Quantum metric spaces as a model for pregeometry

Enrique Alvarez*
Universität Bern, Institut fur Theoretische Physik, Sidlerstrasse 5, CH 3012 Bern, Switzerland

Juli Céspedes and Enric Verdaguer
Grup de Física Teòrica, Universitat Autònoma de Barcelona, 08193 Bellaterra, Barcelona, Spain
(Received 25 July 1991)

A new arena for the dynamics of spacetime is proposed, in which the basic quantum variable is the two-point distance on a metric space. The scaling dimension (that is, the Kolmogorov capacity) in the neighborhood of each point then defines in a natural way a local concept of dimension. We study our model in the region of parameter space in which the resulting spacetime is not too different from a smooth manifold.

PACS number(s): 04.60.+n

I. INTRODUCTION

One of the outstanding unsolved problems in theoretical physics is the description of the quantum effects of the gravitational interaction (cf., for example, Ref. [1] for a general review). Dimensional arguments indicate that the energy scale at which these effects should become apparent is the Planck mass $m_P=\sqrt{c\hbar/G}$, where G is Newton's constant. Assuming the validity of the usual Heisenberg uncertainty principle up to these energy scales, this means that the length scale at which quantum effects should modify the general-relativistic description of spacetime as a differentiable manifold is the Planck length $l_P=\sqrt{G\hbar/c^3}$.

At still smaller distances, $l<l_P$, it is likely that the description of spacetime as a smooth manifold is not adequate. It is actually possible that there is a minimal (physically measurable) length; this is suggested, for example, by the studies of the scattering of strings at very high energies (cf., for example, Ref. [2]). Perhaps the simplest intuitive picture one can get on the "melting" of classical spacetime when one uses "probes" at length scales of the order of l_P is to imagine that the actual "classical" metric is given by the expectation value of a quantum operator: $g_{\mu\nu}=\langle \hat{g}_{\mu\nu} \rangle$. This raises the possibility that the metric (or the vielbein) can become degenerate (det $g=0$), changes signature or even the dimension of the spacetime itself. Of course more drastic possibilities can be imagined, which are not by any means similar to any smooth manifold whatsoever.

We are then led towards the most pressing physical question: what is the mathematical structure most appropriate for describing physics at those scales? There are several possibilities, according, in particular, to the more or less basic role one imagines for the spacetime signature; if one adopts the extreme point of view that this will be an essential ingredient of quantum gravity, in the sense that no analytic continuation to "Euclidean" space is possible or convenient, then methods specifically tailored for signature $s=-2$ are probably most adequate. Among them, let us mention the entire Penrose program of twistors (cf., for example, the book [3]). A strategy with the same starting prejudices is the one which stems from the consideration of the ensemble of partial ordered sets (or posets, for short, cf. [4]); and in general the approach pioneered by Finkelstein (cf., for example, [5]). Another extreme possibility is that only the topology remains, that is, that there is a set of points without any further physical specification in addition to the one which singles out certain subsets as open sets, thereby defining the topology itself; one is then led to consider theories of "quantum topology" such as the ones recently put forward in Ref. [6].

The alternative we want to explore in this paper (and which was first suggested by one of us in Ref. [7]) is much more conservative in character; we shall assume that quantum physics is best formulated in Euclidean space "ab initio," and that a proper analytic continuation to the physical signature can be performed at a late stage of the computations. While there is no deep understanding of why this should be so, it is in agreement with the modern trends in quantum field theory. We shall also assume, in particular, that it still makes sense to speak about "spacetime events" even in the Planck regime, when spacetime is not a manifold anymore. To be sure, our intuitive notion of "event" stems from its description in terms of four coordinates. What we are postulating here is that the concept is more fundamental than the local coordinates themselves.

We would like to endow this set of points with as much mathematical structure as possible, but still without forcing them to condense in a smooth manifold. The most natural framework in this context is a "metric space," i.e., we assume that the set of events has a two-point real function defined on it, with the properties that it is symmetric, that is, $d(x,y)=d(y,x)$; it is zero only when $x=y$, $d(x,y)=0 \Leftrightarrow x=y$; and satisfies the triangular inequality, namely, $d(x,y)\leq d(x,z)+d(z,y)$, $\forall z$.

*On leave from Universidad Autónoma de Madrid, 28049 Madrid, Spain.

To be specific, let us show a simple example of the type of spaces covered by our formalism. Consider the subset of R^3 consisting of $\{z=0\}\cup\{x=y=0\}$ (which is not a differentiable manifold, being the union of the plane xy with the z axis). This set can be endowed with a structure of metric space by defining d_{xy} in an obvious way if both x,y are either in the plane or else in the z axis; and when $p\in R^2$ and $q\in R$, we define $d_{pq}\equiv|z|+\sqrt{x^2+y^2}$. It is plain that this is indeed a metric space which has locally dimensions one or two. The aim of this paper is to study the simplest quantum dynamics of this class of spaces.

This framework is obviously more general than the general-relativistic one, in the sense that every Riemannian manifold is a metric space, with distance given by the "geodesic distance," i.e., $d(x,y)=\int_x^y ds$. This is essentially Synge's "world function" (cf., for example, Ref. [8]). Actually, Synge worked in a pseudo-Riemannian space, with the physical Lorentzian signature, so that he squared the whole thing in order to have a non-negative quantity. The world function $\Omega(x,y)=d^2(x,y)$ is anyway a more convenient quantity than the distance itself, at least in the framework of Riemannian spaces. It is obvious that in order for an arbitrary distance to qualify as a geodesic distance of a smooth manifold it should satisfy necessarily certain additional constraints; for example, given any pair of points (x,y), there is always a third point z which saturates the triangular inequality, that is, $d(x,y)=d(x,z)+d(z,y)$. An arbitrary metric space will not satisfy this property, which means that its distance is not a geodesic distance of a certain smooth manifold.

The preceding analysis implies that a neighborhood of a given point will not be generically homeomorphic to R^n; in other words, the events in the Planck regime will not be described by coordinates except in the simplest examples.

In the general framework just introduced, the distance is a quantum field itself. We have not been able to figure out any general guiding principle, analogous to general covariance, which would allow us to write an action for the bilocal quantum field $d(x,y)$. We shall not distinguish between the quantum variable $d(x,y)$ itself, and its average value $\langle d(x,y)\rangle$, where the average value is defined with respect to the functional measure associated with the action for the field $d(x,y)$. As a consequence the models to be considered in this paper are just the simplest ones, both analytically and from the point of view of numerical simulations.

We would then like to write down an effective field theory which aims to describe the physics corresponding to an energetic regime in which $E\sim m_P$, but definitely not much bigger. It is natural to expect that the semiclassical metric space defined by using as a distance the expectation value of the quantum field $d(x,y)\equiv\langle\hat{d}(x,y)\rangle$, although not smooth, will be "almost smooth" most of the time.

