REVIEWS OF MODERN PHYSICS, VOLUME 73, OCTOBER 2001

Noncommutative field theory

Michael R. Douglas*

Department of Physics and Astronomy, Rutgers University, Piscataway, New Jersey 08855
and Institut des Hautes Etudes Scientifiques, 35 route des Chartres, 91440 Bures-sur-Yvette, France

Nikita A. Nekrasov

Institut des Hautes Etudes Scientifiques, 35 route des Chartres, 91440 Bures-sur-Yvette, France
and Institute for Theoretical and Experimental Physics, 117259 Moscow, Russia

(Published 29 November 2001)

This article reviews the generalization of field theory to space-time with noncommuting coordinates, starting with the basics and covering most of the active directions of research. Such theories are now known to emerge from limits of M theory and string theory and to describe quantum Hall states. In the last few years they have been studied intensively, and many qualitatively new phenomena have been discovered, on both the classical and the quantum level.

CONTENTS

I. INTRODUCTION

Noncommutativity is an age-old theme in mathematics and physics. The noncommutativity of spatial rotations in three and more dimensions is deeply ingrained in us. Noncommutativity is the central mathematical concept expressing uncertainty in quantum mechanics, where it applies to any pair of conjugate variables, such as position and momentum. In the presence of a magnetic field, even momenta fail to mutually commute.

*Electronic address: mrd@physics.rutgers.edu

One can just as easily imagine that position measurements might fail to commute and describe this using noncommutativity of the coordinates. The simplest noncommutativity one can postulate is the commutation relation

$$[x^i, x^j] = i\theta^{ij}, \tag{1}$$

with a parameter θ which is an antisymmetric (constant) tensor of dimension $(\text{length})^2$.

As has been realized independently many times, at least as early as 1947 (Snyder, 1947), there is a simple modification to quantum field theory obtained by taking the position coordinates to be noncommuting variables. Starting with a conventional field theory Lagrangian and interpreting the fields as depending on coordinates satisfying Eq. (1), one can follow the usual development of perturbative quantum field theory with surprisingly few changes, to define a large class of "noncommutative field theories."

It is on this class of theories that our review will focus. Until recently, such theories had not been studied very seriously. Perhaps the main reason for this is that postulating an uncertainty relation between position measurements will *a priori* lead to a nonlocal theory, with all of the attendant difficulties. A secondary reason is that noncommutativity of the space-time coordinates generally conflicts with Lorentz invariance, as is apparent in Eq. (1). Although it is not implausible that a theory defined using such coordinates could be effectively local on length scales longer than that of θ, it is harder to believe that the breaking of Lorentz invariance would be unobservable at these scales.

Nevertheless, one might postulate noncommutativity for a number of reasons. Perhaps the simplest is that it might improve the renormalizability properties of a theory at short distances or even render it finite. Without giving away too much of our story, we should say that this is of course not obvious *a priori* and a noncommutative theory might turn out to have the same or even worse short-distance behavior than a conventional theory.

Another motivation is the long-held belief that in quantum theories including gravity, space-time must change its nature at distances comparable to the Planck scale. Quantum gravity has an uncertainty principle which prevents one from measuring positions to better accuracies than the Planck length: the momentum and energy required to make such a measurement will itself modify the geometry at these scales (DeWitt, 1962). One might wonder if these effects could be modeled by a commutation relation such as Eq. (1).

A related motivation is that there are reasons to believe that any theory of quantum gravity will not be local in the conventional sense. Nonlocality brings with it deep conceptual and practical issues which have not been well understood, and one might want to understand them in the simplest examples first, before proceeding to a more realistic theory of quantum gravity.

This is one of the main motivations for the intense current activity in this area among string theorists. String theory is not local in any sense we now understand, and indeed has more than one parameter characterizing this nonlocality: in general, it is controlled by the larger of the Planck length and the *string length*, the average size of a string. It was discovered by Connes *et al.* (1998) and Douglas and Hull (1998) that simple limits of M theory and string theory lead directly to noncommutative gauge theories, which appear far simpler than the original string theory yet keep some of this nonlocality.

One might also study noncommutative theories as interesting analogs of theories of more direct interest, such as Yang-Mills theory. An important point in this regard is that many theories of interest in particle physics are so highly constrained that they are difficult to study. For example, pure Yang-Mills theory with a definite simple gauge group has no dimensionless parameters with which to make a perturbative expansion or otherwise simplify the analysis. From this point of view it is quite interesting to find any sensible and nontrivial variants of these theories.

Now, physicists have constructed many variations of Yang-Mills theory in the search for regulated (UV finite) versions as well as more tractable analogs of the theory. A particularly interesting example in the present context is the twisted Eguchi-Kawai model (Eguchi and Nakayama, 1983; Gonzalez-Arroyo and Okawa, 1983), which in some of its forms, especially that of Gonzalez-Arroyo and Korthals Altes (1983), is a noncommutative gauge theory. This model was developed in the study of the large-N limit of Yang-Mills theory ('t Hooft, 1974) and we shall see that noncommutative gauge theories show many analogies to this limit (Filk, 1996; Minwalla *et al.*, 2000), suggesting that they should play an important role in the circle of ideas relating large-N gauge theory and string theory (Polyakov, 1987; Aharony *et al.*, 2000).

Noncommutative field theory is also known to appear naturally in condensed-matter theory. The classic example (though not always discussed using this language) is the theory of electrons in a magnetic field projected to the lowest Landau level, which is naturally thought of as a noncommutative field theory. Thus these ideas are relevant to the theory of the quantum Hall effect (Prange and Girvin, 1987), and indeed, noncommutative geometry has been found very useful in this context (Bellissard *et al.*, 1993). Most of this work has treated noninteracting electrons, and it seems likely that introducing field-theoretic ideas could lead to further progress.

It is interesting to note that despite the many physical motivations and partial discoveries we just recalled, noncommutative field theory and gauge theory were first clearly formulated by mathematicians (Connes and Rieffel, 1987). This is rather unusual for a theory of significant interest to physicists; usually, as with Yang-Mills theory, the flow goes in the other direction.

An explanation for this course of events might be found in the deep reluctance of physicists to regard a nonlocal theory as having any useful space-time interpretation. Thus, even when these theories arose naturally in physical considerations, they tended to be regarded only as approximations to more conventional local theories, and not as ends in themselves. Of course such sociological questions rarely have such pat answers and we shall not pursue this one further except to remark that, in our opinion, the mathematical study of these theories and their connection to noncommutative geometry has played an essential role in convincing physicists that these are not arbitrary variations on conventional field theory but indeed a new universality class of theory deserving study in its own right. Of course this mathematical work has also been an important aid to the more prosaic task of sorting out the possibilities, and it is the source for many useful techniques and constructions that we shall discuss in detail.

Having said this, it seems that the present trend is that the mathematical aspects appear less and less central to the physical considerations as time goes on. While it is too early to judge the outcome of this trend and it seems certain that the aspects which traditionally have benefited most from mathematical influence will continue to do so (especially the topology of gauge-field configurations, and techniques for finding exact solutions), we have to some extent deemphasized the connections with noncommutative geometry in this review. This is partly to make the material accessible to a wider class of physicists, and partly because many excellent books and reviews cover the material from this point of view, starting with that of Connes (1994), and including those of Nekrasov (2000) focusing on classical solutions of noncommutative gauge theory, Konechny and Schwarz (2000b) focusing on duality properties of gauge theory on a torus, and Gracia-Bondia *et al.* (2001) and Varilly (1997). We maintain one section which attempts to give an overview of aspects for which a more mathematical point of view is clearly essential.

Although many of the topics we discuss were motivated by and discovered in the context of string theory, we have also taken the rather unconventional approach of separating the discussion of noncommutative field theory from that of its relation to string theory, to the extent that this was possible. An argument against this approach is that the relation clarifies many aspects of the theory, as we hope will become abundantly clear upon reading Sec. VII. However, it is also true that string theory is not a logical prerequisite for studying the theory, and we feel the approach we took better illustrates its internal self-consistency (and the points where this is still lacking). Furthermore, if we hope to use noncommutative field theory as a source of *new* insights into string theory, we need to be able to understand its physics without relying too heavily on the analogy. We also hope this approach will have the virtue of broader accessibility and perhaps help in finding interesting applications outside of string theory. Reviews with a more string-theoretic emphasis include that of Harvey (2001a) which discusses solitonic solutions and their relations to string theory.

Finally, we must apologize to the many whose work we were not able to treat in the depth it deserved in this review, a sin we have tried to atone for by including an extensive bibliography.

II. KINEMATICS

A. Formal considerations

Let us start by defining noncommutative field theory in a somewhat pedestrian way, by proposing a configuration space and action functional from which we could either derive equations of motion or define a functional integral. We shall discuss this material from a more mathematical point of view in Sec. VI.

Conventions. Throughout the review we use the following notations: Latin indices $i,j,k,\ldots$ denote spacetime indices, Latin indices from the beginning of the alphabet $a,b,\ldots$ denote commutative dimensions, Greek indices $\mu,\nu,\ldots$ enumerate particles, vertex operators, etc., while Greek indices from the beginning of the alphabet $\alpha,\beta,\ldots$ denote noncommutative directions.

In contexts where we simultaneously discuss a noncommuting variable or field and its commuting analog, we shall use the "hat" notation: x is the commuting analog to $\hat{x}$. However, in other contexts, we shall not use the hat.

1. The algebra

The primary ingredient in the definition is an associative but not necessarily commutative algebra, to be denoted $\mathcal{A}$. The product of elements a and b of $\mathcal{A}$ will be denoted ab, $a\cdot b$, or $a\star b$. This last notation (the *star product*) has a special connotation, to be discussed shortly.

An element of this algebra will correspond to a configuration of a classical complex scalar field on a *space* M. Suppose first that $\mathcal{A}$ is commutative. The primary example of a commutative associative algebra is the algebra of complex-valued functions on a manifold M, with addition and multiplication defined pointwise: $(f+g)(x)=f(x)+g(x)$ and $(f\cdot g)(x)=f(x)g(x)$. In this case, our definitions will reduce to the standard ones for field theory on M.

Although the mathematical literature is usually quite precise about the class of functions (continuous, smooth, etc.) to be considered, in this review we follow standard physical practice and simply consider all functions that arise in reasonable physical considerations, referring to this algebra as $\mathcal{A}(M)$ or (for reasons to be explained shortly) as M_0. If more precision is wanted, for most purposes one can think of this as $C(M)$, the bounded continuous functions on the topological manifold M.

The most elementary example of a noncommutative algebra is Mat_n, the algebra of complex $n\times n$ matrices. Generalizations of this, which are almost as elementary, are the algebras $\mathrm{Mat}_n[C(M)]$ of $n\times n$ matrices whose matrix elements are elements of $C(M)$, and with addition and multiplication defined according to the usual rules for matrices in terms of the addition and multiplication on $C(M)$. This algebra contains $C(M)$ as its center (take functions times the identity matrix in Mat_n).

Clearly elements of $\mathrm{Mat}_n[C(M)]$ correspond to configurations of a matrix field theory. Just as one can gain some intuition about operators in quantum mechanics by thinking of them as matrices, this example already serves to illustrate many of the formal features of noncommutative field theory. In the remainder of this subsection we introduce the other ingredients we need to define noncommutative field theory in this familiar context.

To define a real-valued scalar field, it is best to start with $\mathrm{Mat}_n[C(M)]$ and then impose a reality condition analogous to the reality of functions in $C(M)$. The most useful in practice is to take the Hermitian matrices $a=a^\dagger$, whose eigenvalues will be real (given suitable additional hypotheses). To do this for general $\mathcal{A}$, we would need an operation $a\to a^\dagger$ satisfying $(a^\dagger)^\dagger=a$ and (for $c\in\mathbb{C}$) $(ca)^\dagger=c^*a^\dagger$, in other words an antiholomorphic involution.

The algebra $\mathrm{Mat}_n[C(M)]$ could also be defined as the tensor product $\mathrm{Mat}_n(\mathbb{C})\otimes C(M)$. This construction generalizes to an arbitrary algebra $\mathcal{A}$ to define $\mathrm{Mat}_n(\mathbb{C})\otimes\mathcal{A}$, which is just $\mathrm{Mat}_n(\mathcal{A})$ or $n\times n$ matrices with elements in $\mathcal{A}$. This algebra admits the automorphism group $GL(n,\mathbb{C})$, acting as $a\to g^{-1}ag$ (of course the center acts trivially). Its subgroup $U(n)$ preserves Hermitian conjugation and the reality condition $a=a^\dagger$. One sometimes refers to these as $U(n)$ noncommutative theories, a bit confusingly. We shall refer to them as *rank n theories*.

In the rest of the review, we shall mostly consider noncommutative associative algebras which are related to the algebras $\mathcal{A}(M)$ by deformation with respect to a parameter θ, as we shall define shortly. Such a deformed algebra will be denoted by M_θ, so that $M_0=\mathcal{A}(M)$.

2. The derivative and integral

A noncommutative field theory will be defined by an action functional of fields $\Phi,\phi,\varphi,\ldots$ defined in terms of the associative algebra $\mathcal{A}$ (it could be elements of $\mathcal{A}$, or vectors in some representation thereof). Besides the algebra structure, to write an action we shall need an integral $\int\mathrm{Tr}$ and derivatives ∂_i. These are linear operations satisfying certain formal properties:

(a) The derivative is a derivation on $\mathcal{A}$, $\partial_i(AB)=(\partial_i A)B+A(\partial_i B)$. With linearity, this implies that the derivative of a constant is zero.

(b) The integral of the trace of a total derivative is zero, $\int\mathrm{Tr}\partial_i A=0$.

(c) The integral of the trace of a commutator is zero, $\int\mathrm{Tr}[A,B]\equiv\int\mathrm{Tr}(A\cdot B-B\cdot A)=0$.

A candidate derivative ∂_i can be written using an element $d_i\in\mathcal{A}$; let $\partial_i A=[d_i,A]$. Derivations that can be written in this way are referred to as inner derivations, while those that cannot are outer derivations.

We denote the integral as $\int\mathrm{Tr}$, as it turns out that, for general noncommutative algebras, one cannot separate the notations of trace and integral. Indeed, one normally uses either the single symbol Tr (as is done in mathematics) or $\int$ to denote this combination; we do not follow this convention here only to aid the uninitiated.

We note that just as condition (b) can be violated in conventional field theory for functions that do not fall off at infinity, leading to boundary terms, condition (c) can be violated for general operators, leading to physical consequences in noncommutative theory which we will discuss.

B. Noncommutative flat space-time

After $\mathrm{Mat}_n[C(M)]$, the next simplest example of a noncommutative space is the one associated with the algebra $\mathbb{R}^d_\theta$ of all complex linear combinations of products of d variables $\hat{x}^i$ satisfying

$$[x^i,x^j]=i\theta^{ij}. \tag{2}$$

The i is present because the commutator of Hermitian operators is anti-Hermitian. As in quantum mechanics, this expression is the natural operator analog of the Poisson bracket determined by the tensor θ^{ij}, the *Poisson tensor* or noncommutativity parameter.

By applying a linear transformation to the coordinates, one can bring the Poisson tensor to canonical form. This form depends only on its rank, which we denote as $2r$. We keep this general as one often discusses partially noncommutative spaces, with $2r<d$.

A simple set of derivatives ∂_i can be defined by the relations

$$\partial_i x^j\equiv\delta_i^j, \tag{3}$$

$$[\partial_i,\partial_j]=0, \tag{4}$$

and the Leibnitz rule. This choice also determines the integral uniquely (up to overall normalization), by requiring that $\int\partial_i f=0$ for any f such that $\partial_i f\neq 0$.

We shall occasionally generalize Eq. (4) to

$$[\partial_i,\partial_j]=-i\Phi_{ij}, \tag{5}$$

to incorporate an additional background magnetic field.

