal, and might in principle be quite variable in time, but present large-redshift observations are indicative of a rate of variation too small to be presently observable and support the validity of $\tilde{K}$ as an approximately stationary large-scale symmetry group for the state of the universe. On the other hand, at the opposite extreme of microscopic observation, the invariant metrics for all of the cited groups appear empirically indistinguishable by direct measurement, differing only by terms of the order of the square of the space curvature; as a result of this, effective Lorentz invariance of elementary *clothed* particle interactions, as observed in the laboratory, is retained.

The question of a theoretical structure for fundamental particles are considered on the basis of general principles of symmetry, causality, and stability. It is argued that the usual assumption in conventional theory to the effect that Minkowski space-time translations have trivial internal actions (i.e., are represented by identity matrices on the particles' spin spaces) is unnatural in $\tilde{M}$, and is empirically established only as an approximation. Exemplifying the theoretical possibility that is thereby raised, a field that appears particularly felicitous in $\tilde{M}$ is proposed for the fundamental fermions. It transforms simply according to the representation of the Poincaré group defined by its natural imbedding in the linearizer (i.e., maximal linear form) $G = SU(2,2)$ of $\tilde{G}$. This becomes identical with the usual spin $\frac{1}{2}$ representation in the limit of vanishing curvature, but otherwise involves a nontrivial action of Minkowski space-time translations, of order R^{-1}, and uniquely defined in natural units; the classification of states is thereby materially different, except at an approximate level.

The action of $\tilde{G}$ on these fields is incompletely reducible except in the vanishing curvature limit. The irreducible quotients of nested invariant subspaces represent spinor fields that have features suggestive of neutrinos, leptons observed as massive, and nuclear constituents. The incomplete reducibility gives rise to nontrivial transitions between these quotient spaces, paralleling empirical decays, and providing a $\tilde{G}$ covariant model for conservation laws that hold only relative to a lower invariant subspace. The indecomposable nesting may in effect tend to shield the higher fermions in a manner suggestive of apparent internal structure. On the other hand, antiparticles are naturally taken as transforming contragradiently to particles, and their invariant subspace nesting relations would correspondingly be the reverse of those for particles. This might indicate a fundamental distinction between particles and antiparticles.

2. BACKGROUND

The present theory has been designated 'chronometric' because it originates in part in the analysis of time and temporal evolution, which analysis may be described in part as a conjunction of the position of Einstein (1921) with symmetry ideas in particle theory.

It was argued that if it is assumed only that space-time $\tilde{M}$ is four-dimensional and that it is endowed with a causal structure, in the sense of a designation of the future directions at any given point, then these concepts involve two distinct but interrelated elements. First, time as a parameter or coordinate, formulated as a real-valued function τ defined on $\tilde{M}$, which increases along future directions. Second, time as a one-parameter group T_t of transformations in $\tilde{M}$ that are evolutionary, in the sense that for $t > 0$, T_t carries every point p of $\tilde{M}$ into its future (i.e., the totality of points connected to p by paths whose direction is always towards the future), and is in addition causality preserving. These concepts are naturally taken to be connected by the equation $\tau(T_t p) = \tau(p) + t$. Such a structure, called a *covariant clock* has associated with it a corresponding energy H, which is the infinitesimal generator of the temporal evolution group T_t, as a functional on the state space of the system in question, if classical, or an Hermitian operator in its state vector space, if quantum.

It is a mathematical fact that these very general constraints on time suffice to determine the usual coordinate x_0 in M_0, together with the one-parameter group T_t: $(x_0, x) \rightarrow (x_0 + t, x)$ as the unique covariant clock on M_0, within an arbitrary causal transformation on M_0 (which consists of the product of a Poincaré with a scale transformation). But it is equally true that despite this *global* result in M_0, there exists *locally* in M_0 a quite distinct covariant clock, not at all equivalent to the conventional one even within a local causal transformation (of which there are a 15-parameter family), but with the remarkable property of osculating closely the conventional clock. Thus physical observation of purely local systems could not distinguish between these two clocks; only observations involving very long times or corresponding distances could do so. The question is in principle quite material; if conservation of energy holds relative to one of the clocks, it can not do so relative to the other.

