International Journal of Theoretical Physics, Vol. 25, No. 5, 1986

Quantum Theoretical Origin of Spacetime Structure

A. R. Marlow[1]

Received October 1, 1985

The possibility that macroscopic spacetime is defined by the observable structure of the simplest nontrivial quantum systems is investigated.

The basic claim of this paper is that modern physics has already uncovered a theoretical explanation for the observed structure of spacetime: such structure is a necessary consequence of the mathematical structure of quantum theory applied to the simplest nonstationary systems the theory allows.

Before going into details, a hypothetical example from the early part of this century might be helpful in clarifying the type of argument to be used. Suppose that after Einstein's discovery that an indefinite metric was necessary in physical theory, it had further been found that the mathematical structure of such metrics required a minimum of four dimensions with a 3-1 splitting. Surely such a circumstance (contrafactual, of course) would have been taken as an explanation of the basic geometric structure of spacetime—an example of Wigner's "unreasonable effectiveness of mathematics." As in all physical explanations, the basic question would merely have been transferred to a deeper level (why indefinite metric?), but with this proviso we could legitimately claim to have an explanation for the dimensionality and signature of spacetime. Now while, in fact, indefinite metrics do not impose such dimensional and signature restrictions, it is the aim of this paper to show that the mathematical structure of quantum theory—complex projective space (ray space)—does impose precisely such restrictions on the observable properties of the simplest systems the theory allows.

[1]Department of Physics, Loyola University, New Orleans, Louisiana 70118.

0020-7748/86/0500-0561$05.00/0

To see this most clearly, I will work in the projective space formulation of quantum theory: pure states of physical systems correspond to one-dimensional subspaces of a complex Hilbert space $\mathscr{H}$, and so can be represented as complex one-dimensional projection operators on $\mathscr{H}$. The standard density operator representation for arbitrary states (pure and statistically mixed) is then formed by taking convex linear combinations of such projections, so that the full set of states for a physical system is the real manifold of positive trace 1 operators W on $\mathscr{H}$. Observables correspond to self-adjoint operators A on $\mathscr{H}$, and the results of experimental measurements of A on a state W are predicted as an average value

$$\langle A \rangle_W = \text{trace}(AW)$$

with uncertainty (standard deviation)

$$\Delta A_W = [\text{trace}(A^2 W) - \text{trace}^2(AW)]^{1/2}$$

Finally, elementary dimensional computations easily show that, if N is the complex dimension of $\mathscr{H}$, the dimension of the real vector space $\mathscr{A}$ of self-adjoint operators (observables) is N^2, while the dimension of the real nonlinear submanifold $\mathscr{S}$ of positive trace 1 operators (states) is $N^2 - 1$.

From these simple considerations alone a striking difference between classical and quantum theory appears: the mathematical structure of classical theory imposes no dimensional restrictions on the systems that can be treated by the theory, while quantum theory specifies a discrete sequence of dimensions for spaces of system observables and submanifolds of system states. It might be objected that, since standard treatments embed everything in an infinite-dimensional Hilbert space and consider all observables and density operators as acting on that space, all such finite-dimensional distinctions are irrelevant, but this objection is misleading. I have presented arguments elsewhere (Marlow, 1984) against the automatic assumption of infinite-dimensional space as the basic representation space for quantum theory, but even granting the representation of all states and observables as operators on such a space, the objection fails. If the evolution of a given system is such that its states W are confined to a finite-dimensional submanifold of the infinite-dimensional manifold of positive trace 1 operators, then there exists a finite-dimensional projection operator P with the property $PW = WP = W$ for all states W of the system, and all infinite-dimensional observables A' "compress" to a finite-dimensional subspace of observables of the form $A = PA'P$ insofar as any measurable predictions for the given system are concerned. This follows, of course, from $P^2 = P$ (idempotence of projections) and the properties of the trace function: $\text{trace}(A'W) = \text{trace}(A'(PWP) = \text{trace}(PA'PW) = \text{trace}(AW)$. Thus we are back to considering the finite-dimensional spaces of observables on finite-dimensional manifolds of states allowed by the mathematical structure of quantum theory.

Purely mathematically, the simplest case permitted by quantum structure corresponds to complex dimension 1 ($\mathscr{H} = C$ itself), with one-dimensional real space of observables isomorphic to R, but zero-dimensional state manifold consisting of the single pure state operator 1, reflecting the fact that complex one-dimensional space serves to define a zero-dimensional complex projective space consisting of a single ray. Whatever speculative role such a single-state system might play in theory, it should have no part in observational physics, since measurements depend crucially on the results of changes of state. Hence, I refer to this case as the trivial or one-point system.