Finally, we shall require a metric, which we shall take to be a constant symmetric tensor g_{ij}, satisfying $\partial_i g_{jk}=0$. In many examples we take this to be $g_{ij}=\delta_{ij}$, but note that one cannot bring both g_{ij} and θ^{ij} to canonical form simultaneously, as the symmetry groups preserved by the two structures, $O(n)$ and $Sp(2r)$, are different. At best one can bring the metric and the Poisson tensor to the following form:

$$g=\sum_{\alpha=1}^{r} dz_\alpha d\bar{z}_\alpha+\sum_b dy_b^2,$$

$$\theta=\frac{1}{2}\sum_\alpha \theta_\alpha\partial_{\bar{z}_\alpha}\wedge\partial_{z_\alpha},\quad \theta_a>0. \tag{6}$$

Here $z_\alpha=q_\alpha+ip_\alpha$ are convenient complex coordinates. In terms of p,q,y the metric and the commutation relations Eq. (6) read as

$$[y_a,y_b]=[y_b,q_\alpha]=[y_b,p_\alpha]=0,$$

$$[q_\alpha,p_\beta]=i\theta_\alpha\delta_{\alpha\beta},\quad ds^2=dq_\alpha^2+dp_\alpha^2+dy_b^2. \tag{7}$$

1. Symmetries of $\mathbb{R}^d_\theta$

An infinitesimal translation $x^i \to x^i + a^i$ on $\mathbb{R}^d_\theta$ acts on functions as $\delta\phi = a^i \partial_i \phi$. For the noncommuting coordinates x^i, these are formally inner derivations, as

$$\partial_i f = [-i(\theta^{-1})_{ij} x^j, f]. \tag{8}$$

One obtains global translations by exponentiating these,

$$f(x^i + \varepsilon^i) = e^{-i\theta_{ij}\varepsilon^i x^j} f(x) e^{i\theta_{ij}\varepsilon^i x^j}. \tag{9}$$

In commutative field theory, one draws a sharp distinction between translation symmetries (involving the derivatives) and internal symmetries, such as $\delta\phi = [A, \phi]$. We see that in noncommutative field theory, there is no such clear distinction, and this is why one cannot separately define integral and trace.

One often uses only $[\partial_i, f]$ and if so, Eq. (8) can be simplified further to the operator substitution $\partial_i \to -i(\theta^{-1})_{ij} x^j$. This leads to derivatives satisfying Eq. (5) with $\Phi_{ij} = -(\theta^{-1})_{ij}$.

The $Sp(2r)$ subgroup of the rotational symmetry $x^i \to R^i_j x^j$ which preserves θ, $R^{i'}_i R^{j'}_j \theta_{i'j'} = \theta_{ij}$ can be obtained similarly, as

$$f(R^i_j x^j) = e^{-iA_{ij}x^i x^j} f(x) e^{iA_{ij}x^i x^j}, \tag{10}$$

where $R = e^{iL}$, $L^i_j = A_{kj}\theta^{ik}$, and $A_{ij} = A_{ji}$. Of course only the $U(r)$ subgroup of this will preserve the Euclidean metric.

After considering these symmetries, we might be tempted to go on and conjecture that

$$\delta\phi = i[\phi, \epsilon] \tag{11}$$

for any ϵ is a symmetry of $\mathbb{R}^d_\theta$. However, although these transformations preserve the algebra structure and the trace,[1] they do not preserve the derivatives. Nevertheless, they are important and will be discussed in detail below.

2. Plane-wave basis and dipole picture

One can introduce several useful bases for the algebra $\mathbb{R}^d_\theta$. For discussions of perturbation theory and scattering, the most useful basis is the plane-wave basis, which consists of eigenfunctions of the derivatives:

$$\partial_i e^{ikx} = i k_i e^{ikx}. \tag{12}$$

The solution e^{ikx} of this linear differential equation is the exponential of the operator $ik\cdot x$ in the usual operator sense.

The integral can be defined in this basis as

$$\int \mathrm{Tr} e^{ikx} = \delta_{k,0}, \tag{13}$$

where we interpret the delta function in the usual physical way (for example, its value at zero represents the volume of physical space).

[1]Assuming certain conditions on ϵ and ϕ; see Sec. VI.A.

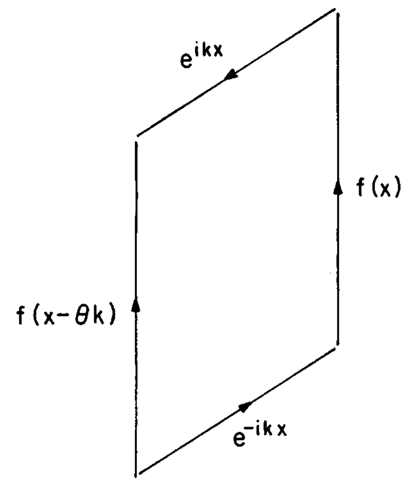

FIG. 1. The interaction of two dipoles.

More interesting is the interpretation of the multiplication law in this basis. This is easy to compute in the plane-wave basis, by operator reordering:

$$e^{ikx} \cdot e^{ik'x} = e^{-(i/2)\,\theta^{ij}k_i k'_j} e^{i(k+k')\cdot x}. \tag{14}$$

The combination $\theta^{ij}k_i k'_j$ appearing in the exponent comes up very frequently, and a standard and convenient notation for it is

$$k \times k' \equiv \theta^{ij} k_i k'_j = k \times_\theta k',$$

the latter notation being used to stress the choice of Poisson structure.

We can also consider

$$e^{ikx} \cdot f(x) \cdot e^{-ikx} = e^{-\theta^{ij}k_i\partial_j} f(x) = f(x^i - \theta^{ij}k_j). \tag{15}$$

Multiplication by a plane wave translates a general function by $x^i \to x^i - \theta^{ij}k_j$. This exhibits the nonlocality of the theory in a particularly simple way and gives rise to the principle that large momenta will lead to large nonlocality.

A simple picture can be made of this nonlocality (Sheikh-Jabbari, 1999; Bigatti and Susskind, 2000) by imagining that a plane wave corresponds not to a particle (as in commutative quantum field theory) but instead to a "dipole," a rigid oriented rod whose extent is proportional to its momentum:

$$\Delta x^i = \theta^{ij} p_j. \tag{16}$$

If we postulate that dipoles interact by joining at their ends (Fig. 1), and grant the usual quantum field theory relation $p = \hbar k$ between wave number and momentum, the rule Eq. (15) follows immediately.

3. Deformation, operators, and symbols

There is a sense in which $\mathbb{R}^d_\theta$ and the commutative algebra of functions $C(\mathbb{R}^d)$ have the same topology and the same size, notions we shall keep at an intuitive level. In the physical applications, it will turn out that θ is typically a controllable parameter, which one can imagine increasing from zero to go from commutative to noncommutative (this does not imply that the physics is continuous in this parameter, however). These are all reasons to study the relation between these two algebras more systematically.

There are a number of ways to think about this relation. If θ is a physical parameter, it is natural to think of $\mathbb{R}^d_\theta$ as a deformation of $\mathbb{R}^d$. A deformation M_θ of $C(M)$ is an algebra with the same elements and addition law (it is the same considered as a vector space) but a different multiplication law, which reduces to that of $C(M)$ as a (multi)parameter θ goes to zero. This notion was introduced by Bayen *et al.* (1978) as an approach to quantization, and has been much studied since, as we shall discuss in Sec. VI. Such a deformed multiplication law is often denoted $f\star g$ or *star product* to distinguish it from the original pointwise multiplication of functions.

This notation has a second virtue, which is that it allows us to work with M_θ in a way that is somewhat more forgiving of ordering questions. Namely, we can choose a linear map S from M_θ to $C(M)$, $\hat{f}\mapsto S[\hat{f}]$, called the *symbol* of the operator. We then represent the original operator multiplication in terms of the star product of symbols as

$$\hat{f}\hat{g}=S^{-1}[S[\hat{f}]\star S[\hat{g}]]. \tag{17}$$

One should recall that the symbol is not "natural" in the mathematical sense: there could be many valid definitions of S, corresponding to different choices of operator ordering prescription for S^{-1}.

A convenient and standard choice is the Weyl ordered symbol. The map S, defined as a map taking elements of $\mathbb{R}^d_\theta$ to $\mathcal{A}(\mathbb{R}^d)$ (functions on momentum space), and its inverse, are

$$f(k)\equiv S[\hat{f}](k)=\frac{1}{(2\pi)^{n/2}}\int \mathrm{Tr}e^{-ik\hat{x}}\hat{f}(\hat{x}), \tag{18}$$

$$\hat{f}(\hat{x})=S^{-1}[f]=\frac{1}{(2\pi)^{n/2}}\int d^nk\, e^{ik\hat{x}}f(k). \tag{19}$$

Formally these are inverse Fourier transforms, but the first expression involves the integral equation (13) on $\mathcal{A}(\mathbb{R}_\theta)$, while the second is an ordinary momentum-space integral.

One can get the symbol in position space by performing a second Fourier transform; e.g.,

$$S[\hat{f}](x)=\frac{1}{(2\pi)^n}\int d^nk\int \mathrm{Tr}e^{ik(x-\hat{x})}\hat{f}(\hat{x}). \tag{20}$$

We shall freely assume the usual Fourier relation between position and momentum space for the symbols, while being careful to say (or denote by standard letters such as x and k) which we are using.

The star product for these symbols is

$$e^{ikx}\star e^{ik'x}=e^{-(i/2)\,\theta^{ij}k_ik'_j}e^{i(k+k')\cdot x}. \tag{21}$$

Of course all of the discussion in Sec. II.B.2 above still applies, as this is only a different notation for the same product, Eq. (14).

Another special case that often comes up is

$$\int \mathrm{Tr}f\star g=\int \mathrm{Tr}fg. \tag{22}$$

4. The noncommutative torus

Much of this discussion applies with only minor changes to define $\mathbf{T}^d_\theta$, the algebra of functions on a noncommutative torus.

To obtain functions on a torus from functions on $\mathbb{R}^d$ we would need to impose a periodicity condition, say $f(x^i)=f(x^i+2\pi n^i)$. A nice algebraic way to phrase this is to instead define $\mathbf{T}^d_\theta$ as the algebra of all sums of products of arbitrary integer powers of a set of d variables U_i, satisfying

$$U_iU_j=e^{-i\theta^{ij}}U_jU_i. \tag{23}$$

The variable U_i takes the place of e^{ix^i} in our previous notation, and the derivation of the Weyl algebra from Eq. (1) is familiar from quantum mechanics. Similarly, we take

$$[\partial_i,U_j]=i\delta_{ij}U_j$$

and

$$\int \mathrm{Tr}U_1^{n_1}\cdots U_d^{n_d}=\delta_{\vec{n},0}.$$

There is much more to say in this case about the topological aspects, but we postpone this to Sec. VI.

C. Field theory actions and symmetries

Field theories of matrix scalar fields are very familiar and are treated in most textbooks on quantum field theory. The matrix generalization is essential in discussing Yang-Mills theory. In a formal sense we shall now make explicit any field theory Lagrangian that is written in terms of matrix fields, matrix addition and multiplication, and the derivative and integral, can be equally well regarded as a noncommutative field theory Lagrangian, with the same equations of motion and (classical) symmetry properties as the matrix field theory.

Let us consider a generic matrix scalar field theory with a Hermitian matrix valued field $\phi(x)=\phi(x)^\dagger$ and (Euclidean) action

$$S=\int d^dx\sqrt{g}\left(\frac{1}{2}g^{ij}\mathrm{Tr}\partial_i\phi\partial_j\phi+\mathrm{Tr}V(\phi)\right), \tag{24}$$

where $V(z)$ is a polynomial in the variable z, $\partial_i=\partial/\partial x^i$ are the partial derivatives, and g^{ij} is the metric. The constraint we require in order to generalize a matrix action to a noncommutative action is that it be written only using the combination $\int\mathrm{Tr}$ appearing in Eq. (24); we do not allow either the integral $\int$ or the trace Tr to appear separately. In particular, the rank of the matrix N cannot appear explicitly, only in the form Tr1 combined with the integral.

Under this assumption, it is an easy exercise to check that if we replace the algebra $\mathrm{Mat}_N[C(M)]$ by a general associative algebra $\mathcal{A}$ with integral and derivative satisfying the requirements above, the standard discussion of equations of motion, classical symmetries, and Noether's theorem all go through without change. The point is that formal manipulations which work for arbitrary matrices of functions can always be made without commuting the matrices. Another way to think about this result is to imagine defining the theory in terms of an explicit matrix representation of the algebra $\mathcal{A}$.

Thus the noncommutative theory with action Eq. (24) has the standard equation of motion

$$g^{ij}\partial_i\partial_j\phi = V'(\phi)$$

and conservation laws $\partial_i J^i = 0$ with the conserved current J^i associated to a symmetry $\delta\phi(\epsilon,\phi)$ determined by the usual variational procedures,

$$\delta S = \int \mathrm{Tr} J^i \partial_i \epsilon.$$

For example, let us consider the transformations Eq. (11). In matrix field theory, these would be the infinitesimal form of a $U(N)$ internal symmetry $\phi \to U^\dagger \phi U$. Although in more general noncommutative theories $[\partial_i,\epsilon] \neq 0$ and these are not in general symmetries, we can still consider their action, and by exponentiation define an analogous $U(N)$ action. We will refer to this group as $U(\mathcal{H})$, the group of unitary operators acting on a Hilbert space $\mathcal{H}$ admitting a representation of the algebra $\mathcal{A}$. In more mathematical terms, discussed in Sec. VI, $\mathcal{H}$ will be a module for $\mathcal{A}$.

Of course, if we do not try to gauge $U(\mathcal{H})$, it could also be broken by other terms in the action, for example, source terms $\int \mathrm{Tr} J\phi$, position-dependent potentials $\int \mathrm{Tr} f V(\phi)$, and so forth.

Application of the Noether procedure to Eq. (11) leads to a conserved current T^i, which for Eq. (24) would be

$$T^i = i g^{ij}[\phi, \partial_j \phi]. \tag{25}$$

As we discussed in Sec. II.B.1, Eq. (11) includes translations and the rotations which preserve θ^{ij}, so T^i can be used to define momentum and angular momentum operators, for example,

$$P_i = -i(\theta^{-1})_{ij} \int \mathrm{Tr} x^i T^0. \tag{26}$$

Thus we refer to it as the *restricted stress-energy tensor*.

One can also apply the Noether definition to the general variation $x^i \mapsto x^i + v^i(x)$, to define a more conventional stress-energy tensor T_{ij}, discussed in Gerhold *et al.* (2000) and Abou-Zeid and Dorn (2001a). In general, the action of this stress tensor changes θ and the underlying algebra and its interpretation has not been fully elucidated at present (see Secs. VI.G and VII.C.2 for related issues).

Finally, there is a stress-energy tensor for noncommutative gauge theory which naturally appears in the relation to string theory, which we shall discuss in Secs. II.D.3 and VII.E.

We could just as well consider theories containing an arbitrary number of matrix fields with arbitrary Lorentz transformation properties (scalar, spinor, vector, and so on). However, at this point we shall only consider a generalization directly analogous to the treatment of higher spin fields in Euclidean and Minkowski space. We shall discuss issues related to curved backgrounds later; at present the noncommutative analogs of manifolds with general metrics are not well understood.

Although one can be more general, let us now assume that the derivatives ∂_i are linearly independent and satisfy the usual flat space relations $[\partial_i,\partial_j]=0$.

Given Poincaré symmetry or its subgroup preserving θ, one can use the conventional definitions for the action of the rotation group on tensors and spinors, which we do not repeat here. In particular, the standard Dirac equation also makes sense over $\mathbb{R}^d_\theta$ and $\mathbf{T}^d_\theta$, so spin-1/2 particles can be treated without difficulty.

The discussion of supersymmetry is entirely parallel to that for conventional matrix field theory or Yang-Mills theory, with the same formal transformation laws. Constraints between the dimension of space-time and the number of possible supersymmetries enter at the point we assume that the derivatives ∂_i are linearly independent. With care, one can also use the conventional superfield formalism, treating the anticommuting coordinates as formal variables which commute with elements of $\mathcal{A}$ (Ferrara and Lledo, 2000).

Finally, as long as time is taken as commutative, the standard discussion of Hamiltonian mechanics and canonical quantization goes through without conceptual difficulty. On the other hand, noncommutative time implies nonlocality in time, and the Hamiltonian formalism becomes rather complicated (Gomis *et al.*, 2001); it is not clear that it has any operator interpretation. Although functional integral quantization is formally sensible, the resulting perturbation theory is problematic as discussed in Sec. IV. It is believed that sensible string theories with timelike noncommutativity exist, discussed in Sec. VII.G.