According to Mach's principle, gravity originates at large distances, but in the absence of precise observations of masses and distances outside the solar system, only the redshift and related observations appear to fall in the category required to distinguish, potentially at least, between the two clocks. Several nontrivial mathematical circumstances facilitate the formulation and treatment of a prediction regarding the redshift and clarify its interpretation. One, the unconventional energy H is always larger than the conventional energy H_0 in any positive energy causally covariant context; and the excess energy is negligible for sufficiently localized states but can be much greater for delocalized ones. Two, Maxwell's equations provide such a context by virtue of their known conformal, and hence causal, invariance, together with the invariance of the canonical inner product between Maxwell wave functions and positivity of (either) energy. Three, H may be interpreted as the natural energy in an Einstein universe in such a way that the conventional energy H_0 is the natural energy in the Minkowski universe that maximally osculates the Einstein universe at the point of observation, and which moreover may be canonically imbedded in the Einstein universe (causally, in particular). Four, the expectation value of H_0 in a photon state ψ depends only on the local character of ψ, while that of H depends on the global structure of ψ. Five, the decomposition of H in the form $H = H_0 + H_1$ is Lorentz covariant, being uniquely characterizable as the decomposition into components that are, respectively, scale covariant and scale contravariant (scale here referring to transformation properties under the scale generator S, taking the form in M_0

$$S = \sum_j x_j(\partial/\partial x_j):\ [H_0, S] = H_0,\ [H_1, S] = -H_1)$$

These circumstances indicate that the observed energy of a photon is properly modeled by H_0 and not by H, but that in the Einstein universe the driving Hamiltonian is naturally taken as H, resulting in the theoretical prediction of an observed redshift z given by the equation $(1+z)^{-1} = \langle H_0\psi_t, \psi_t\rangle / \langle H_0\psi, \psi\rangle$, where ψ is the wave function of the photon and ψ_t is its wave function after propagation through time t in M. A similar expression is applicable if the classical energy is used. Mathematical analysis then shows that $z = \tan^2(t/2) + e$, where t is the time in radians (a causally invariant unit in $\tilde{M}$), and e depends on ψ in principle but is negligible in the frequency range that is observable.

This expression for z is naturally applicable only if the source is stationary relative to the point of observation; otherwise ψ must be replaced by its transform $T\psi$ by the action T on photon wave functions of the causal transformation representing the motion in question. Since possible motions are not directly observable in any model independent fashion, the possibility of making observable deductions from the foregoing redshift-time-of-travel law depends on the stationarity of sources relative to the point of observation, and thereby on their mutual stationarity, if this point is not to have preferential status. Making this assumption of stationarity, apart from motions (such as small random ones) whose quantitative effect is negligible, subject to *a posteriori* validation, parameter free predictions of the luminosity redshift relation follow. More specifically, an appropriate concept of stationarity requires in effect a separation of the cosmos into time and space components; at any given time, the sources are conceived as distributed in space. With the Einstein clock in $\tilde{M}$, this requires that space be represented by a three-dimensional sphere S^3; this admits a unique invariant metric under causal transformations on $\tilde{M}$ that do not affect the time, and the result is in effect a metric structure in $\tilde{M}$, identical to that of the Einstein universe.

Similar parameter free predictions follow for all of the luminosity redshift angular diameter number relations, assuming the appropriate statistical uniformity: in intrinsic luminosity, for sources in space, in the case of the luminosity redshift relation; in spatial distribution of sources, in the case of number counts; etc. Such predictions have been in very good agreement with observations on objectively defined samples that are complete (or random) within well defined limits of flux, redshift, etc., with due allowance for the observational cutoff bias inherent in astronomical observation. In contrast, the corresponding predictions of the expanding universe theory rarely fit galaxy observations well or even in a statistically acceptable manner, for the most part irrespective of the use of adjustable cosmological parameters or ancillary ad hoc hypotheses that are in practice incapable of model independent substantiation (cf. Nicoll and Segal, 1982a, 1982b and refs. therein). This is even more true of the fit to quasar or radio source observations without the introduction of entire adjustable functions, under the rubrics of luminosity and/or number evolution, none of which is appropriate or used in the chronometric theory. In addition a variety of serious anomalies within the expanding universe theory are eliminated in the chronometric theory (cf. Segal, 1980 and refs. therein).