The simplest nontrivial physical system permitted by the mathematical structure of quantum theory, then, is defined by complex dimension $N = 2$ ($\mathscr{H} = C^2$), with a real four-dimensional space of observables A and a real three-dimensional submanifold of states W. To see how this is isomorphic to the standard Dirac 4-spinor formalism for spin-$\frac{1}{2}$ systems, first note that the standard treatments work with a complex four-dimensional column ψ and row ψ^+ representation of states, so that $W = \psi\psi^+$ and $\psi^+\psi$ defines the Hermitian inner product on this representation of states. As always, such a decomposition of the state representation introduces an arbitrary phase factor; in terms of the unique positive square roots of our positive state operators W, we may define

$$\psi = W^{1/2} e^{JS}$$

where S is a scalar and J is any operator such that $J^2 = -I$, $J^+ = -J$ (a complex structure operator, which, it should be emphasized, need not be identical with the complex structure i used to define the complex field C of our underlying Hilbert space). Thus,

$$W = \psi\psi^+$$

and the expression for expectation values becomes

$$\mathrm{trace}(AW) = \mathrm{trace}(A\psi\psi^+) = \mathrm{trace}(\psi^+ A\psi)$$

To get the standard complex column 4-spinor representation of ψ, we need only note that the unit imaginary j, defined on states by $jW = WJ$, commutes with any left multiplication by operators, i.e.,

$$AjW = jAW$$

so that, in terms of the complex numbers defined by j, any ψ can be expressed as a column of complex coefficients relative to a basis for the self-adjoint operators A, say I and the three standard Pauli σ operators (defined using the original unit imaginary i). In terms of this same basis, of course, the linear operator on ψ defined by left multiplication with a general self-adjoint operator A is represented as a 4×4 matrix, and allowing the abuse of notation of writing the same symbols for the operators A, ψ, and their matrix representations, the expectation value formula becomes

$$\psi^{+}A\psi = \text{trace}(\psi^{+}A\psi) = \text{trace}(AW)$$

where ψ ranges through a four-parameter submanifold of j-complex 4-spinors.

To complete the demonstration that the simplest nontrivial systems allowed by the mathematical structure of quantum theory are isomorphic to the standard Dirac 4-spinor systems, we need to investigate the symmetries of the expectation value formula to show that they give the observed relativistic and gauge symmetries of standard theory. To preserve the self-adjointness and complex i-linearity of all observables A and states W, as well as all expectation values, we have

$$\text{trace}(AW) = \text{trace}(S^{+-1}AS^{-1}SWS^{+})$$

where S ranges through the group $GL(2, C)\times\overline{GL(2, C)}$ of invertible i-linear–antilinear transformations on C^2. To express this group clearly as well as for later study of the differentiable structure of the manifold of states (and to distinguish the fundamental complex structure i from the complexification j, which is an artifact of the physicist's practice of using homogenous coordinatizations for complex projective structures), it is necessary to make explicit the underlying real structure of the complex Hilbert space C^2:

$$C^2 = \{R^4, i\}$$

i.e., complex 2-space is real 4-space equipped with a complex structure operator i, expressed relative to the usual coordinates as

$$i = \begin{pmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

Then our symmetry group consists of two disconnected components in the group $GL(4, R)$ of all real, invertible 4×4 matrices: the subgroup $GL(2, C)$ of C-linear elements S, commuting with i, and $\overline{GL}(2, C)$, consisting of all elements S' that anticommute with i (anti-C-linear elements). This second component is not a subgroup of $GL(4, R)$, since products of two elements are i-linear, but has the simple structure

$$\overline{GL(2, C)} = \gamma_4 GL(2, C)$$

that is, every element can be expressed as $S' = \gamma_4 S$, with S in $GL(2, C)$ and γ_4 some fixed element anticommuting with i, conveniently chosen as

$$\gamma_4 = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{pmatrix}$$

Defining $\gamma^4 = \gamma_4^+ = \gamma_4^{-1}$, it is conventional in standard treatments of Dirac 4-spinors to incorporate this operator into the representations of states and observables by the prescriptions

$$\bar{\psi} = \psi^+ \gamma^4, \qquad \gamma^A = \gamma_4 A$$

so that the expectation value formula becomes

$$\langle A \rangle_W = \langle A \rangle_\psi = \bar{\psi} \gamma^A \psi = \psi^+ A \psi = \text{trace}(AW)$$

(It has become even more conventional to incorporate the complexification j by defining the operator γ_0 on 4-spinors ψ as $\gamma_0 \psi = j\gamma_4 \psi = \gamma_4 \psi J$; this makes γ_0 a j-self-adjoint operator on 4-spinors, but for the moment I will forego this extension as an unnecessary intrusion of the complexification j. It amounts to the choice of relativistic spacetime signature $+---$ over signature $-+++$.)