This means in particular, that if we count the number of points in a closed ball centered at the point x and with radius R,

$$N(x,R)\equiv\sum_y\theta(R-d(x,y))\ , \tag{1.1}$$

then, in some places at least, this will scale in the local limit $R\to 0$ as

$$N(x,R)\sim R^{n(x)}\ . \tag{1.2}$$

This definition of dimension is known in the mathematical literature as "Kolmogorov capacity," and is equivalent in many simple situations to the Hausdorff dimension (cf. Ref. [9] for a clear, physicist-oriented, exposition).[1] The picture of spacetime implied by this description is then a set of points that in the vicinity of some particular point looks like a manifold with some value for the dimension; whereas in the neighborhood of some other point it looks again like a smooth manifold, but with a different value for the dimension, and, finally, around some other places (which are not expected to be too frequent if the energy is not much bigger than Planck's mass), it does not look at all like R^n.

In some subset of the parameter space of the model, spacetime will just look as an ordinary differentiable manifold with constant dimension. Not far from this subset (in a convenient topology), we will find other models in which the dimension varies from place to place, but in a smooth way. Those are the models which would describe the physical situation we are interested in: we want to understand what happens when the smooth spacetime manifold disappears and melts into more complicated topological spaces.

In this paper we are not able, unfortunately, to complete the above program. In Sec. II we shall study in some detail some analytically tractable toy models, which give an accurate description of the physical picture we have in mind. In Sec. III some more general toy models which require numerical simulations are discussed; these simulations were performed using simple variations of the Metropolis algorithm. In Sec. IV we establish some general properties of models representing two extreme types, namely, the most similar and the most dissimilar to R^n in a certain precise sense. Finally, our (provisional) conclusions are stated in Sec. V.

A point that will not be further elaborated in this paper, but which is most interesting, is that our general framework in the classical case (tree approximation) is broad enough to cover large classes of singular spacetimes (cf., for example, [10]); besides, as a metric space there is a general way of completing it, in case it is not complete to begin with. A brief presentation of the present ideas can be found in [11].

---

[1] As an example of the situations that we may encounter with such definition, let us consider a discretization of the real line with a density distribution of points given by some smooth bounded function, which may have been dynamically generated. At the points x where that distribution is nonzero we obviously have $n(x)=1$, but if there is a zero of order k at a point x_0 the dimension is $n(x_0)=k+1$. So that this space would be mainly one dimensional except at some points where it changes dimension.

II. SOME ANALYTICAL TOY MODELS

In order to build up some tractable models, one could try to discretize somewhat the problem, that is, to consider that the spacetime has a finite number of points, N. Now there is a small problem here: given any set of points (either finite or infinite), we can endow it with the structure of a metric space, by defining what is called the "discrete metric," namely,

$$d_0(x,y)\equiv 1-\delta_{x,y} \ . \tag{2.1}$$

It is obvious that the corresponding topology is the "discrete topology," in which all subsets of the topological space are open sets. This topology is actually too trivial to be of any physical interest: every function of the space in itself is a continuous one. We certainly do not want our spacetime to have as simple a structure as that. Unfortunately, this is exactly what happens in every finite metric space: let us call

$$\epsilon\equiv \min_{x,y\in M} d(x,y) \quad (x\neq y) \ . \tag{2.2}$$

According to the postulates, this must be a nonzero number. Now, it is very easy to see that the set composed by a single element x is an open ball of radius R (where R can be any real number such that $R<\epsilon$),

$$\{x\}=B_x(R) \ . \tag{2.3}$$

Given that every union of open sets is still an open set, this proves our claim on the triviality of a finite metric space.

This means, of course, that we have to complicate somewhat our model and consider, at least, the limit $N\to\infty$. For N points, it seems reasonable to start with a model in which the dynamical variable is a symmetric $N\times N$ matrix and to associate the distance between points with the matrix elements M_{ij} themselves. But M will represent a distance only if it satisfies the distance conditions: it must be symmetric, with vanishing diagonal and satisfy the triangular inequality. The last one poses a very strong constraint on the matrices, and a way of implementing this constraint is considered in the next section, but it cannot be solved analytically.

However, one can circumvent this difficulty by defining the distance just through the eigenvalues $\{\lambda_i\}$ of M in the following way:

$$d_{ij}=|\lambda_i-\lambda_j| \ .$$

It can be easily shown that this definition satisfies the triangular inequality automatically, as well as the remaining distance properties provided that $\lambda_i\neq\lambda_j$, when $i\neq j$. It would appear that this definition corresponds to an inherently one-dimensional situation, but this is not the case as we shall prove in what follows.

Next, we should postulate an action for the dynamical field, M. Now, the distance is invariant under similarity transformations: the distance one gets from M is the same as that obtained from $\tilde{M}\equiv O^T MO$, where O is an orthogonal matrix. Thus, we can restrict the action to be also invariant under such transformations. It can be proved that any invariant object must be a function of the traces of the first N powers of M. We will choose our toy-model action to be

$$S=\sum_{p=1}^{\infty}\bar{g}_p \mathrm{tr}M^{2p} \ ,$$

although this is not the most general invariant action, it has enough parameters to be a nontrivial model, it is amenable to analytic treatment, and exhibits some of the features we expected from a more general model.

The partition function is defined by

$$Z(g)=\int dM \exp\left[-\sum_{p=0}^{\infty}\bar{g}_p \mathrm{tr}M^{2p}\right] \ . \tag{2.4}$$

In the preceding formula, the coupling constants $\bar{g}_p$ have been rescaled, $\bar{g}_p=g_p/N^{p-1}$. On the other hand, the measure dM is the unitary (or orthogonal in the real case) invariant measure on Hermitian matrices (cf., for example, Ref. [12]). There is a large literature on the subject of the dynamics of random matrices and, in particular, in the references in [13], it has been noticed how the partition function could be expressed for arbitrary N in terms of convenient orthogonal polynomials:

$$P_n(\lambda)\equiv[Z_n(g)]^{-1}\int d\mu(\lambda_1)\cdots d\mu(\lambda_n)\Delta^2\prod_{i=1}^{n}(\lambda-\lambda_i) \ , \tag{2.5}$$

where $d\mu(\lambda)$ is the measure determined by the potential,

$$d\mu(\lambda)\equiv d\lambda \exp[-V(\lambda)] \ , \tag{2.6}$$

where $V(\lambda)$ is the function associated with the potential,

$$V(\lambda)=\sum_{p=1}^{\infty}\bar{g}_p\lambda^{2p} \ , \tag{2.7}$$

$\Delta(\lambda)$ is the Vandermonde determinant,

$$\Delta(\lambda_1,\ldots,\lambda_n)\equiv\prod_{i>j}(\lambda_i-\lambda_j) \ , \tag{2.8}$$

and $Z_n(g)$ refers to the same integral as in the defining formula of the partition function, in the case in which the dimensionality of the matrices is (n,n) (with $\bar{g}_p$ kept fixed). The final result of Bessis, Itzykson, and Zuber for the density of eigenvalues in the limit $N\to\infty$ is the even function

$$u(\lambda)=\frac{1}{\pi}\int_{|\lambda|/2}^{a}\frac{d\xi\, w'(\xi)}{\sqrt{4\xi^2-\lambda^2}} \ , \tag{2.9}$$

where the auxiliary function $w(\xi)$ is defined as

$$w(\lambda)=\frac{i}{2\pi}\oint\frac{dz}{z}\lambda\left[z+\frac{1}{z}\right](\partial V)\lambda\left[z+\frac{1}{z}\right] \ . \tag{2.10}$$