D. Gauge theory

The only unitary quantum field theories including vector fields are gauge theories, and the standard definitions also apply in this context. However, there is a great deal more to say about the kinematics and observables of gauge theory.

A gauge connection will be a one-form A_i, each component of which takes values in $\mathcal{A}$ and satisfies $A_i = A_i^\dagger$. (See Sec. VI.D for a more general definition.) The associated field strength is

$$F_{ij} = \partial_i A_j - \partial_j A_i + i[A_i, A_j], \tag{27}$$

which under the gauge transformation

$$\delta A_i = \partial_i \epsilon + i[A_i, \epsilon] \tag{28}$$

transforms as $\delta F_{ij} = i[F_{ij}, \epsilon]$, allowing us to write the gauge invariant Yang-Mills action

$$S = -\frac{1}{4g^2} \int \mathrm{Tr} F^2. \tag{29}$$

All this works for the reasons already discussed in Sec. II.C.

Gauge invariant couplings to charged matter fields can be written in the standard way using the covariant derivative

$$D_i\phi \equiv \partial_i\phi + i[A_i,\phi]. \quad (30)$$

Finite gauge transformations act as

$$(\partial_i + iA_i, F, \phi) \to U^\dagger(\partial_i + iA_i, F, \phi)U$$

and these definitions gauge the entire $U(\mathcal{H})$ symmetry. One can also use $\mathrm{Mat}_N(\mathcal{A})$ to get the noncommutative analog of $U(N)$ gauge theory, though (at this point) not the other Lie groups.

As an example, we quote the maximally supersymmetric Yang-Mills (MSYM) Lagrangian in ten dimensions, from which $\mathcal{N}=4$ SYM in $d=4$ and many of the simpler theories can be deduced by dimensional reduction and truncation:

$$S = \int d^{10}x \, \mathrm{Tr}(F_{ij}^2 + i\overline{\chi_I}\not{D}\chi^I), \quad (31)$$

where χ is a 16-component adjoint Majorana-Weyl fermion. This action satisfies all of our requirements and thus leads to a wide variety of supersymmetric noncommutative theories. Indeed, noncommutativity is the only known generalization (apart from adding irrelevant operators and taking limits of this) which preserves maximal supersymmetry.

Because one cannot separately define integral and trace, the local gauge invariant observables of conventional gauge theory do not carry over straightforwardly: only $\int \mathrm{Tr}O$ is gauge invariant. We now discuss this point.

1. The emergence of space-time

The first point to realize is that the gauge group in noncommutative theory contains space-time translations. This is already clear from the expression Eq. (8), which allows us to express a translation $\delta A_i = v^j\partial_j A_i$ in terms of a gauge transformation Eq. (28) with $\epsilon = v^j(\theta^{-1})_{jk}x^k$. Actually, this produces

$$\delta A_i = v^j\partial_j A_i + v^j(\theta^{-1})_{ji},$$

but an overall constant shift of the vector potential drops out of the field strengths and has no physical effect in infinite flat space.

Taking more general functions for ϵ will produce more general space-time transformations. As position-dependent translations, one might compare these with coordinate definitions or diffeomorphisms. To do this, we consider the products as star products and expand Eq. (21) in θ, to obtain

$$\delta\phi = i[\phi,\epsilon] = \theta^{ij}\partial_i\phi\partial_j\epsilon + \mathcal{O}(\partial^2\phi\partial^2\epsilon). \quad (32)$$

$$\to \{\phi,\epsilon\}, \quad (33)$$

($\{,\}$ is the Poisson bracket), so at leading order the gauge group is the group of canonical transformations preserving θ (we discuss this further in Sec. II.E.5). Of course, the higher derivative terms modify this result. In fact, the full gauge group $U(\mathcal{H})$ is simpler, as we shall see in Sec. II.E.2.

Another aspect of this unification of space-time and gauge symmetry is that if the derivative is an inner derivation, we can absorb it into the vector potential itself. In other words, we can replace the covariant derivatives $D_i = \partial_i + iA_i$ with connection operators in $\mathbb{R}^d_\theta$,

$$C_i \equiv (-i\theta^{-1})_{ij}x^j + iA_i \quad (34)$$

such that

$$D_i f \to [C_i, f]. \quad (35)$$

We also introduce the "covariant coordinates,"

$$Y^i = x^i + \theta^{ij}A_j(x). \quad (36)$$

If θ is invertible, then $Y^i = i\theta^{ij}C_j$ and this is just another notation, but the definition makes sense more generally.

In terms of the connection operators, the Yang-Mills field strength is

$$F_{ij} = i[D_i, D_j] \to i[C_i, C_j] - (\theta^{-1})_{ij}, \quad (37)$$

and the Yang-Mills action becomes a simple "matrix model" action,

$$S = \mathrm{Tr}\sum_{i,j} \{i[C_i, C_j] - (\theta^{-1})_{ij}\}^2. \quad (38)$$

Now, although we motivated this from Eq. (29), we could look at this the other way around, starting with the action Eq. (38) as a function of matrices C_i and postulating Eq. (34), to derive noncommutative gauge theory (and Yang-Mills theory in the limit $\theta\to 0$) from a matrix model. This observation is at the heart of most of the common ways that noncommutative gauge theory arises in particle physics, as the action Eq. (38) and its supersymmetrization is simple enough to arise in a wide variety of contexts. For example, it can be obtained as a limit of the twisted Eguchi-Kawai model (Eguchi and Nakayama, 1983; Gonzalez-Arroyo and Okawa, 1983), which was argued to reproduce the physics of large N Yang-Mills theory. The maximally supersymmetric version obtained in the same way from Eq. (31), often referred to as the "IKKT model" (Ishibashi *et al.*, 1997), plays an important role in M theory, to be discussed in Sec. VII.

Having so effectively hidden it, we might well wonder how d-dimensional space-time is going to emerge again from Eq. (38). Despite appearances, we do not want to claim that noncommutative gauge theory is the same in all dimensions d. Now in the classical theory, to the extent we work with explicit expressions for A_i in Eq. (34), this is generally not a problem. However, in the quantum theory, we need to integrate over field configurations C_i. We shall need to argue that this functional integral can be restricted to configurations which are similar to Eq. (34) in some sense.

This point is related to what at first appears only to be a technical subtlety involving the θ^{-1} terms in Eq. (38). They are there to cancel an extra term $[(\theta^{-1}x)_i, (\theta^{-1}x)_j]$, which would have led to an infinite constant shift of the action. The subtlety is that one could have made a mistake at this point by assuming that $\mathrm{Tr}[C_i,C_j]=0$, as for finite dimensional operators.

Of course the possibility that $\mathrm{Tr}[C_i,C_j]\neq 0$ probably comes as no surprise, but the point we want to make is that $\mathrm{Tr}[C_i,C_j]$ should be considered a *topological* aspect of the configuration. We shall argue in Sec. VI.A that it is invariant under any variation of the fields which (in a sense) preserve the asymptotics at infinity. This invariant detects the presence of the derivative operators in C_i and this is the underlying reason one expects to consistently identify sectors with a higher dimensional interpretation in what is naively a zero-dimensional theory.

2. Observables

All this is intriguing, but it comes with conceptual problems. The most important of these is that it is difficult to define local observables. This is because, as noted from the start, there is no way to separate the trace over $\mathcal{H}$ (required for gauge invariance) from the integral over noncommutative space. We can easily enough write gauge invariant observables, such as

$$\int \mathrm{Tr}F(x)^n,$$

but they are not local.

A step forward is to define the Wilson loop operator. Given a path L, we write the holonomy operator using exactly the same formal expression as in conventional gauge theory,

$$W_L=P\exp\left(i\int_L d\sigma A[x(\sigma)]\right),$$

but where the products in the expansion of the path ordered exponential are star products. This undergoes the gauge transformation

$$W_L\to U^\dagger(x_1)W_{L(x_1,x_2)}U(x_2),$$

where x_1 and x_2 are the start and end points of the path $L_{1,2}$.

We can form a Wilson loop by taking for L a closed loop with $x_1=x_2$, but again we face the problem that we can only cyclically permute operators, and thus cancel $U^{-1}(x_1)$ with $U(x_1)$, if we take the trace over $\mathcal{H}$, which includes the integral over noncommutative space.

We can at least formulate multilocal observables with this construction, such as

$$\int \mathrm{Tr}O_1(x_1)W[L_{1,2}]O_2(x_2)W[L_{2,3}]\cdots O(x_n)W[L_{n,1}]$$

with arbitrary gauge covariant operators O_i at arbitrary points x_i, joined by Wilson loops. This allows us to control the distance between operators within a single trace, but not to control the distance between operators in different traces.

Actually one can do better than this, using what are called *open* Wilson loops (Ishibashi *et al.*, 2000). The simplest example is

$$W_L(k)\equiv\int \mathrm{Tr}W[L_{1,2}]e^{ik\hat{x}_2}.$$

If the distance between the end points of L and the momentum k satisfy the relation Eq. (16), $\theta^{ij}k_i=(x_1-x_2)$, this operator will be gauge invariant, as one can see by using Eq. (15).

This provides an operator which carries a definite momentum and which can be used to define a version of local correlation functions. There is a pleasing correspondence between its construction and the dipole picture of Sec. II.B.2; not only can we think of a plane wave as having a dipole extent, we should think of the two ends of the dipole as carrying opposite electric charges which for gauge invariance must be attached to a Wilson line.

The straight line has a preferred role in this construction, and the open Wilson loop associated to the straight line with length determined by Eq. (16) can be written

$$W(k)\equiv\int \mathrm{Tr}e^{k\times C}=\int \mathrm{Tr}e^{ik\cdot Y},$$

where Y^i are the covariant coordinates of Eq. (36). This construction can be used to covariantize local operators as follows: given an operator $\mathcal{O}(x)$, transforming in the adjoint, and momentum k, we define

$$W[\mathcal{O}](k)=\int \mathrm{Tr}e^{k\times C}\mathcal{O}(x)=\int \mathrm{Tr}e^{ik_mY^m(x)}\mathcal{O}(x),$$

$$\mathcal{O}[y]=\int d^dk\int \mathrm{Tr}e^{ik\cdot(Y-y)}\mathcal{O}(x). \tag{39}$$

3. Stress-energy tensor

As we discussed above, the simplest analog of the stress-energy tensor in noncommutative field theory is Eq. (25), which generates the noncommutative analog of canonical transformations on space-time. However, in noncommutative gauge theory, this operator is the generator of gauge transformations, so it must be set to zero on physical states. This leads to a subtlety analogous to one known in general relativity: one cannot define a gauge invariant local conserved momentum density. This is compatible with the difficulties we just encountered in defining local gauge invariant observables.

One can nevertheless regard Eq. (26) as a nontrivial global conserved momentum. This is because it corresponds to a formal gauge transformation with a parameter $\epsilon\sim x^i$ which does not fall off at infinity (on the torus, it is not even single valued), and as such can be consistently excluded from the gauge group. This type of consideration will be made more precise in Sec. VI.A.

One can make a different definition of stress-energy tensor, motivated by the relation Eq. (34) between the connection and the noncommutative space-time coordinates, as the Noether current associated to the variation

$$C_i\to C_i+a_i(k)e^{ik\cdot Y}$$

which for the action Eq. (38) can easily be seen to produce

$$T_{ij}(k)=\sum_l \int_0^1 ds \int \mathrm{Tr} e^{isk\cdot Y}[C_i,C_l]e^{i(1-s)k\cdot Y}[C_j,C_l], \tag{40}$$

which is conserved in the sense that $k_m\theta^{mn}T_{nl}(k)=0$ for a solution of the equations of motion. This appears to be the natural definition in string theory, as we discuss in Sec. VII.E.

4. Fundamental matter

Another type of gauge invariant observable can be obtained by introducing new fields (bosons or fermions) which transform in the fundamental of the noncommutative gauge group. In other words, we consider a field ψ which is operator valued just as before, but instead of transforming under $U(\mathcal{H})$ as $\psi\to U\psi U^\dagger$, we impose the transformation law

$$\psi\to U\psi.$$

More generally, we need to define multiplication $a\cdot\psi$ by any element of $\mathcal{A}$, but this can be inferred using linearity.

Bilinears such as $\psi^\dagger\psi$, $\psi^\dagger D_i\psi$, and so on will be gauge invariant and can be used in the action and to define new observables, either by enforcing an equation of motion on ψ or doing a functional integral over ψ (Ambjorn *et al.*, 2000; Rajaraman and Rozali, 2000; Gross and Nekrasov, 2001).

Although in a strict sense this is also a global observable, an important point (which will be central to Sec. VI.D) is that one can also postulate an independent rule for multiplication by $\mathcal{A}$ (and the unitaries in $\mathcal{A}$) on the right,

$$\psi\to\psi a.$$

Indeed, if we take $\psi\in\mathcal{A}$, we shall clearly get a nontrivial second action of this type, since left and right multiplication are different. In this case, we can think of $\psi^\dagger\psi$ as a function on a second, dual noncommutative space. For each $f\in\mathcal{A}$ one obtains a gauge invariant observable $\mathrm{Tr}\psi^\dagger\psi f$, which is local on the dual space in the same sense that an noncommutative field is a local observable in an ungauged theory.

Taking $\psi\in\mathcal{A}$ is a choice. One could also have taken $\psi\in\mathcal{H}$, which does not lead to such a second multiplication law. The general theory of this choice is discussed in Sec. VI.D.

The two definitions lead to different physics. Let us compare the spectral density. If we take the Dirac operator $\gamma^i D_i$ acting on $\psi\in\mathcal{H}_r$, this has $d\rho(E)\sim dE E^{r-1}$, as for a field in r dimensions.

If we take the same Dirac operator with $\psi\in\mathcal{A}$, we would get infinite spectral density. A more useful definition is

$$D\!\!\!/\,\psi=\gamma^i(D_i\psi-\psi\partial_i),$$

which fixes this by postulating an ungauged right action of translations, leading to a spectral density appropriate to $2r$ dimensions.

5. The Seiberg-Witten map

Having discovered an apparent generalization of gauge theory, we should ask ourselves to what extent this theory is truly novel and to what extent we can understand it as a conventional gauge theory. This question will become particularly crucial once we find noncommutative gauge theory arising from open string theory, as general arguments imply that open string theory can always be thought of as giving rise to a conventional gauge theory. Is there an inherent contradiction in these claims?

Seiberg and Witten (1999) proposed that not only is there no contradiction, but that one should be able to write an explicit map from the noncommutative vector potential to a conventional Yang-Mills vector potential, explicitly exhibiting the equivalence between the two classes of theories.

One might object that the gauge groups of noncommutative gauge theory and conventional gauge theory are different, as is particularly clear in the rank 1 case. However, this is not an obstacle to the proposal, as only the physical configuration space—namely, the set of orbits under gauge transformation—must be equivalent in the two descriptions. It does imply that the map between the two gauge transformation laws must depend on the vector potential, not just the parameter.

Thus the proposal is that there exists a relation between a conventional vector potential A_i with the standard Yang-Mills gauge transformation law with parameter ϵ, Eq. (28), and a noncommutative vector potential $\hat{A}_i(A_i)$ and gauge transformation parameter $\hat{\epsilon}(A,\epsilon)$ with noncommmutative gauge invariance $\hat{\delta}\hat{A}_i=\partial_i\hat{\epsilon}+i\hat{A}_i*\hat{\epsilon}-i\hat{\epsilon}*\hat{A}_i$, such that

$$\hat{A}(A)+\hat{\delta}_{\hat{\epsilon}}\hat{A}(A)=\hat{A}(A+\delta_\epsilon A). \tag{41}$$

This equation can be solved to first order in θ without difficulty. Writing $\theta=\delta\theta$, we have

$$\hat{A}_i(A)-A_i=-\frac{1}{4}\delta\theta^{kl}\{A_k,\partial_l A_i+F_{li}\}_+ +\mathcal{O}(\delta\theta^2), \tag{42}$$

$$\hat{\epsilon}(A,\epsilon)-\epsilon=\frac{1}{4}\delta\theta^{kl}\{\partial_k\epsilon,A_l\}_+ +\mathcal{O}(\delta\theta^2), \tag{43}$$

where $\{A,B\}_+\equiv AB+BA$. The corresponding first-order relation between the field strengths is

$$\hat{F}_{ij}-F_{ij}=\frac{1}{4}\delta\theta^{kl}(2\{F_{ik},F_{jl}\}_+ -\{A_k,D_lF_{ij}+\partial_lF_{ij}\}_+).$$

This result even admits a reinterpretation which defines the map to all finite orders in θ. Consider the problem of mapping a noncommutative gauge field $\hat{A}^{(\theta)}$ defined with respect to the star product for θ, to a noncommutative gauge field $\hat{A}^{(\theta+\delta\theta)}$ defined for a nearby choice of θ. To first order in $\delta\theta$, it turns out that the solution to the corresponding relation (41) is again Eqs. (42) and (43), now with the right-hand side evaluated using the star product for θ. Thus these equations can be interpreted as differential equations (Seiberg-Witten equations) determining the map to all orders.