Not only does the chronometric redshift theory fit much better than the expansion theory in the very region in which the latter originated, but statistical analysis shows that the Hubble law is self inconsistent in the low redshift region, if the samples treated are complete to anywhere near the indicated limits. While it may appear surprising in view of widespread popular belief in the Friedmann model, there is simply no known material positive indication for the expansion theory in clearly objective model independent redshift observations. The approximately Planck law form of the CBR is predicted by any temporally homogeneous energy conserving cosmology (cf. Jakobsen et al. 1979). Moreover there are no known empirical counterindications for the chronometric theory. In these circumstances it would appear clearly incorrect on the basis of normal scientific methodology to regard the expansion theory as established and/or to reject the chronometric theory. In any event, for the purposes of the present argument, the latter will be regarded as valid.

3. GRAVITY

A free photon propagated over a long time loses no Einstein energy, but shifts Minkowski energy H_0 into the 'super relativistic' energy H_1. But what is the physical interpretation of the energy H_1 *per se*? Originally (Segal, 1972) it was noted that "The effective interaction Hamiltonian H_1 responsible for the change in frequency thus has no apparent effect on local particle production, but is responsible for significant long range effects." Here it is argued that this energy H_1, which is absent in the Minkowski universe, but occurs naturally in the Einstein universe, is observed primarily as gravity, in the case of massive states, and can be regarded as such in a generalized sense in the case of massless fields.

A first step in this identification of the force represented by H_1 must be its correlation with the Newtonian gravitational potential. Since H_1 is always nonnegative and vanishes effectively only for localized states, it attains its minimum upon the localization of the macroscopic system in which it is sufficiently large as to be observable. Thus H_1 would appear to represent a generally attractive force. Now as a diffuse large scale force, the effect of H_1 on normal macroscopic matter may be expected to be approximable by a potential. In principle the matter in the universe may be expected to be more properly representable by an interacting quantum field in which H_1 should be an Hermitian operator; the potential V corresponding to H_1 then represents an expectation value for H_1 in a many particle state. Consequently the effective potential V should inherit the basic symmetry properties of H_1. Among these properties are invariance under spatial displacements and scale contravariance; the only attractive such potential in R^3 is $-g/r$, where r is the euclidean distance and g is a positive constant dependent on the matter states in question.

A second step consists in the correlation of gravity with the transform under conformal inversion of the aggregate of microscopic forces; the universality of gravity suggests that weak, electromagnetic, strong, and possibly not yet observed forces, are involved. The effects of these forces would naturally depend on the matter in question. Thus in the case of the electron, the mass would appear to be primarily electromagnetic in origin, with an admixture from the weak interaction; and a diffuse electromagnetic cosmic effect would appear to originate from the electron's interaction with the CBR. This suggests that the observed m_e may be approximable by the interaction energy of the electron with the CBR, as theoretically estimated by quantum field theoretic analysis, with the electron modeled as an incoming massless spinor particle in its bare (noninteracting) state, and the photon field as a 3° Planck law radiation field. It may therefore be computable on the basis of QED as adapted from the Minkowski to the Einstein universe, and to a finite temperature photon field by a computation that appears nontrivial but nevertheless within the bounds of realistic possibility.

In the case of nucleons, the nuclear interaction itself would appear to be the primary force whose conformal inversion constitutes the effect of gravity, the interaction with the CBR being of marginal importance. There are two specific indications for the dominance of the nuclear force in this respect. One, the observed mass of matter is proportional to the baryon number, generally within a small fraction of 1 percent, the difference being ascribable to binding energies in part of non-nuclear origin. Two, the argument just given for the phenomenological potential due to gravity applies equally to microscopic forces of many particle nuclear states, and leads to a potential proportional to the distance r. This is just what is observed, apart from a very short distance inverse distance potential whose analytic form appears less well established, and is ascribable to few particle effects not properly modeled as the expectation value of a quantized field in a many particle state.

Inertia here represents the resistance of matter to displacement from its equilibrium position relative to the energetic contents of the universe, and hence is represented quantitatively by the interaction energy of the matter with these contents. In particular, inertial and gravitational mass are inherently identical, and appear entirely consistent with the theories of Mach and Einstein. The Einstein equivalence principle, which in a modal form is describable mathematically as the local infinitesimal conformal invariance of all physical interactions, takes in an integrated form the law of invariance of fundamental fields and interactions under the global causal group. This requires all bare masses to vanish, and reinforces the principle just argued, that all mass is simply the integrated interaction energy with the remainder of the universe. This in turn confirms indications such as may be derived from the theorem of O'Raifeartaigh (1965): it appears unlikely that fundamental empirical masses may be derivable from linear symmetry considerations, or even nonlinear local ones, except insofar as they may ultimately be found to relate to the global distributions of various forms of energy in the universe.

