Unitarity in Higher-Derivative Quantum Gravity

E. T. Tomboulis
Department of Physics, Princeton University, Princeton, New Jersey 08544,
and Department of Physics, University of California, Los Angeles, California 90024
(Received 5 December 1983)

A Euclidean lattice formulation is presented of the general fourth-order gravitational action involving $R_{\mu\nu}^2$ terms. This lattice theory is bounded and reflection positive. The existence of a positive Hilbert space and Hamiltonian, and hence unitarity, follow then from the standard nonperturbative Osterwalder-Schrader construction.

PACS numbers: 04.60.+n, 11.10.Gh

It is well known that the general fourth-order gravitational Lagrangian[1]

$$\mathscr{L}=(-g)^{1/2}\{-\lambda\kappa^{-4}+\gamma\kappa^{-2}R-\alpha^{-2}(R_{\mu\nu}^2-\tfrac{1}{3}R^2)-\beta R^2+(\text{topological terms})\} \tag{1}$$

defines a renormalizable theory.[2] Even more important is perhaps the fact that the theory is asymptotically free,[3] so that it may actually exist as a truly cutoff-independent, interacting continuum field theory in $d=4$. On the other hand, (1) has traditionally been rejected, because the fourth-order derivatives lead to ghosts in the perturbation series about the linearized theory. However, more recent work has given a number of indications that this may be misleading. These come first from the $1/N$ expansion,[4] N being the number of matter fields, where the theory is unitary in the $N\to\infty$ limit[5]; a Hamiltonian quantization of the conformal version of (1),[6] where no ghosts are seen to leading order in a strong coupling expansion; and, at the *classical* level, the zero–total-energy theorem of Boulware, Horowitz, and Strominger[7] for (1) without the Hilbert-Einstein terms.

It would seem that the unitarity picture presented by the linearized theory is simply too naive, and is substantially modified by nonperturbative effects. This is not surprising in an asymptotically free theory, where the actual physical, asymptotic states have, in general, nothing to do with the excitations seen in perturbation theory. The dynamical question of unitarity needs then to be examined at a nonperturbative level. To do so, we will write (1) as a lattice gauge theory of the Lorentz group which, after the Euclidean rotation, has become the compact group O(4). Our formalism is close in spirit to Utiyama's original work (in the continuum),[8] where the connections are the gauge fields of the Lorentz group, and the vierbeins are coupled as additional "matter" fields.[9]

Let us define the theory on a regular hypercubic lattice Λ with spacing a, in $d\geq 4$ Euclidean dimensions. On every bond $b=(n,\hat{\mu})$ originating at site n and extending in the positive μ direction, I introduce the variable $U_b=U_\mu(n)\in \mathrm{O}(4)$, with $U_{-\mu}(n+\hat{\mu})=U_\mu^{-1}(n)$. I will use the four-dimensional spinorial representation constructed, in the familiar manner, from the Clifford algebra generated by the (Euclidean) Dirac matrices γ^a satisfying $\gamma^a\gamma^b+\gamma^b\gamma^a=2\delta^{ab}$, $\gamma^{a\dagger}=\gamma^a$, $a=1,\ldots,4$. We then have $M^{ab}=\frac{1}{4}i[\gamma^a,\gamma^b]=M^{ab\dagger}$ for the group generators, and for elements continuously connected to the identity, we can write $U_\mu(n)=\exp[ia\omega_\mu^{ab}(n)M^{ab}/2]$. ω_μ are the connections. The vierbein fields are introduced via the d variables $H_\mu(n)\equiv\exp[iah_\mu^a\gamma_a]$ which are taken to reside on *sites*, with $H_{-\mu}(n)\equiv H_\mu^{-1}(n)$. The index μ is then a label that introduces one of these variables for every space-time direction μ. Note that $H_\mu(n)$ are unitary matrices. To obtain the general fourth-order Lagrangian, it is convenient to employ auxiliary fields $f_\mu^a(n)$ and $b(n)$ which are introduced as site variables in exactly the same way, i.e., $F_\mu(n)=\exp[iaf_\mu^a\gamma^a]$, $F_{-\mu}=F_\mu^{-1}(n)$, and $B_\mu(n)=\exp[iab(n)h_\mu^a\gamma^a]$, $B_{-\mu}(n)=B_\mu^{-1}(n)$. I then define the quantities