Equation (42) can be solved explicitly for the case of a rank one gauge field with constant F. In this case, it reduces to

$$\delta\hat{F} = -\hat{F}\,\delta\theta\,\hat{F},$$

where Lorentz indices are contracted as in matrix multiplication. It has the solution (with boundary condition F at $\theta=0$)

$$\hat{F} = (1+F\theta)^{-1}F. \tag{44}$$

This result can be used to relate the conventional and noncommutative gauge theory actions at leading order in a derivative expansion, as we shall discuss in detail in Sec. VII.C.

All this might suggest that the noncommutative framework is merely a simpler way to describe theories which could have been formulated as conventional gauge theories, by just applying the transformation $\hat{F} \to F$ to the action. This, however, ignores the possibility that the map might take nonsingular field configurations in the one description, to singular field configurations in the other. Indeed, Eq. (44) gives an explicit example. When $F=-\theta^{-1}$, the noncommutative description appears to break down, as $\hat{F}$ would have a pole. Conversely, F is singular when $\hat{F}=\theta^{-1}$.

As we continue, we shall find many examples in which noncommutative gauge theory has different singular solutions and short-distance properties from conventional gauge theory, and despite this formal relation between the theories it will become clear that their physics is in general rather different.

A solution to the Seiberg-Witten equation was recently found (Liu, 2000; Liu and Michelson, 2001a; Mukhi and Suryanarayana, 2001; Okawa and Ooguri, 2001b). Namely, the following inhomogeneous even degree form on $\mathbb{R}^d$, defined as the integral over the superspace $\mathbb{R}^{d|d}$, is closed:

$$\int d^dk\, d^d\vartheta\, \mathrm{Tr}_{\mathcal{H}}\left[\exp\left(\frac{\mathcal{F}}{2\pi i}\right)\right]\rho(k)\mathcal{F}$$
$$= k_i(Y^i - y^i) - \vartheta_i dy^i + \vartheta_i\vartheta_j[Y^i, Y^j], \tag{45}$$

where y's are the coordinates on $\mathbb{R}^d$, and $\rho(k)$ can be any smooth function such that $\rho(0)=1$ (this slightly generalizes the references, which take $\rho=1$). This expression has an expansion in differential forms on $\mathbb{R}^d$, whose two-form part is the conventional $F+\theta^{-1}$. Deeper aspects of this rather suggestive superspace expression will be discussed in Nekrasov (2001).

E. Bases and physical pictures

The algebra $\mathbb{R}^d_\theta$ of "functions on noncommutative $\mathbb{R}^d$," considered as a linear space, admits several useful bases. Since it is just a product of Heisenberg algebras and commuting algebras, all of this formalism can be traced back to the early days of quantum mechanics, as can much of its physical interpretation. In quantum mechanics, it appears when one considers density matrices (the Wigner functional) and free Fermi fluids (in one dimension, this leads to bosonization and W_∞ algebra).

However, we stress that the noncommutativity under discussion here is *not* inherently quantum mechanical. Rather, it is a formal device used to represent a particular class of interactions between fields, which can exist in either classical or quantum field theory. In particular, an essential difference with the standard quantum-mechanical applications is that these involve linear equations, while we are going to encounter general nonlinear equations.[2]

1. Gaussians and position-space uncertainty

While the plane-wave basis is particularly good for perturbation theory, nonperturbative studies tend to be simpler in position space. However, in noncommutative theory the standard position space basis tends not to be the most convenient, because of the nonlocal nature of the interactions. One is usually better off using a basis which simplifies the product.

One expects the noncommutativity Eq. (1) to lead to a position-space uncertainty principle, which will exclude the possibility of localized field configurations. Although there is truth to this, the point is a bit subtle, as it is certainly possible to use delta functions $\delta^{(d)}(x-x_0)$ as a basis (for the symbols) which from the point of view of the kinetic term is local.

Of course, the star product is not diagonal in this basis. Computing the star product of two delta functions leads to a kernel, which can be used to write an integral representation of the product:

$$(f\star g)(z) = \int d^dx\, d^dy\, K(x,y;z) f(x) g(y),$$
$$K(x,y;z) = \delta(z-x)\star\delta(z-y)$$
$$= \frac{1}{(2\pi)^d}\int d^dk\, e^{ik(z-x)}\delta(z-y-\theta k)$$
$$= \frac{1}{(2\pi)^d \det\theta} e^{i(z-x)\theta^{-1}(z-y)}. \tag{46}$$

In particular, the star product

$$\delta(z)\star f(y) = \frac{1}{(2\pi)^d\det\theta}\int d^dy\, e^{iy\theta^{-1}z} g(y) \tag{47}$$

is a highly nonlocal operation: it is the composition of a Fourier transform with the linear transformation $z \to \theta^{-1}z$.

As in quantum mechanics, one might expect the Gaussian to be a particularly nice basis state, since it is

[2]The special case of the equation $\phi^2=\phi$ defining a projection can arise as a normalization condition in quantum mechanics. Also, somewhat similar nonlinear equations appear in the approximation methods of quantum statistical mechanics.

simultaneously Gaussian in both conjugate coordinates. Let $\psi_{\mathbf{M},\mathbf{a}}$ be a Gaussian with center $\mathbf{a}$, covariance $\mathbf{M}$, and maximum 1:

$$\psi_{\mathbf{M},\mathbf{a}}=\exp[-(x^i-a^i)M_{ij}(x^j-a^j)],$$

which satisfies $\int d^dx|\psi|^2=[\det(2M/\pi)]^{-1/2}$.

The star product of two Gaussians can be easily worked out using Eq. (46). In particular, for concentric Gaussians of width a and b, we have

$$\psi_{(1/a^2)\,\mathbf{1},0}\star\psi_{(1/b^2)\,\mathbf{1},0}=C(a,b)\psi_{[1/\Delta(a,b)^2]\mathbf{1},0} \tag{48}$$

with

$$\Delta(a,b)^2=\frac{a^2b^2+\theta^2}{a^2+b^2},$$

$$C(a,b)=\left(1+\frac{\theta^2}{(ab)^2}\right)^{-d/2}. \tag{49}$$

This result illustrates the sense in which interactions in noncommutative theory obey a position-space uncertainty principle. Formally, we can construct a Gaussian configuration of arbitrarily small width in the noncommutative theory; its limit is the delta function we just discussed. Unlike commutative theory, however, multiplication by a Gaussian of width $b^2<\theta$ does not concentrate a field configuration but instead tends to disperse it. This is particularly clear for the special case $a=b$ of Eq. (49). More generally, the operation of multiplication by a Gaussian $\psi_{b\mathbf{1},0}$ (for any b) will cause the width to approach θ, decreasing ($\Delta<a$) if $a^2>\theta$ and increasing if $a^2<\theta$.

For some purposes, one can think of a Gaussian as having a minimum "effective size" $\max\{a,1/(\theta a)\}$. This is a bit imprecise, however. For example, the result can be a Gaussian with $\Delta(a,b)^2<\theta$, which will be true if and only if $(a^2-\theta)(b^2-\theta)<0$. A better picture is that star product with a small Gaussian is similar to the Fourier transform (47).

The configuration with the minimum effective size is evidently the Gaussian of width $a^2=\theta$. One of its special features is that its product with any Gaussian will be a Gaussian of width θ, and thus a basis can be defined consisting entirely of such Gaussians. This can be done using coherent states and we shall return to this below.

2. Fock space formalism

A nice formal context which provides a basis including the minimal Gaussian is to use as noncommutative coordinates creation and annihilation operators acting on a Fock space. These are defined in terms of the canonical coordinates of Eq. (7) in the usual way,

$$\mathbf{a}_\alpha=\frac{q_\alpha+ip_\alpha}{\sqrt{2\theta_\alpha}},\quad \mathbf{a}^\dagger{}_\alpha=\frac{q_\alpha-ip_\alpha}{\sqrt{2\theta_\alpha}},$$

$$z_\alpha=\sqrt{2\theta_\alpha}\mathbf{a}_\alpha,\quad \overline{z}_\alpha=\sqrt{2\theta_\alpha}\mathbf{a}^\dagger{}_\alpha,$$

$$[\mathbf{a}_\alpha,\mathbf{a}^\dagger{}_\beta]=\delta_{\alpha\beta}. \tag{50}$$

We can now identify elements of $\mathbb{R}^d_\theta$ with functions of the y_a valued in the space of operators acting in the Fock space $\mathcal{H}_r$ of r creation and annihilation operators. The Fock space $\mathcal{H}_r$ is the Hilbert space $\mathcal{H}$ of our previous discussion, and this basis makes the nature of $U(\mathcal{H})$ particularly apparent: it is the group of all unitary operators on Hilbert space. Explicitly,

$$\mathcal{H}_r=\oplus\mathbb{C}|n_1,\ldots,n_r\rangle,$$

$$\mathbf{a}_\alpha|\ldots,n_\alpha,\ldots\rangle=\sqrt{n_\alpha}|\ldots,n_\alpha-1,\ldots\rangle,$$

$$\mathbf{a}^\dagger{}_\alpha|\ldots,n_\alpha,\ldots\rangle=\sqrt{n_\alpha+1}|\ldots,n_\alpha+1,\ldots\rangle,$$

$$\hat{n}_\alpha=\mathbf{a}^\dagger{}_\alpha\mathbf{a}_\alpha. \tag{51}$$

We have also introduced the number operator $\hat{n}_\alpha$. Real functions of the original real coordinates correspond to the Hermitian operators.

In this language, the simplest basis we can use consists of the elementary operators $|\vec{k}\rangle\langle\vec{l}|$. These can be expressed in terms of $\mathbf{a}^\dagger{}_\alpha$ and $\mathbf{a}_\alpha$ as

$$|\vec{k}\rangle\langle\vec{l}|=\sum_{\vec{n}}\prod_\alpha(-1)^{n_\alpha}\frac{\mathbf{a}_\alpha^{\dagger k_\alpha+n_\alpha}\mathbf{a}_\alpha^{l_\alpha+n_\alpha}}{n_\alpha!\sqrt{k_\alpha!l_\alpha!}}.$$

Using Eq. (8), the derivatives can all be written as commutators with the operators $\mathbf{a}_\alpha,\mathbf{a}^\dagger{}_\alpha$. The integral (13) becomes the standard trace in this basis,

$$\int d^{2r}x\,\mathrm{Tr}\to\prod_\alpha(2\pi\theta_\alpha)\mathrm{Tr}. \tag{52}$$

The operators $|\vec{k}\rangle\langle\vec{k}|$, their sums and unitary rotations of these, provide a large set of projections, operators P satisfying $P^2=P$. This is in stark contrast with the algebra $C(M)$ (where M is a connected space) which would have had only two projections, 0 and 1, and this is a key difference between noncommutative and commutative algebras. Physically, this will lead to the existence of new solitonic solutions in noncommutative theories, as we discuss in Sec. III.

A variation on the projection which is also useful in generating solutions is the partial isometry, which by definition is any operator R satisfying

$$RR^\dagger R=R. \tag{53}$$

Such an operator can be written as a product $R=PU$ of a projection P and a unitary U. The simplest example is the shift operator $S^\dagger_\alpha$ with matrix elements

$$S^\dagger_\alpha|\ldots,n_\alpha,\ldots\rangle=|\ldots,n_\alpha+1,\ldots\rangle. \tag{54}$$

It satisfies

$$S_\alpha S^\dagger_\alpha=1$$

and

$$S^\dagger_\alpha S_\alpha=1-\sum_{n_\beta;\beta\neq\alpha}|n_\beta,0_\alpha\rangle\langle n_\beta,0_\alpha|. \tag{55}$$

3. Translations between bases

In this subsection we assume for simplicity that $d=2r$.

The translation between this basis and our previous descriptions involving commutative functions and the star product, as we commented above, can be done using the plane-wave basis and Fourier transform.

A standard tool from quantum mechanics which facilitates such calculations is the coherent state basis (Klauder and Sudarshan, 1968). Note that this is *not* a basis of $\mathbb{R}^d_\theta$ but rather is a basis of $\mathcal{H}_r$. We recall their definition

$$(\xi| = \langle 0| e^{\xi_\alpha \mathbf{a}_\alpha}, \quad |\eta) = e^{\eta_\alpha \mathbf{a}^\dagger{}_\alpha}|0\rangle, \tag{56}$$

a useful formula for matrix elements of $\hat{f}$ in this basis,

$$(\xi|\hat{f}|\eta) = \int f(z,\bar{z}) \frac{d^r z\, d^r \bar{z}}{\Pi_\alpha \pi \theta_\alpha} e^{\xi\cdot\eta - (1/\theta)(\bar{z} - \xi\sqrt{2\theta})\cdot(z - \eta\sqrt{2\theta})}, \tag{57}$$

$$(\bar{z} - \xi\sqrt{2\theta})\cdot(z - \eta\sqrt{2\theta}) = \sum_\alpha (\bar{z} - \xi_\alpha\sqrt{2\theta_\alpha})(z - \eta_\alpha\sqrt{2\theta_\alpha})$$

and a formula which follows from this for matrix elements between Fock basis states with vectors of occupation numbers $\vec{k}$ and $\vec{l}$:

$$\langle \vec{k}|\hat{f}|\vec{l}\rangle = \prod_\alpha \frac{1}{\sqrt{k_\alpha!\, l_\alpha!}} \partial_{\xi_\alpha}^{k_\alpha} \partial_{\eta_\alpha}^{l_\alpha}|_{\xi=\eta=0} (\xi|\hat{f}|\eta). \tag{58}$$

For example, let us consider the projection operator $\hat{f}_0 \equiv |0\rangle\langle 0|$. We have $(\eta|\hat{f}_0|\xi) = 1$, and using Eq. (58) we can reproduce this with $f_0(z,\bar{z}) = 2^r \exp[-z\cdot(1/\theta)\bar{z}]$, so $\hat{f}_0$ is precisely the minimal Gaussian we encountered earlier. One could extend this to use

$$\hat{f}_{\tilde{\eta},\tilde{\xi}} = e^{-\tilde{\xi}\cdot\tilde{\eta}}|\tilde{\eta})(\tilde{\xi}| \leftrightarrow f_{\tilde{\eta},\tilde{\xi}} = 2^r e^{-(\bar{z} - \tilde{\xi}\sqrt{2\theta})\cdot\theta^{-1}(z - \tilde{\eta}\sqrt{2\theta})} \tag{59}$$

as an overcomplete basis for $\mathbb{R}^d_\theta$, consisting of minimal Gaussians with centers $x = (\tilde{z}, \tilde{z}^*)$ [see, e.g., Eq. (15)] multiplied by plane waves with momentum $k = (\tilde{\kappa}, \tilde{\kappa}^*)$, with

$$\tilde{\kappa} = \frac{1}{2\theta}(\tilde{\xi} - \tilde{\eta}^*), \quad \tilde{z} = \sqrt{\frac{\theta}{2}}(\tilde{\xi} + \tilde{\eta}^*)$$

(where * denotes complex conjugation). As another example, the delta function $\delta^{(d)}(x)$ has matrix elements [from Eq. (57)] $(\eta|\hat{\delta}|\xi) = e^{-\eta\xi}$, leading to the expressions

$$\langle \vec{k}|\hat{\delta}|\vec{l}\rangle = \delta_{\vec{k},\vec{l}}(-1)^{|\vec{k}|}, \quad \hat{\delta} = (-1)^{|\hat{\vec{n}}|} \tag{60}$$

with $|\vec{k}| = \Sigma_\alpha k_\alpha, |\hat{\vec{n}}| = \Sigma_\alpha \hat{n}_\alpha$.