This can easily be shown to be equivalent to

$$w(\lambda)\equiv\sum_{p=1}\frac{(2p)!}{p!(p-1)!}g_p\lambda^{2p} \tag{2.11}$$

and the value of the parameter a is defined as a function of the coupling constants by the implicit equation

$$1=w(a)=2g_1a^2+12g_2a^4+60g_3a^6+280g_4a^8+\cdots . \tag{2.12}$$

A beautiful proof of the preceding formulas can be found in the papers of the Saclay group cited in the references. Working out explicitly the integrals, one gets

$$u(\lambda)=\frac{1}{\pi\sqrt{4a^2-\lambda^2}}\sum_{p=0}^{\infty}u_p\lambda^{2p} , \tag{2.13}$$

where

$$\begin{aligned} u_0&=g_1+4a^2g_2+18a^4g_3+80a^6g_4+\cdots ,\\ u_1&=2g_2+6a^2g_3+24a^4g_4+\cdots ,\\ u_2&=3g_3+8a^2g_4+\cdots ,\\ u_3&=4g_4+\cdots . \end{aligned} \tag{2.14}$$

The preceding expressions generalize Wigner's semicircle formula, on the distribution of the set of zeros of Hermite polynomials. We want now to estimate the Kolmogorov capacity in the neighborhood of the point labeled by i. The quantity to be computed is

$$\begin{aligned} N_i(R)&=\sum_j\theta(R-d(i,j))\\ &=\int d\lambda\, u(\lambda)\theta(R-|\lambda_i-\lambda_j|)\\ &=\int_{\lambda_i-R}^{\lambda_i+R}d\lambda\, u(\lambda) \end{aligned} \tag{2.15}$$

which can be immediately worked out to be

$$N_i(R)=\frac{4aR\sin\theta_i}{\pi}\sum_{l=0}u_l(2a\cos\theta_i)^{2l} , \tag{2.16}$$

where we have assumed $R\ll a$, and that λ_i is not close neither to the end points nor to the central, symmetric point. We have defined

$$\cos\theta_i\equiv(\lambda_i+R)/2a . \tag{2.17}$$

In our simplest analytic model the Kolmogorov capacity grows linearly with R (for small R), as long as the point considered is not too close to the center of the spectrum. This means that our semiclassical metric space has almost constant dimension one in this region. This is not the case, however, for any neighborhood of the central point, because there we can write

$$N_0(R)=\frac{4a}{\pi}\sum_{p=0}^{\infty}\frac{up}{2p+1}R^{2p+1} . \tag{2.18}$$

Still, in the generic case, when $R\to 0$, the dominant power is 1, and the central zone is one dimensional as well. But for interactions, of the type $u_p=\delta_{pm}$, then the central region has dimension $2m+1$. This means that if we fine tune our potential, we can get around the central point as high a dimension as desired. It is perhaps worth noticing that due to the fact that all coefficients in Eq. (2.14) are positive, we need some $g_i<0$ in order to get $n(0)>1$.

It is simple enough to derive an explicit formula for the capacity in the vicinity of the end points:

$$N_e(R)=\frac{4a^{1/2}R^{3/2}}{3\pi}\sum_{p=0}u_p(2a)^{2p} .$$

This implies that $n(e)=3/2$.

The picture of the spacetime as implied by our toy model is then a region mainly one dimensional, with a higher-dimensional central neighborhood. This simple model then embodies some of the general characteristics we claimed any physically reasonable model ought to have above the Planck energy. We have, in particular, the possibility of introducing the dimension of spacetime as a physical, dynamically determined, quantum field.

III. MONTE CARLO SIMULATIONS

We do not want now to assume any special symmetry property of the distance, so that we will essentially identify the value of the distance between the points i and j with the element of the random matrix M_{ij}. To be specific,

$$d_{ij}=\sqrt{\epsilon^2(1-\delta_{ij})+4M_{ij}^2} , \tag{3.1}$$

where the small constant ϵ has been introduced to avoid $d_{ij}=0$ when $i\neq j$. As discussed at the beginning of Sec. II, the presence of that constant does not pose a problem because the finite number of points considered in our model implicitly imposes a lower bound for the values of the distance. Note that we have to impose the further restriction $M_{ii}=0$ in order for our definition to be consistent.

To begin with a simple case, we have simulated the action:

$$\begin{aligned} S=&g_1\mathrm{Tr}(M^2)+g_2\mathrm{Tr}(M^4)\\ &+\frac{1}{\xi}\sum_{i,j,k}\exp[\xi(d_{ij}-d_{ik}-d_{jk})] . \end{aligned} \tag{3.2}$$

The first term is analogous to an ordinary mass term. The second term is really an interaction term (the only reason why we have only a fourth-order interaction is one of simplicity; we plan to consider higher powers of M in future work). The last term has been introduced by hand, in order to enforce the triangular inequality: in the limit $\xi\to\infty$, the action blows up unless $d_{ij}-d_{ik}-d_{jk}\leq 0$, $\forall i,j,k=1,\ldots,N$, that is, unless the triangular inequality is satisfied. In this case, the prefactor ensures that the contribution of this term is negligible. This term formally resembles an exponential potential of the Liouville type.

The impossibility of numerically handling an infinite number of points is a primordial handicap of all discretizations. Following the standard philosophy of numerical simulations in gauge theories, we have performed several simulations using different lattice "volumes" in the hope of determining the dependence of our results with the size of the system, and in this way recover the infinite volume limit. Actually, we have been severely limited by the amount of time available in our computer, an IBM 3090 at the UAB (our longest runs took more than 100 minutes of CPU). We have been able to perform simula-

tions with values of N in the range between 12 and 96, at which point the amount of CPU time became prohibitively large.

It is well known (cf., for example, Ref. [14]) that in order to define a continuum theory, with finite mass gap, we have to look for a critical point of our discrete system: when the lattice spacing $a \to 0$, the correlation length $\xi \to \infty$. The way this is done in practice is to plot as many physical observable Θ as possible (in a dimensionless form) for the value a of the cutoff, and, say, for the value $a/2$, as a function of the length scale at which they are defined, r, and a set of bare coupling constants, which we represent generically by $g_0(a)$. Since we assume Θ dimensionless, this implies that $\Theta(2r,a,g_0(a/2)) = \Theta(r,a,g_0(a))$ [plus $O(a^2)$ corrections]. The study of the corresponding curves, plotted for fixed r and a, allows one to distinguish between ultraviolet repulsive and attractive fixed points. Unfortunately, the inherent nonlocality of our fundamental action exceedingly complicates the simulations, so that we only have for the moment preliminary results on the continuum limit.