$$\begin{aligned}
Q_{1,\mu\nu}(n)&\equiv\tfrac{1}{2}\{H_\mu(n)U_{\mu\nu}(n)F_\nu(n)+H_{-\mu}(n)U_{\mu\nu}(n)F_{-\nu}(n)\},\\
Q_{2,\mu\nu}(n)&\equiv\tfrac{1}{2}\{H_\mu(n)U_{\mu\nu}(n)B_\nu(n)+H_{-\mu}(n)U_{\mu\nu}(n)B_{-\nu}(n)\},\\
Q_{3,\mu\nu}&\equiv\tfrac{1}{2}\{H_\mu(n)U_{\mu\nu}(n)H_\nu(n)+H_{-\mu}(n)U_{\mu\nu}H_{-\nu}(n)\},\\
Q_{4,\mu\nu}&\equiv\tfrac{1}{2}\{H_\mu(n)H_\nu(n)+H_{-\mu}(n)H_{-\nu}(n)\},
\end{aligned} \tag{2}$$

with

$$U_{\mu\nu} \equiv U_\mu(n)\,U_\nu(n+\hat\mu)\,U_{-\mu}(n+\hat\mu+\hat\nu)\,U_{-\nu}(n+\hat\nu) \tag{3}$$

the product of U_b's around the plaquette $p=(n,\mu\nu)$, and

$$Q_k(n) \equiv \epsilon^{\mu\nu\kappa\lambda}\{\sum_{l=1}^{4} c_{kl}[Q_{l,\mu\nu}(n)Q_{l,\kappa\lambda}(n) - c_l U_{\mu\nu}(n)U_{\kappa\lambda}(n)]\}, \quad k,l=1,\ldots,4, \tag{4}$$

with $c_k=1$ if $k\neq 4$, $c_4=0$, and $c_{kk}=1$, $c_{12}=-4$, $c_{32}=-8$, all other $c_{kl}=0$. Now let r_μ denote a reflection about the $(d-1)$-dimensional plane through a given site n, and perpendicular to the μ direction. For every pair of directions $[\mu\nu]$, $\mu<\nu$, consider the set $\{r^{(\alpha)}\}$ $\alpha=0,\ldots,3$, of possible such reflections: $r^{(0)}\equiv 1$ (no reflection), $r^{(1)}\equiv r_\mu$, $r^{(2)}\equiv r_\nu$, $r^{(3)}\equiv r_\nu r_\mu$. Let us then define

$$Q_k^{(\alpha,\alpha')}(n) \equiv \epsilon^{\mu\nu\kappa\lambda}\{\sum_{l=1}^{4} c_{kl} r^{(\alpha)} r^{(\alpha')}[Q_{l,\mu\nu}(n)Q_{k,\kappa\lambda}(n) - c_l U_{\mu\nu}(n)U_{\kappa\lambda}(n)]\}. \tag{5}$$

In (5), $r^{(\alpha)}$ and $r^{(\alpha')}$ refer to the $(\mu\nu)$ and $(\kappa\lambda)$ pairs of indices, respectively. $r^{(\alpha)}$, $r^{(\alpha')}$ act on $Q_{k,\mu\nu}Q_{k,\kappa\lambda}$ modifying the definitions (2)–(4) by geometrical reflection in the obvious way. The lattice action is now taken to be

$$A_\Lambda = \sum_{n\in\Lambda}\sum_{\alpha,\alpha'}\sum_k \tfrac{1}{64}\beta_k\{\tfrac{1}{2}\mathrm{tr}[Q_k^{(\alpha,\alpha')}(n)]^2 - \tfrac{1}{8}[\mathrm{tr}\,Q_k^{(\alpha,\alpha')}(n)]^2 + \text{c.c.}\}^{1/2}. \tag{6}$$

It is perhaps worth noting that one is led naturally to this summation over reflections in order to preserve discrete rotation and reflection invariances on the lattice, which are broken by the dual plaquette interaction in (4). This interaction form occurs quite generally in lattice gravitational actions,[10] since $\epsilon^{\mu\nu\kappa\lambda}$ is the only constant tensor available for contracting space-time indices. Also, following the suggestion in Ref. 11, squaring and later taking the square root ensures scalar rather than pseudoscalar properties in the continuum limit.