Let us now express the general radially symmetric function in two noncommutative dimensions in the two bases (we take $\theta = 1/2$). These are functions of $r^2 = p^2 + q^2 \sim z\bar{z} = \hat{n}$, so the general such function in the Fock basis is

$$\sum_{n\geq 0} c_n \hat{f}_n = \sum_{n\geq 0} c_n |n\rangle\langle n|. \tag{61}$$

The corresponding symbols f_n can be found as solutions of the equations $z\bar{z} * f_n = (n+1/2) f_n$ (Fairlie, 1964; Curtright *et al.*, 2000). A short route to the result is to form a generating function from these operators,

$$\hat{f} = \sum_{n\geq 0} u^n |n\rangle\langle n|, \tag{62}$$

whose matrix elements are $(\eta|\hat{f}|\xi) = \exp(u\eta\xi)$. This can be obtained from Eq. (57) by taking for f the generating function

$$f(z\bar{z};u) = \left(\lambda - \frac{1}{2}\right)\exp(\lambda z\bar{z})$$

with $1-u = 1/(\lambda - \frac{1}{2})$. Substituting u for λ in this expression leads to a generating function for the Laguerre polynomials $L_n(4r^2)$ (Bateman, 1953).

The final result for the symbol of Eq. (62) is

$$f = 2\sum_{n\geq 0} u^n L_n(4r^2) e^{-2r^2}.$$

4. Scalar Green's functions

We now discuss the Green's function of the free noncommutative scalar field. This is very simple in the plane-wave basis, in which the Klein-Gordon operator is diagonal:

$$(-\partial_i\partial^i + m^2)e^{ikx} = (k^2 + m^2)e^{ikx}.$$

The Green's function satisfying

$$\left(-\sum_i \frac{\partial^2}{\partial x_i^2} + m^2\right) G(x,y) = \delta(x,y) \tag{63}$$

is then just

$$G(k,k') = \frac{\delta_{k,k'}}{k^2 + m^2}. \tag{64}$$

This result can be easily transformed to other bases using coherent states. We set $\theta_\alpha = \frac{1}{2}$ and start from

$$(\eta| e^{ik\bar{z} + i\bar{k}z}|\xi) = e^{-(1/2)k\bar{k} + i\bar{k}\eta + ik\xi + \eta\xi} \tag{65}$$

which is derived from Eq. (57) by Gaussian integration.

This allows us to derive the matrix elements of the Greens function $\hat{G}$ by Fourier transform:

$$\left(\eta\left|\int d^d k \frac{e^{ikx}}{k^2 + i\epsilon}\right|\xi\right) = \int \frac{d^r k\, d^r \bar{k}\, d^{d-2r} p}{k\bar{k} + p^2 + i\epsilon} e^{-(1/2)k\bar{k} + i\bar{k}\eta + ik\xi + \eta\xi + ipy}.$$

It is convenient to express this as a proper time integral. We can then easily include commuting dimensions as well; let there be $2r$ noncommuting and $d-2r$ commuting dimensions, with momenta k and p, respectively. We include a separation y in the commuting directions for purposes of comparison. We then have

$$(\eta|\hat{G}|\xi) = \int_0^\infty dt \int d^r k\, d^r \bar{k}\, d^{d-2r} p\, e^{-tp^2 + ipy} \times e^{-tk\bar{k} - (1/2)k\bar{k} + i\bar{k}\eta + ik\xi + \eta\xi}. \tag{66}$$

The integrals over momenta are again Gaussian:

$$=(2\pi)^{d/2}e^{\eta\xi}\int_0^\infty \frac{dt}{(2t)^{(d-2r)/2}(2t+1)^r}e^{-2\eta\xi/(2t+1)-y^2/4t}. \quad (67)$$

This is the result. Let us try to compare its behavior in noncommutative position space with that in commutative space. This comparison can be based on Eq. (59), which tells us that $e^{-a^2}(\eta=a|\hat{G}|\xi=a)$ evaluates the Green's function on a Gaussian centered at $x=a$. The IR (long-distance) behavior is controlled by the limit $t\to\infty$ of the proper time integral, and we see that it has the same dependence on a as on y. Thus, in this sense, the IR behavior is the same as for the commutative Green's function.

The UV behavior is controlled by short times $t\to 0$, and at first sight looks rather different from that of the commutative Greens function, due to the shifts $t\to t+1/2$. However, this difference is only apparent and comes because the noncommutative Greens function in effect contains a factor of the delta function Eq. (60), which is hiding the UV divergence. To see this, we can compute $\mathrm{Tr}\hat{\delta}\hat{G}$, which is the correct way to take the coincidence limit.

This can be done in the coherent state basis, but we instead make a detour to expand the Green's function in the Fock basis. This can be done using Eq. (58); one finds that the matrix elements are nonzero only on the diagonal and are a function only of the sum of the occupation numbers $\Sigma_a n_a$—this reflects rotational invariance. Changing variables from t to $\lambda=1/(t+1/2)$, setting $y=0$, and going to the Fock basis, we obtain

$$\langle n|\hat{G}|n\rangle=\int_0^2 d\lambda\lambda^{d/2-2}(2-\lambda)^{r-d/2}(1-\lambda)^n. \quad (68)$$

This result makes it easy to answer the previous question: the sum over modes $\mathrm{Tr}\hat{\delta}\hat{G}=\Sigma_n\langle n|\hat{G}|n\rangle$ produces $(2-\lambda)^{-r}$, which is exactly the UV divergent factor which was missing from Eq. (68).

In effect, the coincidence limit of the Green's function diverges in the same way in noncommutative theory as in conventional theory. This was clear in the original plane-wave basis, and will imply that many loop amplitudes have exactly the same UV divergence structure as they would have had in conventional theory, as we shall discuss in Sec. IV.

We shall find ourselves discussing more general Green's functions in the interacting theory. Given a two-point function $G^{(2)}(k)$ in momentum space, the same procedure can be followed to convert it to the coherent state basis and thus interpret it in noncommutative position space. We shall not try to give general results but instead simply assert that this allows one to verify that the standard relations between asymptotic behavior in momentum and position space are valid in noncommutative space.

First, the long-distance behavior is controlled by the analytic behavior in the upper half plane; if the closest pole to the real axis is located at $\mathrm{Im}\,k=m$, the Euclidean Green's function will fall off as e^{-mr} in position space.

Similarly, the short-distance behavior is controlled by the large-k asymptotics, $G^{(2)}(k)\sim k^{-\alpha}$ implies $G(r)\sim r^{\alpha-d}$.

5. The membrane/hydrodynamic limit

The formal relation between $\mathbb{R}^d_\theta$ and the Heisenberg algebra of quantum mechanics suggests that it should be interesting to consider the analog of the $\hbar\to 0$ limit. Since θ is not $\hbar$, this is not a classical limit, but rather a limit in which the noncommutative fields can be treated as functions rather than operators. This limit helps provide some intuition for the noncommutative gauge symmetry, and has diverse interpretations in the various physical realizations of the theory.

We consider noncommutative gauge theory on $\mathbb{R}^d_\theta$ and take the following scaling limit: let

$$\theta=l^2\theta_0;\quad A=l^{-2}A_0;\quad g^2=g_0^2 l^{-4}, \quad (69)$$

and take $l\to 0$, keeping all quantities with the 0 subscript fixed. This corresponds to weak noncommutativity with strong gauge coupling, fixing the dimensionful combination $\lambda^2=g\theta$.

Assume that θ^{ij} is invertible, and define the functions

$$y^i=x^i+\theta^{ij}A_j(x). \quad (70)$$

In the limit, the Yang-Mills action (29) becomes

$$S_h=\frac{1}{\lambda^4}\int d^dx\sum_{i,j}(\{y^i,y^j\}-\theta^{ij})^2, \quad (71)$$

where

$$\{f,g\}=\theta_0^{ij}\partial_i f\partial_j g$$

is the ordinary Poisson bracket on functions.

Infinitesimal gauge transformations take the form Eq. (32) in the limit. The corresponding finite gauge transformations are general canonical transformations, also called symplectomorphisms, which are diffeomorphisms $x\mapsto\tilde{x}(x)$ which preserve the symplectic form:

$$\theta^{-1}_{ij}d\tilde{x}^i\wedge d\tilde{x}^j=\theta^{-1}_{ij}dx^i\wedge dx^j. \quad (72)$$

In general, not all symplectomorphisms are generated by Hamiltonians as in Eq. (6); those which are not are the analogs of large gauge transformations.

The model Eq. (71) is a sigma model, in the sense that the fields y^i can be thought of as maps from a base space $\mathbb{R}^d$ to a target space, also $\mathbb{R}^d$ in this example. One can easily generalize this to let y be a map from one Poisson manifold (M,θ) to another Poisson manifold (N,π), which must also have a volume form μ and a metric $\|\cdot\|^2$ on the space of bi-vectors. The action Eq. (71) will read

$$S=\int_N \mu\,\|y_*\theta-\pi\|^2$$

and will again have the group $\mathrm{Diff}_\theta(M)$ of symplectomorphisms of M as a gauge group. This generalization also appears very naturally; for example if one starts with $d+k$-dimensional Yang-Mills theory, dimensionally reduces to d dimensions, takes these to be $\mathbb{R}^d_\theta$, and takes the limit, one obtains $M=\mathbb{R}^d$ and $N=\mathbb{R}^{d+k}$.

If some of the coordinates are commuting, one gets a gauged sigma model. The case of a single commuting timelike dimension is particularly simple in the canonical formulation: one has canonically conjugate variables (p_i, y^i), a Hamiltonian $H=\Sigma_i p_i^2 + S_h$, and a constraint $0=J=\Sigma_i\{p_i, y^i\}$. In this form, the construction we just described essentially appears in Hoppe (1982), while its maximally supersymmetric counterpart is precisely the light-cone gauge fixed supermembrane action of de Wit *et al.* (1988). Thus one interpretation of the fields y^i is as the embedding of a d-dimensional membrane into $d+k$-dimensional space.

Although this picture is a bit degenerate if $k=0$, the configuration space of such maps is still nontrivial. The related time-dependent theory describes flows of a fluid satisfying Eq. (72). This is particularly natural in two dimensions, where $\theta_{ij}dx^i dx^j$ is the area form, and these are allowed flows of an incompressible fluid. This hydrodynamic picture has also appeared in many works, for example, Bordemann and Hoppe (1993), and is also known in condensed-matter theory.

We shall use this hydrodynamic picture in what follows to illustrate solutions and constructions of the noncommutative gauge theory. As an example, the hydrodynamic limit of the Seiberg-Witten map Eq. (45) is (Cornalba, 1999)

$$F_{ij}+\theta^{-1}_{ij}=[\{y^k, y^l\}]^{-1}_{ij}. \tag{73}$$

As Susskind (2001) points out, for $d=2$ this is just the translation between the Euler and Lagrange descriptions of fluid dynamics.

Similarly, the hydrodynamic analog of the straight Wilson line is the Fourier transform of fluid density,

$$W[k]=\int d^d x \,\mathrm{Pf}(\theta^{-1}) e^{ik_i y^i(x)}. \tag{74}$$

6. Matrix representations

In a formal sense, any explicit operator representation is a matrix representation. In this subsection we discuss $\mathbf{T}^d_\theta$ as a large-N limit of a finite dimensional matrix algebra, and how this might be used to formulate regulated noncommutative field theory. We take $d=2$ for definiteness, but the ideas generalize.

One cannot, of course, realize Eq. (1) using finite dimensional matrices. One can realize Eq. (23) for the special case with $\theta^{12}=2\pi M/N$, for example, by $U_1=\Gamma_1^M$ and $U_2=\Gamma_2$, where $(\Gamma_1)_{m,n}\equiv\delta_{m,n}\exp[2\pi i(n/N)]$ is the clock matrix, and $(\Gamma_2)_{m,n}\equiv\delta_{m-n,1(\mathrm{mod}N)}$ is the shift matrix.

Products of these matrices form a basis for $\mathrm{Mat}_N(\mathbb{C})$ [we assume $\gcd(M,N)=1$] and in a certain sense this allows us to regard $\mathrm{Mat}_N(\mathbb{C})$ as an approximation to $\mathbf{T}^2_\theta$: namely, if we write

$$\Phi[x]\equiv\sum_{0\le m_1, m_2<N} U_1^{m_1}U_2^{m_2}e^{\pi i\theta^{ij}m_i m_j-2\pi i m_i x^i},$$

we find that the $\Phi[x]$ evaluated at lattice sites $x^i=n^i/N$ provide a basis in terms of which the multiplication law becomes Eq. (46), with the integrals replaced by sums over lattice points. This construction is sometimes referred to as a "fuzzy $\mathbf{T}^2$" and was independently proposed in several works (de Wit *et al.*, 1988; Hoppe, 1989; see also the references in Bars and Minic, 2000) as a starting point for regulated field theory on $\mathbf{T}^2$, since we can regard the bound $0\le m_i<N$ as a UV cutoff.

One can obtain $\mathbf{T}^2_\theta$ with more general θ by taking the limit $M,N\to\infty$ holding $\theta=2\pi M/N$ fixed, and obtain $\mathbb{R}^2_\theta$ by taking g_{ij} (and thus the volume) large in an obvious way. Formulating this limit precisely enough to make contact with our previous discussion requires some mathematical sophistication, however, as there are many distinct algebras which can be obtained from $\mathrm{Mat}_N(\mathbb{C})$ by taking different limits. For example, in Sec. VI.B we shall discuss a sense in which the large-N limit of $\mathrm{Mat}_N(\mathbb{C})$ leads to functions on a sphere, not a torus.

For most physical purposes, the definitions of $\mathbb{R}^d_\theta$ we gave previously are easier to use. On the other hand, making the limit explicit is a good starting point for making a nonperturbative definition of quantum noncommutative field theory and for a deeper understanding of the renormalization group. Physically, one usually thinks of the continuum limit (at least at weak coupling) as describing modes with low kinetic energy, so to decide which algebra will emerge in the limit, we need to consider the derivatives.

Candidate derivatives for fuzzy $\mathbf{T}^2$ are operators D_i satisfying

$$D_i\Phi[x]D_i^\dagger=\Phi[x+e_i],$$

where $(e_i)^j=\delta_i^j/N$ is the lattice spacing; e.g., take $D_1=\Gamma_2^a$ with $aM=-1\,\mathrm{mod}N$ and $D_2=\Gamma_1$. Thus a plausible regulated form of Eq. (24) might be

$$S=\frac{N^2}{2}\mathrm{Tr}\sum_i (D_i\Phi D_i^\dagger-\Phi)^2+\mathrm{Tr}V(\Phi). \tag{75}$$

The same approach could be followed for noncommutative gauge theory, leading to a rather ugly action. One aspires to the elegance of Eq. (38), in which the derivatives emerge from a choice of background configuration. It should be clear at this point (and will be made more so in Sec. VI.A) that one cannot simply take $\mathcal{A}=\mathrm{Mat}_N(\mathbb{C})$ in Eq. (38); the derivatives must be specified somehow.

Perhaps the best proposal along these lines at present is due to Ambjorn *et al.* (1999). We shall just state this as a recipe; the motivation behind it will become clearer upon reading Sec. VI.E.

The starting point is the lattice twisted Eguchi-Kawai model, a matrix model whose dynamical variables are d unitary matrices $U_i\in U(N)$, and the action

$$S=-\frac{1}{g^2}\sum_{i\neq j} Z_{ij}\mathrm{Tr}U_iU_jU_i^\dagger U_j^\dagger.$$

This is Wilson's lattice gauge theory action restricted to a single site, which is a natural nonperturbative analog of Eq. (38), generalized by the twist factors Z_{ij} which are constants satisfying $Z_{ij}=(Z_{ji})^*$.