Our program generates a random "initial configuration," i.e., an initial matrix M, and evaluates the action of that configuration. The configuration's evolution is carried out by a Metropolis algorithm in the following way: the first d_{ij} ("link") is changed, and the action of the new configuration evaluated. If the change in the action ΔS is negative, the change is immediately accepted. If $\Delta S \geq 0$ we accept the change with probability $\exp(-\Delta S)$. This procedure is called a sweep. After sweeping all the $N(N-1)/2$ links (that is, all the components d_{ij}), we are led to a different and nearly independent configuration. The whole process will be called an iteration. Only after an iteration can we average the interesting quantities: $\langle S \rangle$, $\langle d_{ij} \rangle$, $\cdots$. The average over configurations obtained this way is equivalent to average over completely random configurations weighted by $\exp(-S)$. In Fig. 1 we show the evolution of the average distance,

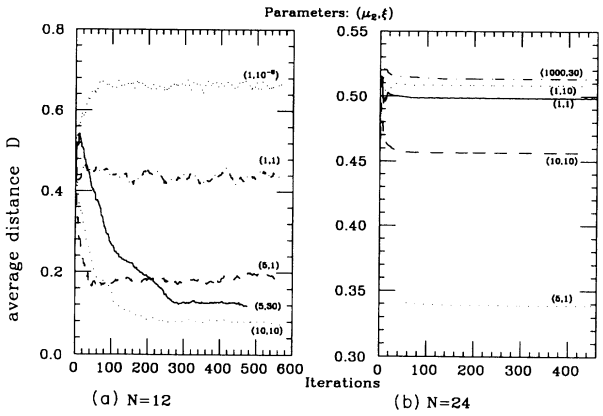

FIG. 1. The average value of the distance, $D \equiv [1/N(N-1)] \sum d_{ij}$, is plotted as a function of the number of iterations. It shows stabilization in the worst cases, after 300 iterations. The parameter μ_2 is just the g_2 coefficient in the main text, and g_1 has been fixed ($g_1 = 1$).

$$D \equiv \frac{1}{N(N-1)} \sum_{i,j<N} d_{ij} , \tag{3.3}$$

the number of iterations of the algorithm as a check of the stabilization of the method. The larger N is, the faster the stabilization takes place, and after that transient period we compute the averages

$$\langle d_{ij} \rangle \equiv \frac{1}{K} \sum_{n=1}^{K} d_{ij}^{(n)} \tag{3.4}$$

of $d_{ij}^{(n)}$, the distance function at the nth iteration. The total number of iterations K depends on the length of the computation, that is, on the number of points considered N.

A reasonable running allows for 20 000 iterations in the case $N=12$, 3000 iterations when $N=24$ and only 500 when $N=48$, because the complexity increases as N^2. The reason for that complexity is the Liouville term. We can think of the points X_i as belonging to a lattice, and the d_{ij} as links between them. The Metropolis algorithm's efficiency lies on the locality of the action, but in our problem d_{ij} links any point with all the others, and, in addition, the action is highly nonlocal. If the action was indeed local, after a change of one "link" d_{ij}, we would expect a change of a few terms in the action. But in our case, a change in one component d_{ij} induces a change in a lot of terms in the action, slowing down the efficiency of the algorithm.

After having computed the $\langle d_{ij} \rangle$ we proceed to compute the Kolmogorov capacity. For each point we construct a set of balls of growing radius centered in that point and we plot $\ln[N(x,R)]$ $\ln(R)$, [see (1.1)] so that from (1.2) it follows that the slope is the dimension at x. This program cannot be applied to all the cases studied, as some of them need special care. When N is small, problems of saturation can be present: as the radius grows, all the points enter into the ball, and only the first points in the plot can be taken into account [Fig. 2(b)]. Furthermore, when N is small ($N=12$), oftentimes the distance is trivial, i.e., $d_{ij} \simeq d = \text{const}$ $\forall i,j$, and obviously the procedure cannot be used because the slope cannot be determined within a reasonable accuracy [Fig. 2(c)]. On the other hand, when N is large ($N=96$) the statistics is clearly insufficient, so that the slope error can be very large.

In Fig. 3 we show the dimensions obtained at each point for several values of the parameters. After a first look at the dimensions plotted in each figure, one of the most striking features is the homogeneity in the dimension values at different points. The dimensions at every point in the toy world have the same order of magnitude, and for certain values of the parameters, almost all the dimension values are equal (within the error margins). This fact is probably due to the use of a purely quartic interaction in the potential. Future work should include the exploration of interaction terms of the type $g_p \mathrm{tr} M^{2p}$, for $p > 2$, as well as for $g_p < 0$.

Our numerical simulations do not provide a clear picture of the phase structure of the system. On general

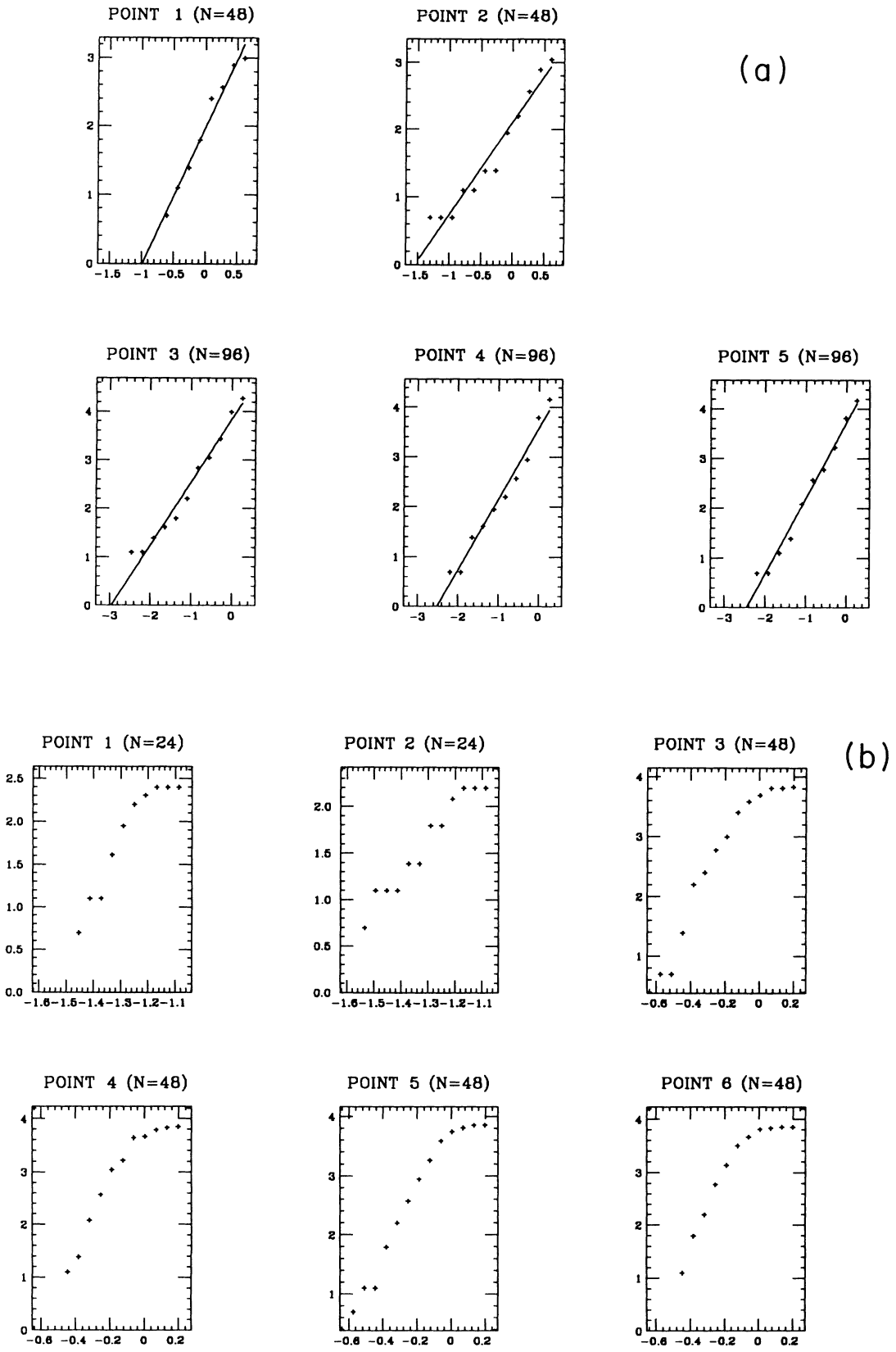

FIG. 2. Three types of fittings necessary to compute the local dimension, $n(x)$: In (a), we have plotted a straightforward, unambiguous case; (b) we have plotted some saturated examples; in (c) some examples of the discrete metric.