The concept of rest mass extends naturally outside the Minkowski framework, but has a different analytic expression. The rest mass of a system is, substantially by definition, the minimal energy of the system in all physically admissible frames. Thus, if ψ denotes the normalized wave function of the system, Γ the group of admissible symmetries, and A the energy operator, the rest mass is given by the equation

$$m = \text{infimum}_{g \text{ in } \Gamma} \langle A\psi_g, \psi_g \rangle, \psi_g = R(g)\psi$$

where $R(g)$ is the operator by which g acts on the system state vector space. In conventional theory based on M_0, Γ is the Poincaré group, and the corresponding mass can be given by the analytically more explicit expression

$$m_0 = \langle (P_0^2 - P_1^2 - P_2^2 - P_3^2)\psi, \psi \rangle$$

However, for a causally invariant theory in $\tilde{M}$ an equally exact and explicit expression is not available, although the mass is no less well defined.

The energy operator A is, then, that denoted as H earlier, i.e., the infinitesimal generator of Einstein temporal evolution. Its application to the case of a bare particle, or other conceptual microscopic system, involves two constituents: $H = H^P + H^i$, where H^P is the Hamiltonian for the bare system itself, and H^i is the fundamental interaction between this system and the cosmic background (CB). In principle, a third constituent should be added: H^{CB}, the Hamiltonian of the CB, in formulating the dynamics, but in the expression for the rest mass this is naturally excluded. Modeling the bare particle by a $\tilde{G}$ invariant field (cf. below), H^P is a well defined self adjoint operator for an irreducible particle species (e.g., spin $\frac{1}{2}$ fields). QED in $\tilde{M}$ is causally invariant with the electron modeled by such a 'massless' particle, and H^i takes the form $\int_{S^3} i_l A_l \, du$, the same as in conventional theory except that integration is over S^3 instead of R^3, and the invariant basis naturally used in $\tilde{M}$ differs from that defined by the Minkowski coordinates. The effect of a causal transformation on these Hamiltonians, and corresponding effect on the derived mass under appropriate assumptions regarding the CB, are direct generalizations of the usual Einstein transformation properties of the mass under Lorentz transformations. It is possible that with the CBR modeled by a classical stochastic 3° radiation field, H^i may have expectation value approximately m_e with a dispersion less than the precision of measurement, and that the expectation value of H^P is unobservably small. Transformation by elements of $\tilde{K}$ will leave these results invariant, while for other elements of $\tilde{G}$, the cited generalization of the conventional momentum dependence of the energy will apply.

Having thus made somewhat more concrete the role of the rest mass in the present theory, it must be noted that the Einstein mass m may differ substantially from the Minkowski mass in contexts in which both apply. Experimental results are of course analyzed and reduced presently entirely in terms of the Minkowski framework; in particular, a particle whose Minkowski linear momenta j are experimentally indicated to vanish is considered at rest. This vanishing does not however imply the vanishing of the causally inverted momenta P_j', whose sum with the P_j yields the Einstein momenta $\tilde{P}_j$, except in the mathematical limit. For a particle with an extended wave function and having some dispersion in its linear momenta P_j, it is possible for the P_j to vanish within experimental limits while the P_j' are large. The relation between these operators resembles that between the Laplacian in R^3 and its transform under conformal inversion; their transformation properties under scaling show that the second may have arbitrarily large expectation values in states in which the Laplacian has any positive but arbitrarily small given limits for its expectation values. The Einstein mass, i.e., the natural one in $\tilde{M}$, is defined as an infimum over a group having 5-parameters additional to that defining the Minkowski mass, and may consequently be much smaller. There is an effect in the opposite direction due to the difference between the Einstein and Minkowski energies, but this difference appears negligible for laboratory particle states. The measurement of the causally inverted quantities is probably beyond the scope of present experimental possibilities, due in part to the interference of local background states, but the present theoretical analysis is nonparametric and so may be adaptable to predictions of the observed (i.e., Minkowski) masses of higher electrons. An earlier but different attempt to model the muon as an extended electron is due to Dirac (1962).