Equation (6) is invariant under local O(4) gauge transformations, where $U_\mu(n)\to V(n)\times U_\mu(n)V^{-1}(n+\hat\mu)$, $V\in \mathrm{O}(4)$, whereas h_μ^a, f_μ^a transform as "isovector matter" fields in the defining real four-dimensional vector representation, and $b(n)$ transforms as a singlet. This is simply the transformation $H_\mu(n)\to V(n)H_\mu(n)V^{-1}(n)$, $F_\mu(n)\to V(n)F_\mu(n)V^{-1}(n)$, $B_\mu(n)\to V(n)B_\mu(n)V^{-1}(n)$. This follows from the well-known fact that γ^a transforms like a vector operator under the group, i.e., $V\gamma^a V^{-1}=D^{ab}\gamma^b$, where V is the spinorial and D the defining four-dimensional representation of some element of O(4).

In Einstein gravity, the equations of motion constrain the connection to be the usual function of the vierbeins. For the general fourth-order theory (1), this is not true even at the classical level, so that if the usual relation (i.e., zero torsion) is to hold, it must be imposed by hand. An appropriate constraint is

$$\chi_{\mu\nu}(n) \equiv \mathrm{Re\,tr}\{i\gamma^a \tfrac{1}{2}[W_{\mu\nu}(n)+\tilde W_{\mu\nu}(n)]\}=0, \tag{7}$$

where $W_{\mu\nu}(n)$, $\tilde W_{\mu\nu}(n)$ denote the product of

$$W_\mu(n)\equiv H_\mu(n)U_\mu(n), \quad \tilde W_\mu \equiv U_\mu(n)H_\mu(n+\hat\mu), \tag{8}$$

respectively, around the plaquette $p=(n,\mu\nu)$.

We may now write down the partition function Z_Λ and the corresponding measure $d\mu_\Lambda$:

$$d\mu_\Lambda = \frac{1}{Z_\Lambda}\prod_{b\in\Lambda} dU \prod_{n\in\Lambda} db(n)\prod_\mu dh_\mu(n)\,df_\mu(n)\prod_{p\in\Lambda}\delta[\chi_p]\exp[-A_\Lambda]. \tag{9}$$

Z_Λ is defined by $\int d\mu_\Lambda = 1$. In (9), dU denotes the invariant Haar measure over the group, and dh, etc., the invariant Lebesgue measure over the site variables. From (7), we see that the effect of a gauge transformation on the constraint is to simply rotate the γ^a's among themselves. Hence, the measure (9) is indeed gauge invariant.

In the naive continuum limit, the Q_1 part of (6), for example, reduces to

$$\epsilon^{\mu\nu\kappa\lambda}\epsilon_{abcd}(\mathscr{R}^{ab}_{1,\mu\nu}\mathscr{R}^{cd}_{1,\kappa\lambda} - 4\mathscr{R}^{ab}_{2,\mu\nu}\mathscr{R}^{cd}_{2,\kappa\lambda} + 3R^{ab}_{\mu\nu}R^{cd}_{\kappa\lambda})$$

with

$$\mathscr{R}^{ab}_{1,\mu\nu} \equiv R^{ab}_{\mu\nu} - 2(h^a_\mu f^b_\nu - h^b_\mu f^a_\nu), \qquad \mathscr{R}^{ab}_{2,\mu\nu} \equiv R^{ab}_{\mu\nu} - 2b(n)(h^a_\mu h^b_\nu - h^b_\mu h^a_\nu),$$

and

$$R^{ab}_{\mu\nu} = -\partial_\mu \omega^{ab}_\nu + \partial_\nu \omega^{ab}_\mu + \omega^{ac}_\mu \omega^{cb}_\nu - \omega^{ac}_\nu \omega^{cb}_\mu,$$

the usual Riemann curvature. The remaining terms in (6) reduce to similar expressions. The auxiliary fields f^a_μ, b are nonpropagating, and can be eliminated. This step imitates the procedure in Ref. 12. The constraint χ_p reduces to the usual zero-torsion condition. The classical Lagrangian (1) is then reproduced[13] with

$$\alpha^{-2} = \beta_1, \quad \beta = \tfrac{1}{3}(\beta_1 + \tfrac{1}{4}\beta_2 + 2\beta_3), \quad \gamma = 4\beta_3, \quad \lambda = (24\beta_3 + 12\beta_4).$$