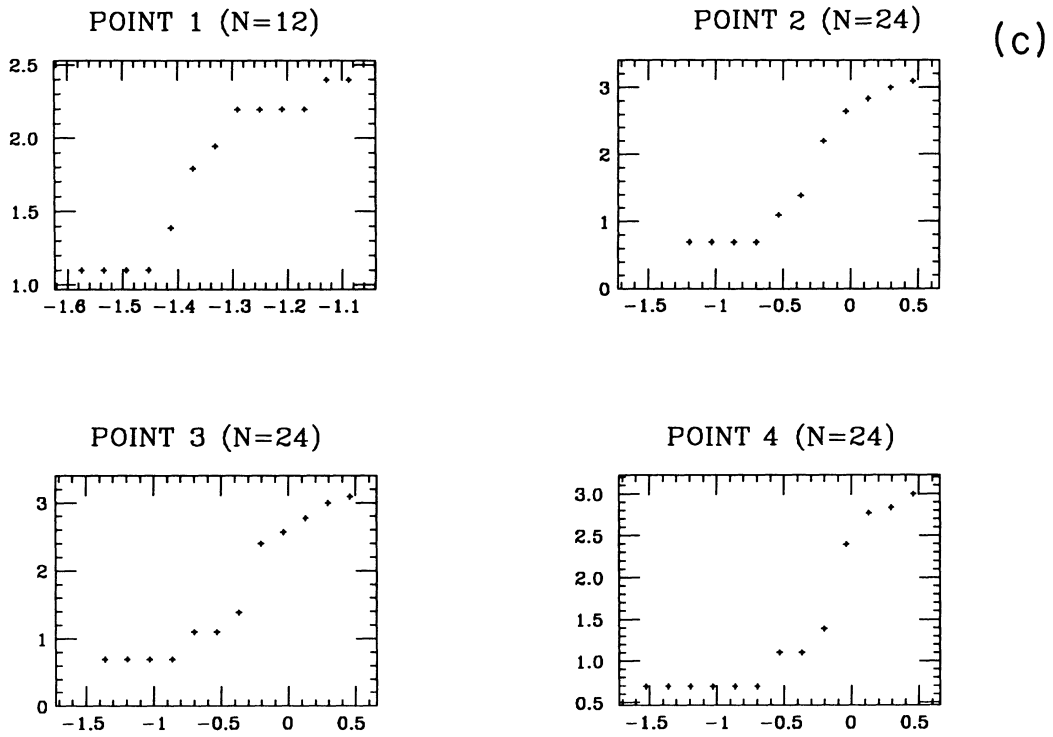

FIG. 2. (*Continued*).

grounds, we expect a phase transition to take place at a "length scale" $l \sim l_P$ between a "disordered phase," in which different pieces of spacetime will have different dimensions or no dimensions at all, and an "ordered phase," in which the dimension will be constant.

A Landau-Ginzburg description of this transition would be provided by an effective potential for the order parameter, $\phi \equiv \partial n(x)$, given by

$$V(\phi)=(l-l_P)/2\phi^2+\lambda\phi^4$$

which shows indeed a symmetry breaking between a phase of zero derivative for the dimension field, the ordered phase at long distances, and a broken phase, with constant nonzero values for the derivative of the dimension field, at short distances. (This phenomenological description is of course only valid for $l \sim l_P$.) What we mean is something like $\lim_{d\to 0}[n(x)-n(y)]/d(x,y)$, which would reduce to a directional derivative in the smooth case.

It is more difficult to find any mechanism to drive the dimension not only to a constant, but also to an integer. It is true, of course, that in the random dynamical models of Nielsen and Ninomiya (cf. [15]), in which they started from a non-Lorentz-invariant gauge theory (invariant, however, under translations), and computed what amounts essentially to the renormalization-group flow of the theory. The β functions worked in the sense of getting deviations from Lorentz invariance smaller at long distances; this means, of course, that even if the "fundamental" theory is non-Lorentz invariant, for a large class of theories, the effective low-energy action possesses the aforementioned symmetry. Actually, they argued subsequently that the assumption of gauge invariance was not essential, and that one should be able to recover an effective gauge invariance at long distances independently of the detailed form of the short-distance dynamics.

We can perhaps speculate on a possible generalization of their claims in the sense that there are more symmetry for integer values of the dimension than for noninteger ones. [We mean more symmetry in the sense that in the former case there is a Lie group with a natural action, such as O(N); whereas the symmetry acting on fractal sets is usually of the discrete type.]

IV. DYNAMICS OF METRIC SPACES

The toy models discussed in the two previous sections illustrate how one can dynamically describe in the quantum context metric spaces which do not behave as smooth manifolds. In this section we would like to be more specific and construct an action to characterize the transition between a discrete space (as spacetime could be before Planck's time) and a smooth space. One such action could be of the form

$$S = bS(\text{discrete}) + (1-b)S(\text{smooth}) , \tag{4.1}$$

where $b\,(0 \le b \le 1)$ is the parameter governing the transition.

In order to get some intuition of this action, the first thing to do is to obtain its explicit form in two extreme cases: the "discrete action" S(discrete) describing classically the "discrete metric," that is,