Now defining C_0 as the complex group based on i (all nonzero complex numbers $e^{\rho + i\varepsilon}$), one can further decompose $G = GL(2, C) \times \overline{GL(2, C)}$ into the product

$$GL(2, C) \times \gamma_4 GL(2, C) = C_0 \times SL(2, C) \times \gamma_4 (C_0 \times SL(2, C))$$

Temporarily setting aside the group C_0 of complex scalars, the fact that the nonscalar symmetry group of the simplest quantum systems is $SL(2, C)$ should strike one as highly significant, given the important role of that group in modern physics: applied to the real 4-space of observables A, the elements S of $SL(2, C)$ define exactly the restricted Lorentz group of physical boosts and rotations, using the definition

$$L_S A = S^+ A S$$

[Since S and $-S$ both define the same Lorentz transformation, one arrives of course at the usual description of $SL(2, C)$ as the twofold covering group of the restricted Lorentz group.]

Lacking any other known physical interpretation of $SL(2, C)$, we are compelled to adopt the standard interpretation of this group as it applies to the simplest systems permitted by the structure of quantum theory. Thus, of the six standardly chosen independent generators of $SL(2, C)$, the three that are i-self-adjoint and anticommute with γ_4,

$$\sigma_1 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}$$

$$\sigma_2 = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \\ 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}$$

$$\sigma_3 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$

[the standard Pauli matrices, expressed here in full real 4×4 form as elements of $GL(4, R)$ instead of the usual more compressed complex 2×2 notation], as observables must be interpreted as representing the three usual independent directions of observed physical space along which systems and observers may be boosted, while the three anti-self-adjoint generators, $i\sigma_1 = \sigma_2\sigma_3$, $i\sigma_2 = \sigma_3\sigma_1$, and $i\sigma_3 = \sigma_1\sigma_2$, commuting with γ_4 and generating i-unitary transformations, obviously must correspond to the orthogonal planes of rotation allowed by observed physical space and axially define the same three observable directions as the boosts.

Besides their role as generators of Lorentz transformations on the simplest quantum systems, the operators σ_k defined above serve as three elements of a complete basis for the real 4-space of observables on such systems. The fourth basis element,

$$\sigma_0 = I = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

obviously has distinguished characteristics: it commutes with every operator, and, under $SL(2, C)$-induced Lorentz symmetries, it transforms as the usual time-energy axis. To see this, first note that the elements R defining spatial rotations in $SL(2, C)$ are i-unitary, with $R^+ = R^{-1}$, while the boosts B are positive operators with $B^+ = B$. Thus the observable axis σ_0 is invariant under all spatial rotations

$$R\sigma_0 R^+ = RR^{-1}\sigma_0 = \sigma_0$$

while transforming under boosts to

$$B\sigma_0 B^+ = BB^+\sigma_0 = B^2\sigma_0$$

Writing boosts in $SL(2, C)$ in the standard velocity parameter representation,

$$B = \exp[(\upsilon/2)\sigma_{\theta\varphi}]$$

where

$$\sigma_{\theta\varphi} = (\sin\varphi\cos\theta)\sigma_1 + (\sin\varphi\sin\theta)\sigma_2 + (\cos\varphi)\sigma_3$$

we have

$$B^2 = \exp(\upsilon\sigma_{\theta\varphi})$$

so that

$$B\sigma_0 B^+ = B^2\sigma_0 = (\cosh\upsilon)\sigma_0 + (\sinh\upsilon)\sigma_{\theta\varphi}$$

as required of the time-energy axis in observable 4-space, where $u = \tanh\upsilon$ defines the standard relativistic speed change involved in the boost.