This action is supplemented by the constraints

$$\Omega_i U_j \Omega_i^\dagger = e^{2\pi i \delta_{ij} r_i/N} U_j$$

which for suitable matrices Ω_i and constants r_i can be shown to admit $U_i=D_i\tilde{U}_i$ as solutions, where D_i are derivatives as above, and $\tilde{U}_i$ are unitary elements of $\mathbf{T}_\theta^d$.

Following the ideas above, one can show that this reproduces perturbative noncommutative gauge theory, and captures interesting nonperturbative structure of the model, namely, the Morita equivalence described in Sec. VI.F, for rational θ. It would be interesting to justify this further, for example, by detailed analysis in two dimensions.

III. SOLITONS AND INSTANTONS

Field theories and especially gauge theories admit many classical solutions: solitons, instantons, and branes, which play important roles in nonperturbative physics.

In general, solutions of conventional field theory carry over to the analogous noncommutative theory, but with any singularities smoothed out, thanks to position-space uncertainty. In particular, the noncommutative rank 1 theory has nonsingular instanton, monopole, and vortex solutions, and in this respect (and many others) is more like conventional Yang-Mills, not Maxwell theory.

An even more striking feature of noncommutative theory is that solitons can be stable when their conventional counterparts would not have been. As we shall see, noncommutativity provides a natural mechanism for stabilizing objects of size $\sqrt{\theta}$.

Another striking feature is how closely the properties of noncommutative gauge theory solutions mirror the properties of corresponding Dirichlet brane solutions of string theory. We shall discuss this aspect in Sec. VII.D.

A. Large θ solitons in scalar theories

In commutative scalar field theories in two and more spatial dimensions, there is a theorem which prohibits the existence of finite-energy classical solitons (Derrick, 1964). This follows from a simple scaling argument: upon shrinking all length scales as $L\to\lambda L$, both kinetic and potential energies decrease, so no finite-size minimum can exist.

This argument will obviously fail in the presence of a distinguished length scale $\sqrt{\theta}$ and in fact one finds that for sufficiently large θ, stable solitons can exist in the noncommutative theory (Gopakumar *et al.*, 2000).

The phenomenon can be exhibited in $2+1$ dimensions and we consider such a field theory with action (24). It is convenient to work with canonically commuting noncommutative coordinates, defined as $x^1+ix^2=z\sqrt{\theta}$ and $x^1-ix^2=\bar{z}\sqrt{\theta}$. In terms of these, the energy becomes

$$E=\int d^2z\,\frac{1}{2}(\partial\phi)^2+\theta V(\phi).$$

In the limit of large θV, the potential energy dominates, and we can look for an approximate solitonic solution by solving the equation

$$\frac{\partial V}{\partial \phi}=0. \tag{76}$$

For example, if we consider a cubic potential, this equation would be

$$V'(\phi)=g\phi^2+\phi=0. \tag{77}$$

While in commutative theory these equations would admit only constant solutions, in noncommutative theory the story is rather more interesting. It is simplest in the Fock space basis, in which the field can be taken to be an arbitrary (bounded) operator on $\mathcal{H}_r$, and for which the multiplication is just operator multiplication. Since ϕ is self-adjoint it can be diagonalized, so we can immediately write the general solution of Eq. (77):

$$\phi=-\frac{1}{g}U^\dagger P U,$$

where U is a unitary and $P^2=P$ is a projection operator, characterized up to unitary equivalence by $\mathrm{Tr}\,P$ (which must be finite for a finite-energy configuration) or equivalently the number of unit eigenvalues. As discussed in Sec. II.E.3, the diagonal operators correspond to radially symmetric solutions, from which unitary rotations produce all solutions.

The simplest solution of this type uses the operator $P_0=|0\rangle\langle 0|$, of energy $2\pi\theta V(-1/g)=\pi\theta/3g^2$. It is remarkable that this energy depends only on the value of the potential at the critical point, and nothing else.

As we discussed in Sec. II.E.2, this is a Gaussian of width $\sqrt{\theta}$ which squares to itself under star product. We see that the scaling argument for instability is violated because of the position-space uncertainty principle, which causes the energy of smaller Gaussians to increase.

The general solution in this sector is $\phi=-(1/g)\,U^\dagger|0\rangle\langle 0|U$. This includes Gaussians with arbitrary centers, the higher modes discussed in Sec. II.E.2, and various "squeezed states." Neglecting the kinetic term, they are all degenerate and are parametrized by an infinite-dimensional moduli space $\lim_{N\to\infty}U(N)/U(N-1)$. This infinite degeneracy will, however, be lifted by the kinetic term. The story is similar for solutions with $\mathrm{Tr}\,P=n$; this moduli space has a limit in which the solutions approach n widely spaced $n=1$ solitons, with exponential corrections.

This example and all of its qualitative features generalize immediately to an arbitrary equation Eq. (76). Its most general solution is

$$\phi=U^\dagger\left(\sum_i \lambda_i P_i\right)U, \tag{78}$$

where the λ_i are the critical points $V'(\lambda_i)=0$, and the P_i are a set of mutually orthogonal projections whose sum is the identity. The analysis also generalizes in an obvious way to higher dimensional theories with full spatial noncommutativity ($d=2r+1$), by tensoring projections in each Fock space factor.

Solutions for which all λ_i are minima of $V(\lambda)$ are clearly locally stable, if we neglect the kinetic energy term. We now consider the full energy functional, which in the Fock basis becomes

$$E=\mathrm{Tr}[z,\phi][\phi,\bar{z}]+\theta\mathrm{Tr}V(\phi).$$

The kinetic term breaks $U(\mathcal{H}_r)$ symmetry and might be expected to destabilize most of the infinite-dimensional space of solutions Eq. (78). On the other hand, for low modes ($|n\rangle\langle n|$ with $n\sim 1$), the kinetic energy will be $O(1)$, so for sufficiently large θ a stable solution should survive. This was checked by Gopakumar *et al.* (2000) by an analysis of linearized stability, with the result (for the $n=1$ solutions in $d=2+1$) that only the minimal Gaussian $|0\rangle\langle 0|$ and its translates are stable.

The solution cannot exist at $\theta=0$ and it is interesting to ask what controls the critical value θ_c at which it disappears (Zhou, 2000; Durhuus *et al.*, 2001; Jackson, 2001). One can easily see the rather surprising fact that this does not depend directly on the barrier height. This follows because we can obtain a family of equivalent problems with very different barrier heights by the rescaling $\phi\rightarrow a\phi$ and $E\rightarrow E/a^2$, all with the same θ_c.

Rather, the condition for noncommutative solitons to exist is that noncommutativity be important at the scale set by the mass of the ϕ particle in the asymptotic vacuum, i.e., $\theta V''\gg 1$. Consider the symmetric ϕ^4 potential. By the above argument, θ_c can only depend on V'' at the minimum; numerical study leads to the result

$$\theta_c V''(0)=13.92. \tag{79}$$

There is some theoretical understanding of this result, which suggests that this critical value is roughly independent of the shape of the potential. It is plausible that only radially symmetric configurations are relevant for stability and if one restricts attention to this sector, the equation of motion reduces to a simple three-term recurrence relation for the coefficients c_n in Eq. (61),

$$(n+1)c_{n+1}-(2n+1)c_n+nc_{n-1}=\frac{\theta}{2}V'(c_n), \quad n\geq 0. \tag{80}$$

Suppose that $V(\phi)$ is bounded below, the vacuum is $\phi=0$, and we seek a solution which approximates the $\theta=\infty$ one soliton solution ($c_0=\lambda;\, c_n=0, n>0$). Finiteness of the energy requires $\lim_{n\rightarrow\infty}c_n=0$, and we can get the large n asymptotics of such a solution by ignoring the nonlinear terms in $V'(\lambda)$; this leads to

$$c_n\sim n^{1/4}e^{-\sqrt{n\theta V''(0)}}. \tag{81}$$

This shows that c_n varies smoothly when $\theta V''(0)\ll 1$, and can be approximated by a solution of the differential equation analog of Eq. (80). However, corresponding to the nonexistence of a solution in the commutative theory, one can show on very general grounds that no solution of this differential equation can have the boundary value $c(0)=\lambda$, even with the nonlinear terms in V' included. A nontrivial soliton is possible only if this continuous approximation breaks down, which requires the control parameter $\theta V''(0)$ to be large.

In particular, a nontrivial one soliton solution must have a discontinuity $|c_0-c_1|\gg 0$, and matching this on to Eq. (81) for c_1 provides a lower bound for $\theta V''(0)$. This analysis leads to an estimate which is quite close to Eq. (79).

Multisoliton solutions have been studied recently in Lindstrom *et al.* (2000), Gopakumar *et al.* (2001), and Hadasz *et al.* (2001). Solutions on $\mathbf{T}^d_\theta$ have been studied in Bars *et al.* (2001).

B. Vortex solutions in gauge theories

Derrick's theorem does not hold in gauge theories and as is well known, the Abelian Higgs model (Maxwell theory coupled to a complex scalar field) has vortex solutions, which (among other applications) describe the flux tubes in superconductors (Nielsen and Olesen, 1973).

We proceed to discuss analogous solutions in the noncommutative gauge theory (Nekrasov and Schwarz, 1998; Bak, 2000; Gross and Nekrasov, 2000a; Polychronakos, 2000; Aganagic *et al.*, 2001;) We work in $2+1$ dimensions with $\theta_{xy}=1$ and look for time-independent solutions in $A_0=0$ gauge. Using Eq. (35) and the Fock basis, the energy is

$$E=\frac{2\pi}{g^2}\int dt\mathrm{Tr}\ \frac{1}{2}F^2+\sum_{i=1,2}D_i\phi D_i\phi^\dagger+V(\phi). \tag{82}$$

Here

$$F=[C,\bar{C}]+1$$

and ϕ is a complex scalar field (satisfying no Hermiticity condition).

We first note that unlike Maxwell theory, even the pure rank one noncommutative gauge theory admits finite-energy solitonic solutions. The pure gauge static equation of motion is

$$0=[C,[C,\bar{C}]].$$

Of course, it is solved by the vacuum configuration $C=\bar{z}$ and $\bar{C}=z$, and gauge transformations of this,

$$C=U^\dagger\bar{z}U, \quad \bar{C}=U^\dagger zU. \tag{83}$$

What is amusing is that U does not need to be unitary in order for this transformation to produce a solution (Harvey, Kraus, and Larsen, 2000; Witten, 2000). It needs to satisfy $UU^\dagger=1$, but $U^\dagger U$ need not be the identity. This implies the partial isometry condition Eq. (53) and is a bit stronger.

The simplest examples use the shift operator Eq. (54): $U=S^m$. To decide whether these are vortex solutions, we should compute the magnetic flux. This is

$$F=(S^\dagger)^m[\bar{z},z]S^m+1=1-(S^\dagger)^mS^m=\sum_{n=0}^{n-1}P_n,$$

where $P_n \equiv |n\rangle\langle n|$. The total flux is $\mathrm{Tr}F = m$. Thus noncommutative Maxwell theory allows nonsingular vortex solutions, sometimes called fluxons, without needing a scalar field or the Higgs mechanism.

Physically, we might interpret this as an noncommutative analog of the commutative gauge theory vortex $A_i = g^{-1}\partial_i g$ with $g = e^{im\theta}$. This is pure gauge except at the origin where it is singular, so we might regard this as an example of noncommutative geometry smoothing out a singularity.

The soliton mass M is proportional to $\mathrm{Tr}F^2 = m$. To restore the dependence on the coupling constants, we first note that (as for any classical soliton) the mass is proportional to $1/g^2$. This quantity has dimensions of length in $2+1$ dimensions, so on dimensional grounds the mass must be proportional to $1/\theta$, consistent with the nonexistence of the fluxon in the conventional limit. In the conventions of Eq. (82),

$$M = \frac{\pi m}{g^2 \theta}. \tag{84}$$

The most general solution is slightly more general than this; it is

$$C = (S^\dagger)^m \bar{z} S^m + \sum_{n=0}^{m-1} c_n(x^0) |n\rangle\langle n|, \tag{85}$$

$$\bar{C} = (S^\dagger)^m z S^m + \sum_{n=0}^{m-1} \bar{c}_n(x^0) |n\rangle\langle n|. \tag{86}$$

The $2m$ functions $c_n, \bar{c}_n$ must satisfy $\partial_0^2 c_n = \partial_0^2 \bar{c}_n = 0$ and can be seen to parametrize the world lines of the m fluxons. This is particularly clear for $m=1$ by recalling Eq. (34).

A peculiar feature to note is that fluxons exist with only one sign of magnetic charge, F aligned with θ. It is also rather peculiar that they exert no force on one another; the energy of the configuration is independent of their locations.

Even more peculiar, the equations $\partial_0^2 c_n = \partial_0^2 \bar{c}_n = 0$ admit as solutions $c_n = x_n + v_n t$ and $\bar{c}_n = \bar{x}_n + \bar{v}_n t$ with no upper bound on v. In other words, the fluxons can move faster than light (Bak *et al.*, 2000; Hashimoto and Itzhaki, 2000). Of course θ defines a preferred rest frame, and there is no immediate contradiction with causality in this frame.

These peculiarities may be made more palatable by the realization that all of these solutions (with no scalar field) are unstable, even to linearized fluctuations. For example, the $m=1$ solution admits the fluctuation

$$C + T = S^\dagger \bar{z} S + t S^\dagger P_0, \tag{87}$$

where t is a complex scalar parametrizing the fluctuation and $S^\dagger P_0 = |1\rangle\langle 0|$. One can straightforwardly compute

$$[C+T, \bar{C}+\bar{T}] + 1 = P_0 + |t|^2 (P_1 - P_0).$$

The total flux $\mathrm{Tr}F$ is constant under this variation, while the energy is proportional to

$$\mathrm{Tr}F^2 = (1-|t|^2)^2 + |t|^4$$

which exhibits the instability. This is of course just like the commutative case; the flux will tend to spread out over all of space in the absence of any other effect to confine it.

There is no stable minimum in this topological sector. This follows if we grant that $\mathrm{Tr}F$ is topological and cannot change under any allowed variation of the fields. This of course depends on one's definitions, but in conventional gauge theory one can make a definition which does not allow flux to disappear. Starting with a localized configuration, the flux may disperse or go to infinity, but one can always enlarge one's region to include all of it, because of causality. Similarly, energy cannot be lost at infinity.

Since noncommutative theory is not causal and indeed a fluxon can move faster than light, this argument must be reexamined. The belief at present is that flux and energy are also conserved in noncommutative gauge theory; they are conserved locally and cannot run off to infinity in finite time. In the case of the fluxon, a given solution will have some finite velocity, and as it disperses, it would be expected to slow down, so that once it has spread over length scales large compared to $\sqrt{\theta}$ conventional causality will be restored. This point could certainly use more careful examination, however.

This physical statement underlies the conventional definition which leads to topological sectors characterized by total flux; one only considers variations of the fields which preserve a specified falloff at infinity. An analogous definition can be made in noncommutative theory, as we shall discuss in Sec. VI.A.

One can use the same idea to generate exact solutions to the noncommutative Abelian Higgs theory, even with a general scalar potential $V(\phi)$. Now one starts with the scalar in a vacuum configuration $V'(\phi)=0$ and $D_i\phi = 0$ (so $\phi \propto 1$,) and again applies an almost gauge transformation with $UU^\dagger = 1$, but $U^\dagger U \neq 1$. The same argument as above shows that this will be a solution for any U. This "solution generating technique" has been used to generate many exact solutions (Harvey *et al.*, 2000; Hashimoto, 2000; Schnabl, 2000; Tseng, 2000; Bergman *et al.*, 2001; Hamanaka and Terashima, 2001)

Of course, the properties of these solutions, including stability, depend on the specific form of the scalar potential and the choice of matter representation; both adjoint matter with $D_i\phi = [C_i, \phi]$ and fundamental matter as defined in Sec. II.D.3 have been studied.

A particularly nice choice of potential (Jatkar *et al.*, 2000) is

$$V = \frac{1}{2}(\phi\phi^\dagger - m^2)^2,$$

as in this case the energy can be written as the sum of squares and a total derivative as in Bogomolny (1976), leading to a lower bound $E \geqslant |\mathrm{Tr}F|$. Solutions saturating this bound are called BPS and are clearly stable.