$$d_{ij} = 1-\delta_{ij} , \tag{4.2}$$

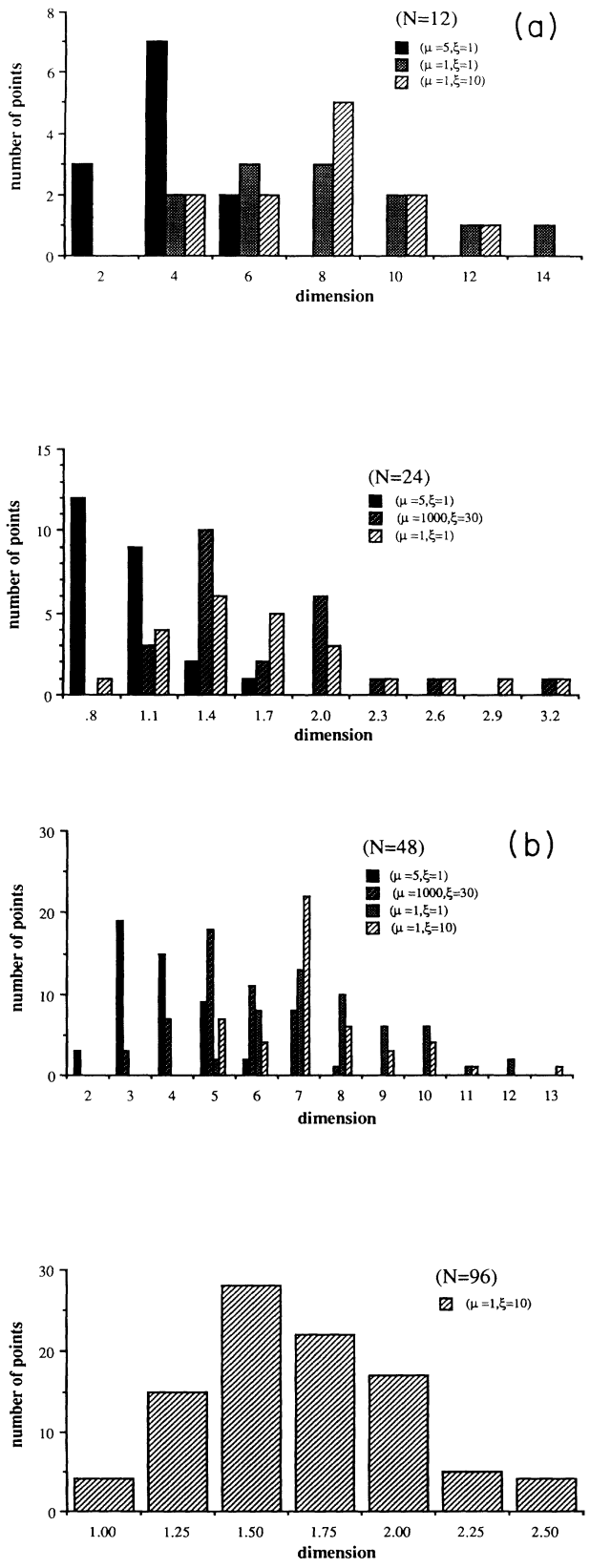

FIG. 3. We have plotted here the local dimension as a function of the point; for different values of N, and for different simulations. As in Fig. 1, $g_1=1$ and μ is the g_2 coefficient in the text.

and the "smooth action" S(smooth) describing the "smooth metric," that is, one which derives from a Riemannian distance through a formula of the type

$$d_{ij}=\int_i^j ds \ . \tag{4.3}$$

Spaces endowed with the discrete metric are in some sense the most dissimilar to R^n; for example, as we have already seen in Sec. II, every mapping of the space in itself is a continuous one. Discrete spaces are the easiest ones to characterize. In matrix form $d=E-1$, where E is the dyadic matrix: $E_{ij}=1\ \forall i,j$. E satisfies

$$E^p=N^{p-1}E \ ,$$

where N is the number of points of our space. (Note that this last formula does not imply $E=N\mathbb{1}$ because $\det E=0$.) But we can use it instead to derive a necessary condition for the metric d_{ij} to represent a discrete metric space, namely,

$$d^p=\left[\frac{(N-1)^p-(-1)^p}{N}\right]d + \frac{N-1}{N}[(N-1)^{p-1}+(-1)^p]\mathbb{1} \ . \tag{4.4}$$

It would seem that we have to impose an infinite number of equations (for $p=1,2,\ldots$), but actually all of them are consequences of the first one:

$$d^2=(N-2)d+(N-1)\mathbb{1} \tag{4.5}$$

which derives from an action

$$S=\mathrm{Tr}\left[\frac{d^3}{3}-\frac{N-2}{2}d^2-(N-1)d\right] \ . \tag{4.6}$$

But the discrete metric matrix (discrete matrix) is not the most general solution to Eq. (4.5), since we may construct a solution from any one, say D, by transforming it with an arbitrary orthogonal matrix O: O^TDO. This is a consequence of the invariance of Eq. (4.5) under similarity transformations. Thus in order to obtain the discrete matrix we must set some constraints to the solutions of (4.5) and modify the action (4.6) accordingly. The most obvious constraints is to set the diagonal elements to zero, since we want this matrix to satisfy the properties of a distance. Now, one can show easily that the most general solution of (4.5) with the constraint

$$d_{ii}=0, \quad i=1,\ldots,N \tag{4.7}$$

is the discrete matrix. In fact, let D be a solution, since it is symmetric it may be diagonalized by means of an orthogonal matrix O: $D=O^TMO$. Then Eq. (4.5) for the diagonal matrix reads

$$M^2-(N-2)M-(N-1)\mathbb{1}=0 \tag{4.8}$$

and therefore the eigenvalues of D, λ_i, satisfy $\lambda_i^2-(N-2)\lambda_i-(N-1)=0$, with solutions -1 and $N-1$. Since D has vanishing diagonal, it is traceless (and so is M). The traceless condition fixes the eigenvalues multiplicities, and we can choose those to be $\lambda_1=N-1$,

$\lambda_2 = \cdots = \lambda_N = -1$. Then the components of D, in terms of M, are

$$D_{ij} = \sum_{k,l} O_{ki} M_{kl} O_{lj} = -\delta_{ij} + N O_{1i} O_{1j} \ . \tag{4.9}$$

Now the constraints $D_{ii} = 0$ determine the O_{1i} coefficients to be $O_{1i} = 1/\sqrt{N}$, $i = 1, \ldots, N$ so that, finally, D is the discrete matrix d.

As a consequence, we have to implement in the action the constraints (4.7). This can be easily done, for example, by introducing N Lagrange multipliers, μ_i:

$$S(\text{discrete}) = \text{Tr}\left[\frac{d^3}{3} - \frac{(N-2)}{2} d^2 - (N-1)d\right] - \sum_i \mu_i d_{ii} \ . \tag{4.10}$$

Note that this action is not bounded from below due to the cubic term and this may cause problems in the path-integral quantization. However one may prove that the action,

$$S(\text{discrete}) = \text{Tr}\left[\frac{d^4}{4} - \frac{N^2-3N+3}{2} d^2 - (N-1)(N-2)d\right] - \sum_i \mu_i d_{ii} \tag{4.11}$$

leads to Eq. (4.4) for $p = 3$ and that the solution of such equation with the constraints (4.7) is the discrete metric, provided N is not divisible by 3.