In the usual treatment of Dirac 4-spinor systems, incorporating the i-antilinear element γ_4 (or $j\gamma_4 = \gamma_0$), since for arbitrary $SL(2, C)$ elements S one has

$$\gamma_4 S^+ = S^{-1}\gamma_4$$

the observables

$$\gamma^A = \gamma_4 A = A^\mu \gamma_4 \sigma_\mu = A^\mu \gamma_\mu$$

obviously transform as

$$\gamma^{A'} = \gamma_4 S^+ A S = S^{-1}\gamma_4 A S = S^{-1}\gamma^A S$$

while the conjugate spinors $\bar{\psi}$ transform as

$$\bar{\psi}' = \psi^+ S^+ \gamma^4 = \psi\gamma^4 S^{-1} = \bar{\psi} S^{-1}$$

Thus, the role of the distinguished time-energy observable axis $\sigma_0 = I$ is taken over by $\gamma_4\sigma_0 = \gamma_4$, but the physical content is identical in either representation, and the observables on the simplest nontrivial quantum systems transform as relativistic 4-vectors under the Lorentz transformations defined by $SL(2, C)$.

Since the simplest nontrivial quantum systems intrinsically exhibit the observational structure standardly associated with relativistic spacetime, assuming that our environment is populated with such systems and that the only meaningful operational definition of a point or event in spacetime is in terms of observations made on a quantum system "occupying" the point, there is no mystery about the origin of the observed structure of spacetime: each elementary system defines a local Minkowski 4-space observable structure, and collections of such systems, treated according to standard many-particle methods, should exhibit the macroscopic structure of a manifold obtained by gluing together local Minkowski structures, i.e., an Einsteinian pseudo-Riemannian manifold. The precise mathematical treatment of such ensembles of elementary quantum systems is at present being formulated by the author, but the general idea should be clear: local Minkowski structure is well explained by systems that individually exhibit such structure in their space of observables.

To get some tentative indication as to how such a quantum theory of spacetime might be elaborated, first assume the most general form of system representation that implements the present assumption that the macroscopic world is built up from observations made on ensembles of the simplest nontrivial quantum systems: require that the density operator representing such an ensemble be of the form of a convex linear combination

$$W=\sum_n p_n W_n$$

where p_n represents the classical probability of detecting the nth system and each operator W_n has the form

$$W_n = q_n|\psi_n\rangle\langle\psi_n| + (1-q_n)(I_n - |\psi_n\rangle\langle\psi_n|)$$

with I_n a complex two-dimensional (real four-dimensional) projection, $|\psi_n\rangle$ a complex unit vector varying through the range of I_n, and q_n a number specifying the classical probability of detecting the state corresponding to $|\psi_n\rangle$. Thus, in terms of the standard polarization vector for the pure state $|\psi_n\rangle$

$$\sigma_n = 2|\psi_n\rangle\langle\psi_n| - I_n$$

with $|\psi_n\rangle\langle\psi_n| = \frac{1}{2}(I_n+\sigma_n)$, we have

$$\begin{aligned} W_n &= q_n/2(I_n+\sigma_n)+(1-q_n)/2(I_n-\sigma_n) \\ &= \tfrac{1}{2}(I_n + r_n\sigma_n) \end{aligned}$$

with $r_n = 2q_n - 1$. Obviously, as $|\psi_n\rangle$ runs through all unit vectors in the range of I_n, the operator σ_n varies over the unit 2-sphere of traceless self-adjoint operators with square I_n. This is, of course, just a simple statement of the proof that complex one-dimensional projective space is isomorphic to the real 2-sphere. Hence, in terms of copies σ_{nk} of the standard Pauli operators, defined now separately on the real four-dimensional range R_n^4 of each projection $I_n = \sigma_{n0}$, it is possible to use the polar angle representation specified earlier to write the possible states of the subsystems of the ensemble as

$$W_n(r_n, \theta_n, \varphi_n) = \tfrac{1}{2}(\sigma_{n0} + r_n\sigma_n(\theta_n, \varphi_n))$$

where, since q_n ranges from 0 to 1, r_n ranges from -1 to 1. Thus, to determine the ensemble W completely it is necessarily to specify the four parameters p_n, r_n, θ_n, φ_n for each subsystem W_n, where the last three variables may be chosen independently for each n, while the p_n are tied together by the requirement that $\Sigma\, p_n = 1$. Such a system can be described as a quantum gas made up of one-dimensional complex projective space systems.