We have produced a locally O(4)-invariant theory, which reproduces (1) in the naive continuum limit. As always, such a construction is not unique, and, indeed, various alternative actions can be given, all differing by terms which vanish as $a \to 0$. The particular choice (9) was dictated by relative simplicity, and, most importantly, by the fact that, at least for $a > 0$, it satisfies certain axioms of Euclidean field theory.[14] Indeed, first note that (6), being constructed out of unitary matrices, is bounded. Hence the expectation of any bounded observable calculated in the measure (9) is regular and bounded. Also, with appropriate, e.g., periodic, boundary conditions, the theory is obviously translation invariant. Furthermore, and this is the crucial property for our purposes, (9) satisfies reflection positivity about $(d-1)$-dimensional planes *with sites.* Indeed, let $\Lambda_\pm$ denote the part of Λ to the "right" or "left," respectively, of such a plane π, and let $\Lambda_0 = \Lambda \cap \pi$. Let $\mathscr{F}_\pm$ be any function of the field variables with support only in $\Lambda_\pm \cup \Lambda_0$, and θ the reflection about π, defined by

$$\theta \mathscr{F}_+(U, H, \ldots) = \bar{\mathscr{F}}_+(\theta U, \theta H, \ldots),$$

where the bar denotes complex conjugation, and

$$\theta U_b \equiv U_{rb}, \quad \theta H_\mu(n) \equiv H_{r\mu}(rn),$$

and similarly for $F_\mu(n)$, $B_\mu(n)$. Here r denotes geometrical reflection about π, and for a bond $b = \langle n,m \rangle = (n, \hat{\mu})$ along $\hat{\mu}$ between nearest-neighbor sites n and m let us write

$$rb = \langle rn, rm \rangle \equiv (rn, r\hat{\mu}).$$

Then it can be verified that

$$\langle \mathscr{F}_+(\theta \mathscr{F}_+) \rangle \geq 0 \tag{10}$$

for arbitrary $\mathscr{F}_+$, with the expectation $\langle \cdots \rangle$ calculated in the measure (9). Now there is a standard construction[14] that leads from (10) to the definition of a physical Hilbert space $\mathscr{H}$ with positive norm, and the demonstration that the transfer matrix $T_{\mathscr{H}}$ acting on $\mathscr{H}$ is a positive[15] self-adjoint operator of norm less than 1, i.e., $T_{\mathscr{H}} = e^{-tH}$, $t \in Z$, with $H = H^\dagger \geq 0$. H is then the (positive) Hamiltonian. These properties hold for arbitrary lattice spacing $a > 0$, and, therefore, should also persist in the limit $a \to 0$, if it exists.

Gauge-invariant Green's functions computed in the measure (9) will consist of loops formed out of U, H. Green's functions such as those considered in the usual continuum perturbation theory, e.g., strings of $h^a_\mu h^b_\nu \delta_{ab} \equiv g_{\mu\nu}$'s, can also be computed. Note that both kinds are group singlets. For any such n-point function $G_n = G_n(\underline{x}_1, t_1, \ldots, \underline{x}_n, t_n)$, then, the customary analytic continuation $t_j \to it_j$ defines a Minkowski function W_n. The Minkowski theory is defined by the set $\{W_n\}$. The Hilbert space $\mathscr{H}$ defined directly from (10) is the Hilbert space of the Minkowski theory, and the spectrum positivity condition is satisfied.[14]

We have constructed a lattice theory with postive transfer matrix acting on a Hilbert space of positive norm, i.e., a unitary theory. The argument, relying as it does on a property of the complete measure, given in closed, regularized form on the lattice, is completely nonperturbative. Furthermore, it is independent of the values of the couplings in (6), and hence of any phase transitions that may occur in the system, i.e., it holds in all phases. By the same token, however, it provides no information about the actual structure of the physical states. In particular, the mechanism by which the negative-energy excitations of the short-distance regime, described by perturbation theory, must disappear in the physical, asymptotic states remains obscure.