In the case of photons, the lack of a well defined position particularly after long distance propagation would appear to vitiate a concept of potential, but their H_1 energy could have effects comparable to those of conventional gravity. For example, the nonthermal processes in active galactic nuclei may use this energy as one source of fuel, which may thereby account in part for what has on occasion been described as an apparent anomalous source of energy in the nuclei.

4. FUNDAMENTAL FERMIONS

The assumption of covariance with respect to the full causal group $\tilde{G}$ of $\tilde{M}$ implies that the fundamental fields on $\tilde{M}$ are determined by the action on them of the 'little' group $\tilde{G}_p$ that leaves fixed any given point p of $\tilde{M}$, and indeed by the action of $\tilde{G}_p$ on the 'spin space' at the point p. For a field with a finite number of components, the latter action defines a finite dimensional representation of $\tilde{G}_p$; and conversely, any given such representation determines a well defined species of field on $\tilde{M}$. The little group $\tilde{G}_p$ is here locally isomorphic to the scale extended Poincaré group, or more exactly it is isomorphic in the large to the simply connected two-fold covering group $\tilde{P}$ of this group. It follows that physical fields of the usual type on $\tilde{M}$ are in 1-1 correspondence with the finite dimensional representations of $\tilde{P}$, similar representations (i.e., those differing only by conjugacy by a fixed matrix) being considered equivalent.

In conventional theory based on M_0, these 'inducing' representations are trivial on the translation subgroup of $\tilde{P}$, and nontrivial only on the homogeneous subgroup of $\tilde{P}$, since this latter group is the little group. In other cosmos there is no apparent physically conceptual reason to assume that all physical fields transform trivially internally under such space-time translations as may leave the inducing point p fixed. In particular, in $\tilde{M}$ the little group has 5 additional dimensions, and includes a subgroup of $\tilde{P}$ isomorphic to the space-time translation group, and there is no *a priori* physical reason why this subgroup should transform spin spaces trivially. To avoid possible confusion, it should be noted that while it makes no essential difference in the mathematical formalism how the point p is chosen, the physical interpretation requires that be space-like relative to the point of observation, from which it must be distinct; and is otherwise arbitrary. In particular, with $\tilde{M}$ formulated as $R^1 \times S^3$ and the South pole S taken as the point of observation, p is conveniently taken to be the North pole N.

Of course it seems essential for empirical purposes that any action of the space-time translation group in M_0 on spin spaces be slight, since conventional theory in which this action is trivial fits well in many respects. This would appear to raise the prospect of an additional small parameter adjusting the level of the action of this group, and thus a departure from the nonparametric unicity that has thus far characterized the present theory. However, it will be shown that there is a natural such action, in which it is the smallness of the space curvature that limits the empirical implications of the nontrivial action, and this takes a unique form in natural units.

The action in question is obtained by a modification of the spin action, where from the term 'spannor' (for spinor + spanner, = 'wrenched' spinor) for the inducing representation derives. If one takes the spin representation of $\tilde{P}$, defined as trivial on the translation subgroup and as usual on homogeneous transformations, and induces from this to the corresponding field on $\tilde{M}$, one obtains fields conventionally regarded as spinor fields, associated with an appropriate form of the Dirac equation, etc. These spinor fields on $\tilde{M}$ are found to be closely related to those in M_0; so closely in fact that the latter may be regarded as restrictions of the former to M_0 as imbedded in $\tilde{M}$, in fundamental respects. There are some significant differences, such as the triviality of the action of a new global invariant symmetry η on M_0, but not on $\tilde{M}$; but this action could be implanted on the spinors in M_0 by transference from the action on spinors in $\tilde{M}$. Whether on $\tilde{M}$ or on M_0, such spinor particles may quite appropriately be characterized simply as spin $\frac{1}{2}$ particles in the usual empirical sense (at least locally, i.e., apart from effects relating to discrete symmetries or to cosmology). The important implicit assumption in the reduction of observational data that is made conventionally is that all elementary particles have a definite spin, that this reflects their transformation properties under the homogeneous Lorentz group, determining the structure of their spin spaces; and the replacement of M_0 by $\tilde{M}$ does not basically alter this assumption, whether for particles of spin $\frac{1}{2}$ or higher spins.

This is however not the case for spannor fields, which appear as ordinary spinor fields in the limit of vanishing space curvature, and transform in the same way under the subgroup $\tilde{K}$, but which in their exact form are quite distinct. The existence of these fields illustrates the fact that the experimental determination of the spin of a particle is model dependent, being based implicitly on the Minkowski framework, and that the actual spin space of the particle could reasonably and consistently with experiment be significantly more complicated.