Placing no restrictions on the observables available to an observer other than the requirement of general quantum theory that they be self-adjoint operators A on some Hilbert space containing the observed system, one obtains from the expectation value formula

$$\bar{A}_W = \sum_m p_n \operatorname{trace}(AW_n) = \sum_n p_n \operatorname{trace}(I_n A I_n W_n)$$

where the compressed operator $A_n = I_n A I_n = \sigma_{n0} A \sigma_{n0}$ is an observable in the 4-space defined by the nth subsystem:

$$A_n = A_n^\mu \sigma_{n\mu} \qquad \text{(summation over } \mu = 0, 1, 2, 3)$$

Hence, as far as an observer can determine from any information allowed by quantum theory, there is no loss of generality in assuming pairwise orthogonality for the system projections I_n, i.e.,

$$I_n I_m = \delta_{nm} I_n$$

and an orthogonal direct sum structure for all observables:

$$A = \sum_n A_n$$

since there would be no observational way of distinguishing the simple structure postulated here from any other that might be assumed. (More specifically, observable structures and possible relations among system projections I_n split into equivalence classes of observationally indistinguishable alternatives, from among which we have made a particularly simple choice. If further restrictions on observers or interactions between subsystems were assumed, some other structural choice might be more appropriate, but until such assumptions are made explicit, there is no reason to change the choice made above.)

Thus, general observables reduce equivalently to direct sums of 4-vector observables A_n on subsystems, and since, by hypothesis, the parameters p_n in the general density operator W represent classical probabilities (in principle arbitrarily refinable), a sophisticated observer with extensive experience of the system could reduce all but one of these probabilities essentially to zero, and so study the 4-vector structure of the individual subsystems. Such an observer could go on to uncover the privileged status of the "time-energy" observable axis I_n for each subsystem, with its constant expectation value, and learn to fix one such axis as origin for difference "coordinate" observables $A_m - I_n$ describing other subsystems. The observer would, of course, find that all such coordinate vectors could vary only along four independent directions, and thus the picture of a four-dimensional background space in which all systems were imbeded would emerge. The investigation of precisely how the full relativistic worldview could arise is still in progress, and so I end my sketch of observables on a simple quantum gas here, but enough has been said to emphasize the main point: whatever structure is assigned to observers, the mathematics of quantum theory imposes strong system-dependent restrictions on the possible results of observations, so that assuming our world is largely populated with the simplest quantum systems inevitably leads to the 4-space structure uncovered by actual experience. (For comparison with standard relativistic treatments, of course, γ_{4n} operators must also be incorporated, as detailed earlier.)

Just as with $SL(2, C)$, the remaining complex scalar component C_0 of the symmetry group $GL(2, C)$ can be given the standard interpretation. The compact part $U(1) = \{e^{i\varepsilon}\}$, treated as a gauge group, with ε an extra coordinate over and above the 4-space defined by system observables (including of course the self-adjoint state operators) gives rise to a 4-potential form ω and a corresponding antisymmetric field strength $F = d\omega$, so that electromagnetic interactions have a natural place in a model of spacetime based on the simplest quantum systems. The noncompact component of positive reals $\{e^{\rho}\}$ would then serve as the standard normalization factor of the usual treatments.

Finally, suppose everything above were agreed upon. What difference would it make? After all, quantum theory has done quite well in its 60-year history by having macroscopic spacetime simply put in by hand as a classical substrcuture over which wave functions, operators fields, etc., exist. Would a model that derives spacetime structure as a property of the simplest quantum systems be expected to offer any new insights, predictions, or directions for research? Clearly, yes. Any thēory deriving spacetime from quantum substructure would be expected to have within itself the means of relating curvature of the resulting spacetime to properties of the substructure, say some measure of quantum system density, and thus of giving a testable quantum theory of gravitation. Another direction for thought is implicit in even the oversimplified model outline above: if spactime is an epiphenomenon of observations made on a universe abounding with the simplest quantum systems, then more complicated quantum systems should live in our observed spactime only to the extent that they are embedded in a sea of the simplest systems. This fits well with the best theory we have for ordinary matter as clumpings of elementary systems (electrons) around cores of more complicated quantum systems (nucleons). It seems especially intriguing that the next most complicated systems allowed by quantum theory, based on a complex projective 2-space of pure states, have nine-dimensions for their space of observables (with an $8+1$ splitting) and transformation group $GL(3, C)$ containing $SU(3)$ as special unitary subgroup. Thus, even the simplest model of the type proposed gives an explanation for why nuclear matter, when isolated in the form, say, of a neutron star, starts exhibiting the behavior of a singularity in the structure of ordinary spacetime: such matter essentially lives in higher dimensions, and cannot be fitted smoothly into four dimensions. In any case, with all the expected advantages from a theory that presents spacetime as an explained phenomenon rather than simply an assumed necessity, it seems clear that, before writing off the structure of the simplest quantum systems as mere coincidence, we should thoroughly investigate the possibility that such systems define the overall structure of the observed universe.

REFERENCE

Marlow, A. R. (1984). *International Journal of Theoretical Physics*, **23**(9), 863–886.