Equation (9) possesses (1) as its naive continuum limit. Does this persist in the renormalized limit? Though, of course, an actual construction of the continuum in $d=4$ is beyond present techniques, a lot can be learned about this issue in an asymptotically free theory by comparing the lattice weak coupling perturbation series as $a \to 0$ to that in the continuum theory. Such a comparison will allow one to pick a specific renormalization-group trajectory in the coupling-constant space of O(4) gauge theories that leads to the four-parameter subspace of Eq. (6). As we saw above, this is the subspace containing the generally covariant, attractive fixed point in the ultraviolet that leads to the correct continuum limit as $a \to 0$. In this connection it need hardly be pointed out that the lattice was introduced as a convenient regulator that allows precise statement and rigorous demonstration of certain nonperturbative assertions; it is, however, rather awkward when it comes to recovering general covariance in the continuum.

Another interesting question concerns the inclusion of the Einstein-Hilbert terms in (1) or (6). Note that the argument in this paper goes through with or without these terms in (6). It would be attractive not to include them, and to hope that they will be induced in the effective action as a long-distance effect.[16] Lattice nonperturbative techniques such as mean-field theory, approximate recursion relations, and Monte Carlo computations may be useful in investigating this possibility, and, more generally, exploring the phase structure of this theory.

I would like to thank T. Banks, A. Strominger, and especially L. Jaffe for discussions. This work is supported by the National Science Foundation under Grant No. PHY 80-19754.

[1]Dimensional couplings are written in units of an appropriate (e.g., Planck) mass $1/\kappa$.

[2]K. S. Stelle, Phys. Rev. D **16**, 953 (1977).

[3]E. T. Tomboulis, Phys. Lett. **97B**, 77 (1980); F. S. Fradkin and A. A. Tseytlin, Nucl. Phys. **B201**, 469 (1982).

[4]E. T. Tomboulis, Phys. Lett. **70B**, 361 (1977).

[5]Unitarity violations present in higher orders in $1/N$ can also be avoided (Ref. 4) if one adopts the Lee-Wick prescription.

[6]M. Kaku, Phys. Rev. D **27**, 2819 (1983).

[7]D. G. Boulware, G. T. Horowitz, and A. Strominger, Phys. Rev. Lett. **50**, 1726 (1983).

[8]R. Utiyama, Phys. Rev. **101**, 1597 (1956).

[9]There are several more recent, and even elegant, formulations of gravity as a gauge theory [e.g., T. W. B. Kibble, J. Math. Phys. **2**, 212 (1960); S. W. MacDowell and F. Mansouri, Phys. Rev. Lett. **38**, 739 (1977); Y. Ne'eman and T. Regge, Phys. Lett. **74B**, 54 (1978); J. Theirry-Mieg, Lett. Nuovo Cimento **23**, 498 (1978); M. Kaku, P. K. Townsend, and P. Van Nieuwenhuizen, Phys. Lett. **69B**, 304 (1977); P. C. West, Phys. Lett. **76B**, 569 (1978)] and they can be transcribed on the lattice [L. Smolin, Nucl. Phys. **B148**, 333 (1979); A. Das, M. Kaku, and P. K. Townsend, Phys. Lett. **81B**, 11 (1979); Kaku, Ref. 6]. However, the noncompactness of the groups used generally leads to unbounded lattice actions, and hence ill-defined functional integrals.

[10]Smolin, Ref. 9; Das, Kaku, and Townsend, Ref. 9; Kaku, Ref. 6.

[11]West, Ref. 9.

[12]Kaku, Townsend, and Van Nieuwenhuizen, Ref. 9.

[13]Our fields have conventional dimensions of mass. The usual formalism with dimensionless vierbeins is regained by a trivial rescaling $h^a_\mu \to \kappa^{-1} h^a_\mu$.

[14]K. Osterwalder and R. Schrader, Commun. Math. Phys. **31**, 83 (1973), and **42**, 281 (1975); J. Glimm and A. Jaffe, *Quantum Physics* (Springer-Verlag, New York, 1981); for lattice gauge theories, D. Brydges, J. Fröhlich, and E. Seiler, Ann. Phys. (N.Y.) **121**, 227 (1979).

[15]Strictly speaking, the standard argument from (10) shows that $0 < T \leq 1$ in the continuum. On the lattice, it leads to $0 \leq T \leq 1$; to show $T > 0$, if desired, an additional argument is then required. See, e.g., M. Lüscher, Commun. Math. Phys. **54**, 283 (1977).

[16]S. L. Adler, Rev. Mod. Phys. **54**, 729 (1982); A. Zee, Phys. Rev. D **23**, 858 (1981).