For adjoint matter and $\theta m^2 = 2$, the exact solution discussed above is BPS. For fundamental matter, the bound can be attained in two ways.[3] One can have

$$F=m^2-\phi\phi^\dagger, \quad \overline{D}\phi=0,$$

which has positive flux solutions. For $\theta m^2=1$ the exact solution as above with $\phi=S^\dagger$ is BPS. Bak *et al.* (2000) show that BPS solutions only exist for $\theta m^2\leqslant 1$, and argue that for $\theta m^2>1$ the exact solution is a stable non-BPS solution.

One can also have

$$F=\phi\phi^\dagger-m^2, \quad D\phi=0.$$

This has been shown to have a negative flux solution for any value of θm^2 (Jatkar *et al.*, 2000; Lozano *et al.*, 2001) which in the $\theta\to 0$ limit reduces to the conventional Nielsen-Olesen solution.

The vortex solutions admit a number of direct generalizations to higher-dimensional gauge theory. If one has $2r$ noncommuting coordinates, one can make simple direct products of the above structure to obtain solutions localized to any $2(r-n)$-dimensional hyperplane. One can also introduce additional commuting coordinates, and it is not hard to check that the parameters $c_n, \overline{c_n}$ in the solution Eq. (85) must then obey the wave equation $0=\eta^{ij}\partial_i\partial_j c$ (respectively, $\overline{c}$) in these coordinates. This is as expected on general grounds (they are Goldstone modes for space-time translations and for the symmetries $\delta\phi=\epsilon$) and fits in with the general philosophy that a soliton in $d+1$-dimensional field theory which is localized in $d-p$ dimensions should be regarded as a p brane, a dynamical object with a $p+1$-dimensional world volume which can be described by fields and a local effective action on the world-volume. While the vortex in $3+1$-dimensional gauge theory, which is a string with $p=1$, may be the most familiar case, the story in more dimensions is entirely parallel.

One can get a nontrivial fluxon solution in the hydrodynamic limit (as in Sec. II.E.5 with $k=0$) by rescaling the magnetic charge as $m\sim l^{-d}$. The solution becomes

$$y^i=x^i\sqrt{1-\frac{L^d}{r^d}}, \quad r>L,$$
$$y^i=0, \quad r\leqslant L,$$
$$r^2=\sum_i (x^i)^2. \tag{88}$$

The vortex charge $m\propto(L/l)^d$ is no longer quantized, but it is still conserved.

C. Instantons

To obtain qualitatively new solutions of gauge theory, we must move on to four Euclidean dimensions. As is well known, minima of the Euclidean action will be self-dual and anti-self-dual configurations,

$$(P^{\mp})_{kl}^{mn}F_{mn}=0, \tag{89}$$

$$(P^{\pm})_{kl}^{mn}\equiv\frac{1}{2}\delta_k^m\delta_l^n\pm\frac{1}{4}\varepsilon^{mn}{}_{kl}, \tag{90}$$

where P^+ and P^- are the projectors on self-dual and anti-self-dual tensors. These solutions are classified topologically by the instanton charge:

$$N=-\frac{1}{8\pi^2}\int \mathrm{Tr}F\wedge F. \tag{91}$$

Note that θ breaks parity symmetry, and the two types of solution will have different properties. Let θ be self-dual; then one can obtain self-dual solutions by the direct product construction mentioned in the previous subsection, while the anti-self-dual solutions turn out to be the noncommutative versions of the Yang-Mills instantons.

Instantons play a central role in the nonperturbative physics of Yang-Mills theory (Schafer and Shuryak, 1998) and have been studied from many points of view. The most powerful approach to constructing explicit solutions and their moduli space is the so-called ADHM construction (Atiyah *et al.*, 1978), which reduces this problem to auxiliary problems involving simple algebraic equations. Although when it was first proposed, this construction was considered rather recherché by physicists (Coleman, 1985), modern developments in string theory starting with Witten (1996) have placed it in a more physical context, and in recent years it has formed the basis for many practical computations in nonperturbative gauge theory, see, e.g., Dorey *et al.* (2000). We shall explain the stringy origins of the construction in Sec. VII.D.

It turns out that the ADHM construction can be adapted very readily to the noncommutative case (Nekrasov and Schwarz, 1998). Let us quote the result for anti-self-dual gauge fields, $P^+F=0$, referring to Nekrasov (2000) for proofs, further explanations, and generalizations. See also Furuuchi (2000).

To construct charge N instantons in the $U(k)$ gauge theory we must solve the following auxiliary problem involving the following finite-dimensional matrix data. Let X^i, $i=1,2,3,4$, be a set of $N\times N$ Hermitian matrices, transforming as a vector under $SO(4)$ space rotations and in the adjoint of a dual gauge group $U(N)$. Let λ_α, $\alpha=1,2$ be a Weyl spinor of $SO(4)$, transforming in (N,k) of $U(N)\times U(k)$. Instanton solutions will then be in correspondence with solutions of the following set of equations:

$$0=(P^+)_{ij}^{kl}([X^i,X^j]+\overline{\lambda}\sigma^i\overline{\sigma}^j\lambda-\theta^{ij}\mathbf{1}_{N\times N}). \tag{92}$$

These equations admit a $U(N)$ symmetry $(X^i,\lambda_\alpha)\mapsto(g_N X^i g_N^{-1},\lambda_\alpha g_N^{-1})$, and two solutions related by this symmetry lead to the same instanton solution. Thus the moduli space of instantons is the nonlinear

[3]In the references, the two types of solution are sometimes referred to as *self-dual* and *anti-self-dual*, but not consistently. It seems preferable to speak of positive flux (F aligned with θ) and negative flux.

space of solutions of Eq. (92) modulo $U(N)$ transformations. By counting parameters, one finds that it has dimension $4Nk$.

The only difference with the conventional case is the shift by θ. This eliminates solutions with $\lambda=\bar{\lambda}=0$, which would have led to a singularity in the moduli space. In conventional Yang-Mills theory, instantons have a scale size parameter which can be arbitrary (the classical theory has conformal invariance), and the singularity is associated with the zero size limit. The scale size is essentially $\rho\sim\sqrt{\bar{\lambda}\lambda}$, and in the noncommutative theory is bounded below at $\sqrt{\theta}$, again illustrating that position-space uncertainty leads to a minimal size for classical solutions.

To find the explicit instanton configuration corresponding to a solution of Eq. (92), one must solve an auxiliary linear problem: find a pair (ψ_α,ξ), with the Weyl spinor ψ_α taking values in $N\times k$ matrices over $\mathcal{A}_\theta$, and ξ being a $k\times k$ $\mathcal{A}_\theta$-valued matrix, such that

$$X^i\sigma_i\psi+\lambda\xi=0,$$

$$\psi^\dagger\psi+\xi^\dagger\xi=\mathbf{1}_{k\times k}. \tag{93}$$

Then the instanton gauge field is given by

$$A_i=\psi^\dagger\partial_i\psi+\xi^\dagger\partial_i\xi, \tag{94}$$

$$C_i=-i\theta_{ij}(\psi^\dagger x^j\psi+\xi^\dagger x^j\xi). \tag{95}$$

One can make these formulas more explicit in the $k=N=1$ case. We follow the conventions of Eq. (6). Let

$$\Lambda=\frac{\sum_\alpha \theta_\alpha(\hat{n}_\alpha+1)}{\sum_\alpha \theta_\alpha\hat{n}_\alpha},\quad D_1=C_1+iC_2,\quad D_2=C_3+iC_4.$$

Then

$$D_\alpha=\frac{1}{\sqrt{\theta_\alpha}}S\Lambda^{-1/2}\bar{z}_\alpha\Lambda^{1/2}S^\dagger,$$

where the operator S is the shift operator in Eq. (54).

This solution is nonsingular with size $\sqrt{\theta}$. One can take the limit $\theta\to0$ to return to conventional Maxwell theory, obtaining a configuration which is pure gauge except for a singularity at the origin. Whether or not this counts as an instanton depends on the underlying definition of the theory; in string theory we shall see later that it is.

D. Monopoles and monopole strings

One might at this point suspect that in general noncommutative gauge theory solitons look quite similar to their Yang-Mills counterparts, if perhaps less singular. The monopole will prove an exception to this rule.

We consider static field configurations in $3+1$ gauge theory with an adjoint scalar field ϕ, and the energy functional

$$E=\frac{1}{4g_{\rm YM}^2}\int \mathrm{Tr}\sum_{1\le i<j\le3}(F_{ij})^2+\sum_{i=1}^{3}(\nabla_i\phi)^2. \tag{96}$$

Just as in the conventional theory, one can rewrite this as a sum of a total square and a total derivative:

$$E=\frac{1}{4g_{\rm YM}^2}\int d^3x\,\mathrm{Tr}(\nabla_i\phi\pm B_i)^2\mp Z, \tag{97}$$

$$Z=\frac{1}{4g_{\rm YM}^2}\int d^3x\,\mathrm{Tr}D_i(B_i\star\phi+\phi\star B_i). \tag{98}$$

The total derivative term Z depends only on the boundary conditions and in the conventional case would have been proportional to the magnetic charge of the soliton. Minimizing the energy with fixed Z leads to $E=|Z|$ and the equations (Bogomolny, 1976)

$$\nabla_i\phi=\pm B_i. \tag{99}$$

We are only going to consider nonsingular configurations, and the first observation to make is that this implies that the diagonal part of the total magnetic charge is zero,

$$0=Q=\int d^3x\,\mathrm{Tr}[D_i,B^i],$$

by the Bianchi identity, analogous to the conventional theory.

This would seem to rule out the possibility of a nonsingular rank 1 monopole solution. Nevertheless, it turns out that such a solution exists, as was discovered by explicit construction (Gross and Nekrasov, 2000b) making use of the Nahm equations (Nahm, 1980). These equations are very analogous to the ADHM equations and admit an equally direct noncommutative generalization. They are ordinary differential equations and their analysis is somewhat intricate; we refer to Nekrasov (2000) for this and approach the solution in a more elementary way.

We start with the fact that in three spatial dimensions (we assume only spatial noncommutativity), since the Poisson tensor has at most rank 2, there will be a commutative direction. Call its coordinate x^3. This theory will admit the vortex solution of Sec. III.B, a string extending to infinity along x^3, say

$$B_3=|0\rangle\langle0|.$$

This can be used to find a simple rank 1 solution of Eq. (99) by postulating $A_3=0$ and

$$\phi=x^3|0\rangle\langle0|.$$

Now, we might not be tempted to call this solution a monopole. However, it is only one point in a moduli space of solutions. The linearized equation of motions around the solution take the form

$$D^iD_i\delta\phi=0,\quad D^iD_i\delta A_j=0,\quad D_j\delta A_j+[\phi,\delta\phi]=0.$$

Besides variations of the location of the vortex in the (x^1,x^2) plane and constant shifts of ϕ, these also include

$$\delta\phi\propto e^{-(x^3)^2/\theta}|0\rangle\langle1|.$$

Turning on this mode corresponds to splitting the string in two, as one can see by looking at the eigenvalues of the operator $\phi+\delta\phi$.

This linearized variation extends to a finite variation of the solution which can be used to send one-half of the string off to infinity. The remaining half is a rank 1 monopole attached to a *physical* analog of the Dirac string, the flux tube of Sec. III.B stretching from the monopole to infinity carrying magnetic flux and energy, and canceling the monopole magnetic flux at infinity. Its energy diverges, but precisely as the string tension Eq. (84) times the length of the string. Thus the noncommutative gauge theory has found a clever way to produce a solution despite the absence of $U(1)$ magnetic charge in three dimensions.

One can express the solution in closed form in terms of error functions (Gross and Nekrasov, 2000b). Self-dual solutions to gauge theory generically admit closed-form expressions; the deep reason for this is the integrable structure of these equations (Ablowitz and Clarkson, 1991). The noncommutative deformation respects this; for example, the noncommutative Bogomolny equations for axially symmetric monopoles are equivalent to non-Abelian Toda lattice equations.

Multimonopoles also exist, and as for the vortices, this moduli space will have limits in which it breaks up into $U(1)$ monopoles. In particular, the noncommutative analog of the 't Hooft–Polyakov monopole of the $U(2)$ theory is better thought of as a multimonopole solution, with two centers associated to the two $U(1)$ subgroups, connected by a string.

A simple observation following from the form of the noncommutative Nahm equations is that (in contrast to the instantons) this multimonopole moduli space is the same as that of the conventional theory. Since the conventional solutions and moduli space were already nonsingular, this fits with the general picture of desingularization of solitons by noncommutativity, but is still nontrivial (it was not forced by symmetry).

There are pretty string theory explanations for all of this, discussed in Sec. VII.D. In string theory, monopoles turn out naturally to be associated with D strings extending in a higher dimension. In commutative gauge theory, the string extends perpendicular to the original dimensions, and its projection on these is pointlike. Noncommutativity tilts the string in the extra dimensions, leading to a projection onto the original dimensions which is itself a string.

IV. NONCOMMUTATIVE QUANTUM FIELD THEORY

In this section we discuss the properties of quantized noncommutative field theories. We start by developing the Feynman rules and explain the role of planarity in organizing the perturbation theory (Filk, 1996).

We then give examples illustrating the basic structure of renormalization: the UV properties are controlled by the planar diagrams, while nonplanar diagrams generally lead through what is called UV/IR mixing (Minwalla *et al.*, 2000) to new IR phenomena. The limit $\theta\to 0$ in these theories is nonanalytic. We cite a range of examples in scalar and gauge theory which illustrate the possibilities; there is some physical understanding of its consequences. We also discuss the question of gauge invariance of the effective action.

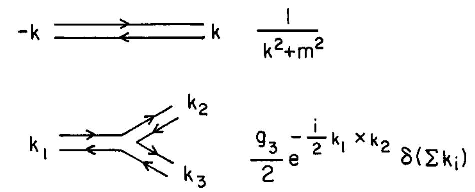

FIG. 2. Double line notation and phase factors.

We then discuss a variety of results for physical observables, including finite-temperature behavior and a universal high-energy behavior of the Wilson line operator. We also outline the Hamiltonian treatment of noncommutative gauge theory on a torus, and conclude with miscellaneous other results.

A. Feynman rules and planarity

We start by considering the theory of a single scalar field ϕ with the action

$$S=\int \mathrm{Tr}\frac{1}{2}(\partial\phi)^2-V(\phi), \quad V(\phi)\equiv\frac{m^2}{2}\phi^2+\sum_{n>2}\frac{g_n}{n!}\phi^n.$$

Functional integral quantization can be done in the standard way and leads to Feynman rules which are almost the same as for conventional scalar field theory. In particular the propagator is the same thanks to Eq. (22). The only difference is an additional momentum dependence in the interaction vertices following from Eq. (21): a ϕ^n vertex with successive incoming momenta k_μ has the phase factor

$$\exp\left(-\frac{i}{2}\sum_{1\leqslant\mu<\nu<n} k_\mu\times k_\nu\right), \tag{100}$$

where [as in Eq. (21)] $k\times k'\equiv\theta^{ij}k_i k'_j$. The same holds in the presence of derivative interactions, multiple fields, and so on.

The factor Eq. (100) is not permutation symmetric but is only cyclically symmetric. This is just as for a matrix field theory, and the same double line notation can be used to represent the choice of ordering used in a specific diagram ('t Hooft, 1974; Coleman, 1985). In this notation, propagators are represented as double lines, while the ordering of fields in a vertex is represented by connecting pairs of lines from successive propagators. See Fig. 2.

This additional structure allows defining faces in a diagram as closed single lines and thus the Euler character of a graph can be defined as $\chi=V-E+F$, with V, E, and F the numbers of vertices, edges, and faces, respectively. For connected diagrams, $\chi\leqslant 2$ with the maximum attained for planar diagrams, those which can be drawn in a plane without crossing lines. A diagram with $\chi=2-2g$ for $g\geqslant 1$ is nonplanar of genus g and can be drawn on a surface of genus g without crossings.