To define the spannors, recall that $\tilde{P}$ can be regarded not only as the essential causal group of M_0, but as a subgroup of the essential causal group $\tilde{G}$ of $\tilde{M}$, M_0 being identifiable with an orbit in $\tilde{M}$ under the action of $\tilde{P}$. On the other hand $\tilde{G}$ is locally identical to its linear form G, which is representable as the group $SU(2,2)$ of 4×4 matrices. This defines a local representation of $\tilde{P}$ that extends in fact to a global representation by such matrices. The vectors in the four-dimensional representation space of this faithful representation of $\tilde{P}$ are defined as the spannors, apart from their conformal weight, which will presently be left arbitrary. Cospannors are defined as vectors in the dual space, and so transform contragrediently.

To see that in the limit of vanishing curvature the spannor representation has a limit that is trivial on the translation subgroup of $\tilde{P}$, and in fact just the usual spin representation, applied to the homogeneous Lorentz subgroup of $\tilde{P}$, an explicit presentation of the representation is required. It is convenient to work with a local presentation of $\tilde{M}$ as $U(2)$, on which G acts by carrying the given matrix Z in $U(2)$ into $(AZ+B)(CZ+D)^{-1}$, where A, B, C, and D are 2×2 complex matrices such that $\begin{pmatrix} A & B \\ C & D \end{pmatrix}$ is in $SU(2,2)$, and defines the group element g in G in question. $\tilde{P}$ may then be represented as the subgroup of G leaving fixed the matrix -1 in $U(2)$, from which point the spannor fields on $U(2)$ are induced. To parametrize $\tilde{P}$, M_0 is identified as usual with the space of 2×2 skew Hermitian matrices $h(2)$, in terms of which any transformation T in $\tilde{P}$ is uniquely expressible as the product (from right to left) of the following constituent transformations: (i) L in $SL(2, C)$, sending $h \to LhL^*$; (ii) the scale transformation $h \to e^t h$, t being real; (c) the space-time translation $h \to h + 4f$, where f is arbitrary in $h(2)$. The corresponding element $U(T)$ of $SU(2,2)$ then takes the form $U(T) = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$, where

$$A = [\cosh(t/2)(1+f) - \sinh(t/2)f] \frac{L^{*-1} + L}{2}$$

$$+ [-\sinh(t/2)(1+f) + \cosh(t/2)f] \frac{L^{*-1} - L}{2},$$

$$B = [-\sinh(t/2)(1+f) + \cosh(t/2)f] \frac{L^{*-1} + L}{2}$$

$$+ [\cosh(t/2)(1+f) - \sin(t/2)f] \frac{L^{*-1} - L}{2},$$

$$C = [-\cosh(t/2)f - \sinh(t/2)(1-f)] \frac{L^{*-1} + L}{2}$$

$$+ [\sinh(t/2)f + \cosh(t/2)(1-f)] \frac{L^{*-1} - L}{2},$$

$$D = [\sinh(t/2)f + \cosh(t/2)(1-f)] \frac{L^{*-1} + L}{2}$$

$$+ [-\cosh(t/2)f - \sinh(t/2)(1-f)] \frac{L^{*-1} - L}{2}.$$

Now setting R for the radius of space S^3 in laboratory units (keeping $\hbar = c = 1$), f is replaced by f/R, while t and L are unchanged. It follows that as $R \to \infty$, $U(T) \to U_0(T)$, where $U_0(T) = \begin{pmatrix} A_0 & B_0 \\ C_0 & D_0 \end{pmatrix}$ and

$$A_0 = \frac{1}{2}(e^{-t/2} L^{*-1} + e^{t/2} L) \qquad B_0 = \frac{1}{2}(-e^{-t/2} L^{*-1} + e^{t/2} L)$$

$$C_0 = \frac{1}{2}(-e^{-t/2} L^{*-1} + e^{t/2} L) \qquad D_0 = \frac{1}{2}(e^{-t/2} L^{*-1} + e^{t/2} L)$$

Introducing the matrix $\Gamma = 2^{-1/2}\begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$, $U_0(T)$ takes the form $\Gamma\begin{pmatrix} e^{-1/2tL^{*-1}} & 0 \\ 0 & e^{1/2tL} \end{pmatrix}\Gamma^{-1}$, showing that on the homogeneous Lorentz group U_0 is equivalent to the usual spin representation.