The "smooth metric," on the other hand, is more difficult to characterize. In the simplest case of R^n, for example, the usual Euclidean distance

$$d(x,y) = \left[\sum_{i=1}^{N} (x^i - y^i)^2\right]^{1/2} \tag{4.12}$$

satisfies the equation

$$d \pounds_{\mathbf{u}}^2 d = 1 - (\pounds_{\mathbf{u}} d)^2 \ , \tag{4.13}$$

where the (directional) derivative is defined through

$$\pounds_{\mathbf{u}} d \equiv u^\mu \partial_\mu d \ . \tag{4.14}$$

A remarkable thing is that the former Eq. (4.13) is valid for an arbitrary vector $\mathbf{u}$. An obvious generalization of (4.14) for an arbitrary metric space is

$$\pounds_{zy} d(x,y) \equiv \lim_{d(z,y)\to 0} \frac{d(x,z) - d(x,y)}{d(z,y)} \ . \tag{4.15}$$

A natural question in this context is the following: to what extent does our posited Eq. (4.13) determine uniquely the "smooth" metric? (That is, the Euclidean metric in the example we are considering.) If we make a change of dependent variables, and use

$$\tilde{\Omega}(x^\mu, y^\nu) \equiv d^2(x^\mu, y^\nu) \tag{4.16}$$

then (4.13) implies that

$$\pounds_{\mathbf{u}}^2 \tilde{\Omega} = 2 \ . \tag{4.17}$$

The general solution of these equations can be easily found choosing $\mathbf{u}$ along the n independent directions:

$$\partial_i^2 \tilde{\Omega} = 2, \quad i = 1, \ldots, n \tag{4.18}$$

and noting that a particular solution for $\tilde{\Omega}$ is $\tilde{\Omega}_p = \sum_i x_i^2$, and the most general solution for the homogeneous equations is a polynomial of the x_i's, containing each variable at most once in each term (so that the second derivative vanishes). The smooth distance is exactly the particular solution, so we can throw away the homogeneous solution by imposing as boundary condition:

$$\lim_{r\to\infty} \left[\tilde{\Omega} - \sum_i x_i^2\right] = 0 \ . \tag{4.19}$$

A typical action which would posses (4.13) as its classical equation of motion and which we will take as our provisional prototype of an action characterizing "smooth spaces," is

$$S(\text{smooth}) \equiv \sum_{x,y} d_{x,y}^2 (\pounds_{zy} d_{x,y})^2 + d_{x,y}^2 \ . \tag{4.20}$$

Path-integral quantization with action (4.1) and the corresponding Liouville term to enforce triangular inequality in the numerical simulations is much more difficult to implement than in the toy models considered in Sec. III. This is due to the presence of the directional derivative terms in (4.20) and it will be the subject of further research.

Since the metric action (4.20) is supposed to fix the dynamics of a smooth space, i.e., the type of metric space one would expect to find after Planck's time, we would like to see how one might recover on some low-energy limit the Einstein-Hilbert action. Since we were forced to introduce the operator $\pounds$ to characterize the smooth distance we will first consider this operator and its connection with the usual directional derivative.

In an arbitrary metric space, there will in general be several inequivalent ways of defining the operator $\pounds$; they correspond to all the independent directions one can choose in the tangent space of a manifold at a given point. In the case our space is a linear one, i.e., when $d(x,y) = ||x - y||$ we can also define derivatives of functions defined on the metric space, $y = f(x)$ at the point $x = x_0$ as the unique linear mapping Df_{x_0} such that $g(x) = f(x_0) + Df_{x_0}(x - x_0)$ is tangent to $f(x)$ at $x = x_0$, i.e., such that $\lim_{x\to x_0} ||f(x) - g(x)|| / ||x - x_0|| = 0$. This definition of derivative reduces in the smooth manifold case, in a particular coordinate patch, to the usual Jacobian matrix: $Df = \partial_i f^j$, and is thus a matrix with n rows and m columns. Directional derivatives are obviously additive for linear spaces. It can happen, in particular, that there are exactly n independent ways of defining the directional derivative. Let us call them $\nabla_a d_{x,y}; a = 1, \ldots, n$. (They would be equal to $e_a^\mu \partial_\mu d_{x,y}$ in a smooth manifold, where e_a^μ, $a = 1, \ldots, n$ are the n vectors of a "vielbein.") In the standard framework of Einstein's general relativity, the square of the geodesic distance Ω (Synge's world function) enjoys several useful

properties in the limit $x \to y$:

$$\lim_{x\to y}\Omega=0, \quad \lim_{x\to y}\partial_\mu\Omega=0 ,$$
$$\lim_{x\to y}\partial_\mu\partial_\nu\Omega=g_{\mu\nu}, \quad \lim_{x\to y}\partial_\mu\partial_\nu\partial_\rho\Omega=0 , \tag{4.21}$$
$$\lim_{x\to y}\partial_\mu\partial_\nu\partial_\rho\partial_\sigma\Omega=\tfrac{3}{2}(R_{\sigma\nu\rho\mu}+R_{\sigma\mu\rho\nu}) .$$

This means that the world function can be used as an alternative variable to the usual metric (or vierbein), and actually, an action completely equivalent to the usual Einstein-Hilbert one is

$$S=\int d(\mathrm{vol})\lim_{x\to y}(-3/2\kappa^2)(\partial^\mu\partial^\rho\Omega)(\partial^\nu\partial^\sigma\Omega)\partial_\mu\partial_\nu\partial_\rho\partial_\sigma\Omega , \tag{4.22}$$

where the inverse matrix of second derivatives is defined through

$$(\partial^\mu\partial^\nu\Omega)\partial_\nu\partial_\sigma\Omega=\delta^\mu_\sigma .$$

We can now be more precise about the long-distance ("low-energy") limit of our theory. Our point of view will be most conservative: we shall assume, that our theory behaves like an ordinary quantum field theory, which means in particular that there should exist an effective low-energy Lagrangian obtained from the fundamental one by keeping local terms, and performing an expansion in derivatives around them. In the situation where this hypothesis holds true, we are guaranteed that the lowest-dimension local scalar is the trace of the Ricci tensor: that is, the Einstein-Hilbert term will always dominate in the regime (cf. Ref. [7]). But in order to recover general relativity, we need to implement its most important symmetry as well, namely, general covariance. This means that a good guiding principle for constructing candidate actions in our generalized theory of metric spaces should be the independence on the basis chosen in the space of linearly independent directional derivatives. (Of course this restriction is very difficult to implement in general, and has a clear meaning only in the case in which this space has integer dimension $n \geq 1$.)

V. CONCLUSIONS

We have presented in this paper a model of the spacetime which one might expect physically to work not too far above Planck's mass.

The intuitive picture one gets from our results is a topological space, with some pieces in it which resemble locally some R^n (that is, quasismooth regions, in which semiclassical physics should work well); some transition regions, from one value of the local dimension, n, to another, which are much more difficult to describe in a semiclassical way; and, finally, some "quantum nonsmooth" pieces, which do not resemble at all R^n. Our framework is able to embrace all these regions in an overall coherent scheme.

We have not yet discovered, however, what the fundamental symmetry principle of this theory should be. In other words, we lack the principles of general covariance and low-energy expansions of effective actions in terms of the number of derivatives (because we do not have such things as coordinates in our framework). The models studied so far should then be considered as a preliminary exploration of the whole approach.

A point worth making is that several topologically equivalent distances give different values for the local dimension. For example, if we define $d'(x,y)\equiv d(x,y)^k$, where $0<k<1$, then, as is easily checked, (H,d) and (H,d') (where H is the set of spacetime points) are equivalent topologically, and yet the local dimension changes in $1/k$. Although our explicit models are not invariant under those transformations, it is tempting to speculate that this is the symmetry used to rescale the dimension to an integer, once the phase of constant dimension is reached.

On a more speculative note, we would like to close by stressing a fascinating analogy with the celebrated work of the French mathematician Connes (cf. Ref. [16]) on the application of his own theory of noncommutative geometry to physics. He starts with some abstract algebra, and defines the physical spacetime as a derived object, reconstructed from the spectrum of operators defined on that algebra; essentially, as far as we understand, the Dirac operator. He actually defines it in an abstract way as the generator of the K-cohomology, and using heavily a very special notion of trace, namely, the Dixmier trace, which gives zero on trace class operators, he is able to relate the logarithmic divergences present in usual renormalizable quantum field theories, with certain nontrivial homology classes.