The spannor and cospannor fields on M thus transform under the full causal group $\tilde{G}$, and form a nontrivial variant of the conventional spinor fields, which can be regarded as the vanishing curvature limit of the former fields (cf. Dirac 1936, Segal 1959a, and Budini 1979). The spannor fields admit an invariant subspace transforming under $\tilde{G}$ equivalently to two-component spinor fields, but this admits no $\tilde{G}$ invariant complementary subspace. On restriction either to $\tilde{K}$ or to $\tilde{P}$, the invariant subspace is invariantly complemented, but the respective complements are irremediably distinct. Thus quantum numbers that are built up from the generators of $\tilde{G}$ may appear Hermitian in the invariant subspace $\mathbf{S}_0$, and in the quotient of the entire spannor space $\mathbf{S}$ *modulo* the subspace, without being in fact Hermitian on the entire space $\mathbf{S}$. The same phenomenon is present in the spinor fields themselves (Segal et al., 1981); in spannors there is a further level to the hierarchy of invariant subspaces, which is suggestive of basic nuclear particles, as the spinor subspace is of leptons.

The hierarchy of subspaces must be constrained not only by covariance considerations, but also by stability and causality. An arbitrary $\tilde{G}$ covariant field on $\tilde{M}$ need not, and typically will not, admit a positive energy $\tilde{G}$ invariant subfield. The particles represented by such a field would not appear stable, and the field would appear inappropriate for fundamental physical purposes. In the case of spannor fields, the important constraint of the existence of a nontrivial positive energy subspace properly inclusive of a putative lepton subspace has been established by S. M. Paneitz.

Causality and phenomenological applicability would appear to require the development of wave equations for the propagation of spannor fields of definite mass, consistently with the given causal structure in $\tilde{M}$. This means the equations should be hyperbolic, with characteristic cones the same as those defining the causal structure in $\tilde{M}$. Any given $\tilde{G}$ invariant category of physical field should be decomposable into a direct sum of $\tilde{K}$ invariant subspaces, each having such a propagation and mass feature. There is no need to postulate such a wave equation, since it derives automatically from the structure of the fields. It can be defined, apart from a constant mass term, by the differential operator representing the action of the $\tilde{K}$-invariant element of the enveloping algebra of the Lie algebra of $\tilde{K}$, $L_{-1,0}{}^2 - \sum_{i,j=1}^{4} L_{i,j}{}^2$. Positivity of the energy insures reality of the relevant mass.

The structure of the stable (or, having one-sided energy spectrum) invariant subspaces of the spannor fields depends in part on the conformal weight, but in all events is incompletely reducible. There is a chain of subspaces, each nested indecomposably in the higher one. The irreducible quotient spaces, or factors, are similar to spinor fields, and in fact appear representative of the leptons observed as massive, of neutrinos, and of nuclear constituents. For the present, it will suffice to designate spannor particles simply as 'x-ons'; due to the incomplete reducibility, there will be transitions between these, as a consequence of the spannor field structure itself. Thus these fields, as a causally invariant ensemble, have more structure than any representation as a direct sum of irreducibly invariant ('elementary') particle species.

These features parallel empirically observed ones that are not easily or economically explained by conventional theory, such as conservation of strangeness only modulo the weak interactions. The indecomposability of the spannor fields gives rise to apparent particle production, consistently with causality and covariance constraints, that can not clearly be modelled in conventional Lagrangian terms, except in the classical limit. A full discussion of this aspect must involve the treatment of the quantization of spannor fields, the physical specification of the conformal weight and assignments to empirically observed particles, etc., which must be deferred to another occasion. Here it will merely be mentioned that spannor forces would naturally be mediated by boson fields transforming according to the positive-energy subfields of the spannor x co-spannor field, say y-ons. The y-ons would transform indecomposably also, and split up on formation of successive quotients of nested invariant subspaces into 'free bosons', and should include photons and particles presently modelled as massive and of spin 1. Higher tensor products analogous to and deforming into higher spin particles from a conventional point of view may also intervene physically. There is moreover the possibility of deriving gauge groups along lines parallel to those for QED, connecting degeneracy of invariant inner products with gauge actions, and perhaps providing a more fundamental treatment of isospin and other internal quantum numbers.

The inclusion relations of the invariant subspace of the stable anti-spannor fields are the reverse of those for the spannors. From a group-theoretic standpoint, the photon and neutrino fields are distinguished by a kind of exceptional smallness. It seems possible that they are properly identifiable with lower subspaces of one of the fields, and with upper subspaces of the co-field; and this could affect their observability, the particles in higher subspaces being perhaps screened by those in lower spaces. If so, this might be explicative of an only apparent preponderance of particle over anti-particle abundance; there is an objective mathematical distinction suggesting that the latter may be less easily observed in an identifiable state.

5. DISCUSSION

Quantitative empirical validation of these ideas would appear most cogently accessible through a treatment of *lepton* masses. Obstacles are considerable analytic complexity and field theoretic divergences, although a one-particle approximation may be effective in the case of higher electrons. Another possible route for further validation is the derivation of strangeness and isospin, or effective forms thereof, from the structure of the spannor fields.

From a cosmological standpoint the interesting question of the dynamics of the inertial frame of the universe is unresolved. However, this frame is only approximate, and the issue is thus not a theoretically foundational one. Many more large redshift observations are needed before there is any possibility to detect a porabilistically significant motion of the frame via its effect on the redshift luminosity number relations of quasars.

If the ideas presented here are basically sound, it would mean that the cosmos, the fundamental fields in it, and their interactions have a conceptually most natural and simple structure. The complexity of the cosmos would reside in its *state* E. It would appear unsound to postulate its description by a vector in any well defined cosmological Hilbert space; being presumably totally devoid of exact symmetries, there is not even the beginnings of an effective precise labeling of the state. This state E may at best be representable as a state in the mathematical sense of Clifford-Weyl C^{*-} algebra (cf. Segal, 1959) built from x-on and y-on fields and antifields.

Rigorously, the state E must be presumed indivisible, 'free' objects (photons, etc.) having only approximate validity. However on a sufficiently coarse scale, E appears empirically to be approximately $\tilde{K}$ invariant for a well defined maximal essentially compact subgroup $\tilde{K}$ of $\tilde{G}$, and to be effectively parametrizable in part. But to the approximate extent that subsystems may be regarded as isolated and possessing individual wave functions relative to such a background state, their wave functions are nonvanishing throughout the cosmos, as a consequence of stability considerations. The localization and individuality of systems that are postulated in conventional analysis appear to be entirely approximate and anthropocentric.

ACKNOWLEDGMENTS

I thank H. D. Doebner, J. F. Nicoll, and S. M. Paneitz for discussions, and the latter for permission to refer to mathematical results derived by him or in our joint work. Mathematical work related to this presentation was supported in part by the NSF. Thanks are also due the Humboldt Foundation for facilitating aspects of the present study.

REFERENCES

P. Budini (1979). *Czechoslovak Journal of Physics*, **B29**, 6.

P. A. M. Dirac (1936). *Annals of Mathematics*, **37**, 429.

P. A. M. Dirac (1962). *Proceedings of the Royal Society of London*, **268A**, 57.

A. Einstein (1921). *Sitzungsber. Preuss. Akad. Wiss.*, 261.

H. P. Jakobsen, M. Kon, and I. E. Segal (1979). *Physical Review Letters*, **42**, 1788.

H. Minkowski (1908). Address at Assembly of German Natural Scientists.

J. F. Nicoll and I. E. Segal (1982). (a) *Astrophysical Journal*, in press. (b) *Proceedings of the National Academy of Science*, in press.

L. O'Raifeartaigh (1965). *Physical Review Letters*, **14**, 575.

S. M. Paneitz and I. E. Segal (1982). *Journal of Functional Analysis*, in press.

I. E. Segal (1972). *Astron. & Astrophys.*, **18**, 143.

I. E. Segal, H. P. Jakobsen, B. Ørsted, S. M. Paneitz, and B. Speh (1981). *Proceedings of the National Academy of Sciences*, **78**, 5261.

I. E. Segal (1980). Ap. J. **238**, 38.

I. E. Segal (1959a). In *Les problemes mathematiques de la theorie quantique des champs*, publ. C. N. R. S. Paris, p. 57.

I. E. Segal (1959b). *Mat.-fys. Medd. K. Danske Videnskabernes Selskab* 31, no. 12, 1.