Our analytic models in Sec. II are essentially a physicist's implementation of the above ideas. We also started with an algebra (which we identified with the algebra of ordinary matrices), and defined spacetime as a derived concept, either using the spectrum of the matrices, or the entries themselves. Our general theories, on the other hand, have been constructed through a limiting process. We did not try to construct them directly. It is obvious that the overall approach is quite similar, and we would not be surprised if some deeper interconnection between our work and noncommutative geometry is to be found in future work; we believe, in particular, that the "continuum theory" corresponding to our discrete limiting procedures will imply some models of noncommutative geometry.

Given the fact that there seems to exist some deep relations between Connes' work and twistor theory as well, it could be that all conceivable generalizations of geometry at very short distances are intimately related.

ACKNOWLEDGMENTS

We are indebted to Orlando Alvarez, Luis Alvarez Gaumé, Marià Baig, Joaquín Díaz Alonso, Domènec Espriu, César Gómez, Pedro Gosdzinsky, and Holger Nielsen for many useful (and sometimes lively) discussions. E.V. acknowledges support from DGICT (Spain) and hospitality at the Enrico Fermi Institute of the University of Chicago. This work has been partially supported by CICYT (Spain).

[1] E. Alvarez, Rev. Mod. Phys. **61**, 561 (1989).
[2] G. Veneziano, Report No. CERN TH-5581, 1989 (unpublished); D. J. Gross, Princeton Report No. PUPT-1108, 1988 (unpublished).
[3] R. Penrose and W. Rindler, *Spinors and Space Time* (Cambridge University Press, Cambridge, England, 1989).
[4] L. Bombelli, J. Lee, D. Meyer, and R. D. Sorkin, Phys. Rev. Lett. **59**, 521 (1987); G. Brightwell and R. Gregory, *ibid.* **66**, 260 (1991).
[5] D. Finkelstein, Int. J. Theor. Phys. **28**, 441 (1988).
[6] C. J. Isham, Imperial College Report No. Imperial/TP/88-89/16, 1989 (unpublished); Class. Quantum Grav. **6**, 1509 (1989); C. J. Isham, Y. Kubyshin, and P. Renteln, *ibid.* **7**, 1053 (1990).
[7] E. Alvarez, Phys. Lett. B **210**, 73 (1988).
[8] J. L. Synge, *Relativity: The General Theory* (North-Holland, Amsterdam, 1960); B. S. DeWitt, *Dynamical Theory of Groups and Fields* (Gordon and Breach, New York, 1965).
[9] J. P. Eckmann and D. Ruelle, Rev. Mod. Phys. **57**, 617 (1985).
[10] G. Ellis and B. G. Schmidt, Gen. Relativ. Gravit. **8**, 915 (1977).
[11] E. Alvarez, J. Cespedes, and E. Verdaguer, in *ERE-90,* edited by C. Bona, J. Carot, and L. Mas, (World Scientific, Singapore, 1991).
[12] M. L. Mehta, *Random Matrices and the Statistical Theory of Energy Levels* (Academic, New York, 1967); S. Chadha, G. Mahoux, and M. L. Mehta, J. Phys. A **14**, 579 (1981); M. L. Mehta, Commun. Math. Phys. **79**, 327 (1981).
[13] E. Brezin, C. Itzykson, G. Parisi, and J. B. Zuber, Commun. Math. Phys. **59**, 35 (1978); C. Itzykson and J. B. Zuber, J. Math. Phys. **21**, 411 (1980); D. Bessis, C. Itzykson, and J. B. Zuber, Adv. Appl. Math. **1**, 109 (1980).
[14] M. Creutz, *Quarks, Gluons and Lattices* (Cambridge University Press, Cambridge, England, 1985).
[15] H. B. Nielsen and M. Ninomiya, Nucl. Phys. **B141**, 153 (1978).
[16] A. Connes, Report No. IHES/M/89/69 (unpublished); *Geometrie Non-conmutative* (Intereditions, Paris, 1989).

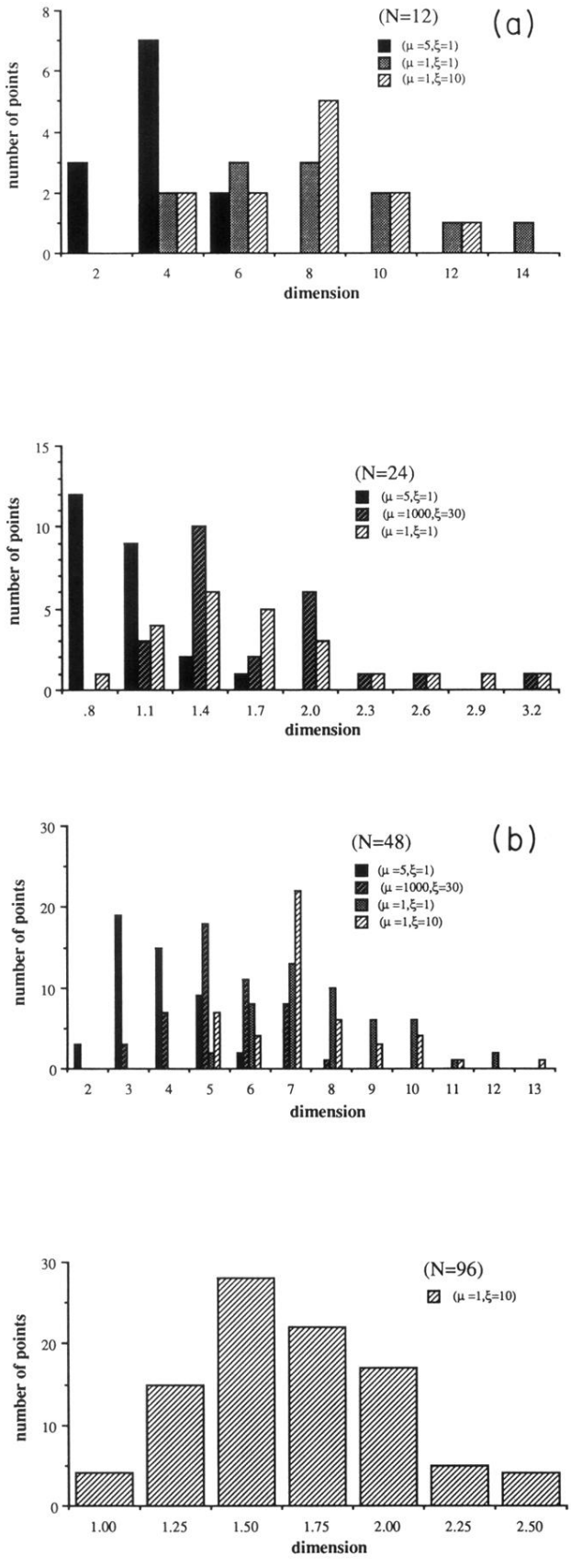

FIG. 3. We have plotted here the local dimension as a function of the point; for different values of N, and for different simulations. As in Fig. 1, $g_1 = 1$ and μ is the g_2 coefficient in the text.