In matrix field theory, summing over indices provides a factor N^F, while including appropriate N dependence

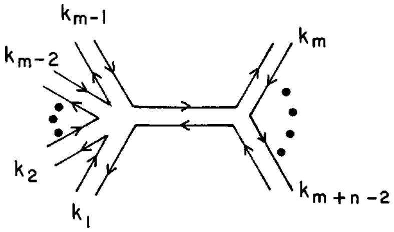

FIG. 3. The contraction $(\mathrm{Tr}\,\phi^m)(\mathrm{Tr}\,\phi^n)\leftrightarrow \mathrm{Tr}\,\phi^{m+n-2}$.

in the action (an overall prefactor of N will do it) completes this to a factor N^χ. This is the basis for the famous 't Hooft limit in which the dynamics of certain field theories is believed to reduce to that of a free string theory: as $N\to\infty$, planar diagrams dominate.

In noncommutative field theory, planarity plays an important role in organizing the phase factors Eq. (100). The basic result is that a planar diagram in noncommutative theory has the same amplitude as that of conventional theory multiplied by an overall phase factor Eq. (100), where the momenta k_i are the ordered external momenta. This can be seen by checking that the diagrams in Figs. 3 and 4 can each be replaced by a single vertex while preserving the phase factor, and that these operations can be used to reduce any planar diagram with n external legs to the single vertex $\mathrm{Tr}\phi^n$.

Since this additional phase factor is completely independent of the internal structure of the diagram, the contribution of a planar diagram to the effective action will be *the same* in noncommutative theory and in the corresponding $\theta=0$ theory (the noncommutative phase factor will be reproduced in the course of evaluating the vertex in the effective action). In particular, divergences and renormalization will be the same for the planar subsector as in a conventional theory.

We can obtain the phase factor for a nonplanar diagram by comparison with the planar case. Given a specific diagram, we choose a way to draw it in the plane, now with crossing propagators. Let $C^{\mu\nu}$ be the intersection matrix, whose $\mu\nu$ matrix element counts the number of times the μth (internal or external) line crosses the νth line (with sign; a crossing is positive as in the figure). Comparing this diagram with the corresponding planar diagram obtained by replacing each crossing with a vertex, we find that it carries the additional phase factor

$$\exp\left(-\frac{i}{2}\sum_{\mu,\nu} C^{\mu\nu}k_\mu\times k_\nu\right). \tag{101}$$

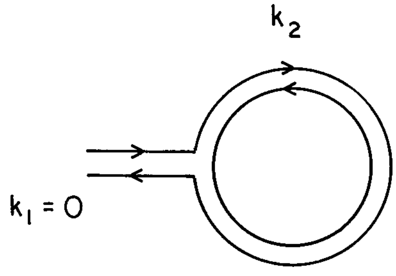

FIG. 4. Tadpoles come with no phase factor.

Although the matrix $C^{\mu\nu}$ is not uniquely determined by the diagram, the result Eq. (101) is.

Since the phase factor Eq. (101) depends on the internal structure of the diagram, nonplanar diagrams can have very different behavior from their $\theta=0$ counterparts. Since the additional factor is a phase, one would expect it to improve convergence of integrals, leading to better UV behavior. This expectation will be borne out below, leading to the principle that the leading UV divergences (in particular the beta function) come from the planar diagrams (no matter what the rank) and are thus the same as for the large-N limit of a matrix field theory.

In practice, this principle might be obscured by the presence of other divergences. A more precise statement one can make is that in the limit $\theta^{ij}\to\infty$, with fixed external momenta, UV and IR cutoffs, integrating over this phase factor would cause the general nonplanar diagram to vanish. Thus this limit would keep only planar diagrams. This observation goes back to the early works on the twisted Eguchi-Kawai models, as mentioned in the Introduction, and indeed was the main focus of interest in these works.

B. Calculation of nonplanar diagrams

The most important features are already visible in the one-loop renormalization of the scalar field theory propagator. We consider two examples, the ϕ^4 theory and the ϕ^3 theory in d Euclidean noncommutative dimensions. We shall not be careful about $\mathcal{O}(1)$ numerical factors in our discussion and often substitute generic positive real constants c, c', etc., to better make the qualitative points. We shall try to keep track of signs.

The ϕ^4 theory has two one-loop self-energy diagrams, one planar and one nonplanar. They contribute

$$\Gamma^{(2)}_{1\,planar}=\frac{g_4}{3(2\pi)^d}\int\frac{d^dk}{k^2+m^2}, \tag{102}$$

$$\Gamma^{(2)}_{1\,nonplanar}=\frac{g_4}{6(2\pi)^d}\int\frac{d^dk}{k^2+m^2}e^{ik\times p}. \tag{103}$$

The planar contribution is proportional to the one-loop diagram of the $\theta=0$ theory, and is divergent for $d\geqslant 2$. If we introduce a momentum space cutoff Λ, we shall find

$$\Gamma^{(2)}_{1\,planar}=\frac{g_4}{3(4\pi^2)^{d/2}}(\Lambda^{d-2}+\cdots),$$

a mass renormalization which can be treated by standard renormalization theory.

Compared to this, the nonplanar contribution has an oscillatory factor $e^{ik\times p}$, arising from the noncommutative nature of space-time, which will play the role of a cutoff. On general grounds, this cutoff will come in for momenta k such that $d/dk(k\times p)\sim\mathcal{O}(1)$, i.e., $k^i\sim\Lambda_p \equiv|\theta^{ij}p_j|^{-1}$, and produce

$$\Gamma^{(2)}_{1\,nonplanar}\sim g_4(c\Lambda_p^{d-2}+\cdots).$$

This cutoff goes to infinity as we take the limit $p \to 0$ of the external momentum. To get a finite result in this limit, we shall need to keep the original cutoff Λ in the calculation as well. This can be done in many ways; the usual approach is a proper time cutoff $e^{-1/\Lambda^2 t}$ in the integrand of Eq. (67) which leads to the result

$$\Lambda_p^2 \equiv \frac{1}{1/\Lambda^2 + p \circ p}, \tag{104}$$

$$p \circ p \equiv p_i \theta^{ik} \theta^{jk} p_j \tag{105}$$

(in fact this is the simplest function with the qualitative behavior we want). The notation $p \circ p$ for the quantity controlling the cutoff is standard; it is positive definite if $\theta^{0i}=0$ (purely spatial noncommutativity). If θ^{ij} has maximal rank and all its eigenvalues are $\pm\theta$, then $p \circ p = \theta^2 p^2$. If there are commutative directions as well, they will not enter into this quantity.

Adding the classical and one-loop contributions to the two-point function leads to a 1PI effective action,

$$\Gamma[\phi] = \int d^d p \, \frac{1}{2} \phi(p) \phi(-p) \Gamma^{(2)}(p) + \cdots, \tag{106}$$

$$\Gamma^{(2)}(p) = p^2 + M^2 + \frac{c g_4}{(p \circ p + 1/\Lambda^2)^{d-2}} + \cdots, \tag{107}$$

where $M^2 = m^2 + c' g_4 \Lambda^{d-2}$ is the renormalized mass. At $d=2$ the powerlike divergences become logarithmic in the usual way.

The novel feature of this result is that the limit $\Lambda \to \infty$ is finite, and the UV divergence of the nonplanar diagram has been eliminated, but only if we stay away from the IR regime $p \to 0$. The limiting theory has a new IR divergence, arising from the UV region of the momentum integration. This type of phenomenon is referred to as UV/IR mixing and is very typical of string theory, but is not possible in local field theory. However, it is possible in noncommutative theory, thanks to the nonlocality.

One way to see that this is not possible in local field theory is to observe that it contradicts the standard dogma of the renormalization group. Let us phrase this in terms of a Wilsonian effective action, defined by integrating out all modes at momentum scales above a cutoff Λ. The result is an effective action

$$S_{Wilson}[\phi;\Lambda] = \int d^4 x \, \frac{Z(\Lambda)}{2} [(\partial\phi)^2 + m^2(\Lambda)\phi^2] + \frac{g^2(\Lambda) Z^2(\Lambda)}{4!} \phi^4. \tag{108}$$

Renormalizability in this framework means that one can choose the functions $Z(\Lambda)$, $m(\Lambda)$, and $g^2(\Lambda)$ in such a way that correlation functions computed with this Lagrangian have a limit as $\Lambda \to \infty$, and correlation functions computed at finite Λ differ from their limiting values by terms of order $1/\Lambda$ for all values of momenta.

However, one sees from Eq. (107) that this is manifestly untrue of noncommutative scalar ϕ^4 theory, and indeed the generic noncommutative field theory with

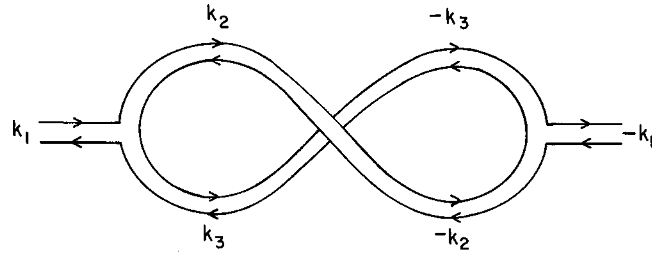

FIG. 5. A nonplanar diagram.

UV divergences. The two-point function at any finite value of Λ differs drastically from its $\Lambda \to \infty$ value, for small enough momenta $[(p\theta)^2 \ll 1/\Lambda^2]$.

The arguments above lead to the general expectation that a nonplanar diagram with UV cutoff dependence $f(\Lambda^2)$ will obtain an IR divergence $f(1/p \circ p)$ in the noncommutative theory. One might further expect that, if the conventional amplitude had no IR divergences, the leading IR divergence of the noncommutative theory would take precisely this form. We shall refer to this as the standard UV/IR relation. It is typical but not universal and in particular is violated in gauge theory.

Another example is the one-loop self-energy diagram in ϕ^3 theory (Fig. 5). The expectations we just stated are indeed valid (van Raamsdonk and Seiberg, 2000) and we find

$$\Gamma^{(2)}(p) = p^2 + M^2 + \frac{c g_3^2 M^2}{(p \circ p + 1/\Lambda^2)^{d-4}} + \cdots, \tag{109}$$

where M^2 is again the renormalized mass. At $d=4$ this becomes a logarithmic divergence.

The effects we have seen so far drastically change the IR behavior and we refer to this case as strong UV/IR mixing. There are also theories, particularly supersymmetric theories, in which the IR behavior is not qualitatively different from conventional theory, and we can speak of weak UV/IR mixing. This would include models where renormalization is logarithmic and only affects the kinetic terms (assuming the standard UV/IR relation), such as the $d=4$ Wess-Zumino model (Girotti *et al.*, 2000).

In proceeding to higher loop orders, one faces the danger that the new IR divergences will mix with other divergences in an uncontrollable way. In fact the IR divergences under discussion are rather similar to those appearing in thermal field theory, in that they give a large effective mass to low momentum modes (we shall discuss this point further below), and can be addressed by similar techniques (Fischler *et al.*, 2000; Gubser and Sondhi, 2000; Griguolo and Pietroni, 2001a).

One easy way to do this is to add and subtract a counterterm to the bare action which represents the leading divergence, so that it can be taken into account in the bare propagator. This leads to the next correction to Eq. (107) being

$$\Gamma^{(2,1)}(p) = c' g_4 \int \frac{d^d k \, (1 + e^{ik \times p})}{k^2 + M^2 + c g_4 (p \circ p)^{2-d}} - \frac{c g_4}{(p \circ p)^{d-2}}$$

with the IR divergence explicitly subtracted.

The main feature of this result is that the IR divergent term in the propagator (for $d \geqslant 2$) causes it to vanish at low momenta, and thus the resummed loop corrections are controlled in this region.

In fact, at weak coupling these corrections to $\Gamma^{(2)}(p)$ are small for all momenta. They do exhibit nonanalyticity in the coupling similar to that in finite-temperature field theory (Griguolo and Pietroni, 2001a), beginning at $\mathcal{O}(g_4^3 \ln g_4)$ in the massive theory and $\mathcal{O}(g_4^{3/2})$ in the massless theory. This brings new difficulties into the perturbation theory; one could try to work at finite coupling by using a self-consistent Hartree equation, or any of the many other approximation methods available for field theory.

In any case, this discussion appears to justify considering Eqs. (107) and (109) as valid approximations to the Green's function at weak coupling, and we discuss physics following from this idea in the next subsection. It falls short of a proof, which would require discussing (at the least) all Green's functions which obtain large UV renormalizations in conventional theory, such as the four-point function in ϕ^4 theory in $d \leqslant 4$. The argument we just gave generalizes to some extent to these diagrams as their IR corrections will be controlled by the propagator in the same way. This might fail for special values of external momenta, however, and a real proof of these ideas has not yet appeared. Relevant work includes Chepelev and Roiban (2000).

See also Chen and Wu (2001), Kinar *et al.* (2001), and Li (2000).

C. Physics of UV/IR mixing

The UV/IR mixing observed by Minwalla *et al.* (2000) appears to be the main qualitative difference between conventional and noncommutative perturbation theory. One cannot say that its full significance, and the issue of whether or not it spoils renormalizability or leads to other inconsistencies, is well understood at present. On the other hand, there is an emerging picture which we shall outline.

A reason for caution at this point is that, for the reasons we just discussed, we cannot blindly assume that the general framework of the renormalization group, which underlies most of our understanding in the conventional case, is applicable. This is not to say that it is obviously inapplicable, but that to justify it one must show how to clearly separate UV and IR phenomena, taking into account the higher loop subtleties mentioned above.

If this can be done, justifying the idea that the high-energy behavior is controlled by the planar diagrams, then the question of whether a noncommutative theory is renormalizable will have the same answer as for the corresponding conventional theory in the planar limit (normally this limit has the same UV behavior as the finite N theory). It has recently been argued by Griguolo and Pietroni (2001b) that this can be done in the Wilsonian renormalization-group approach to proving renormalizability (Polchinski, 1984). There are also arguments against this in certain cases (Akhmedov *et al.*, 2000), which we return to below.

We assume for the sake of discussion that this is indeed true and now discuss the physics of strong UV/IR mixing. Before we start, we note that the supersymmetric theories of primary interest to string theorists have weak or no UV/IR mixing, and the considerations we are about to discuss, except for those regarding unitarity, have not played a role in this context so far. On the other hand, they are likely to be important in other applications.

Once one observes new IR effects at the quantum level, the first question one must ask is whether the original perturbative vacuum $\phi = 0$ (the disordered phase, in the language of statistical field theory) is stable, or whether these effects are a signal that we are expanding around the wrong vacuum. This question could be answered if we knew the exact quantum effective action; in particular the perturbative vacuum will be at least metastable if the inverse propagator $\Gamma^{(2)}(p)$ in Eq. (106) is positive at all momenta. This includes the usual condition on the effective potential $V''(0) > 0$ but since the effects we are discussing modify the dispersion relation so drastically, we need to entertain the possibility that a phase transition could be driven by modes with nonzero momentum, perhaps leading to a stripe phase as one finds in certain condensed-matter systems.

This is a real possibility here, as is clear from Eqs. (107) and (109). One expects, and it might be possible to prove, that with certain hypotheses, the standard UV/IR relation will hold for the exact (cutoff) quantum effective action. This would mean that $\Gamma^{(2)}(p) \sim \Delta M^2(1/p\theta)$ in theories whose planar limit has a UV divergent mass renormalization $\Delta M^2(\Lambda)$.

Then the relevant question is whether this mass renormalization is positive or negative (for large Λ). If it is negative, the perturbative vacuum is clearly unstable to condensing $p \to 0$ modes. On the other hand, if it is positive, the resulting dispersion relation will make the low p modes stiff and there might or might not be a phase transition, but if there is, it will be driven by a mode with $p \neq 0$. Gubser and Sondhi (2000) have argued that such a phase transition is indeed expected (in scalar field theory in $d > 2$) and will be first order.

If we grant that the perturbative vacuum is stable, we can go on to ask about the meaning of the propagator $\Gamma^{(2)}(p)^{-1}$. This discussion will depend on whether the noncommutativity is purely spatial ($\theta_{0i} = 0$) or has a timelike component.

It is not hard to see that timelike noncommutativity combined with UV/IR mixing leads to violations of perturbative unitarity. This shows up in unphysical branch cut singularities in loop diagrams (Gomis and Mehen, 2000) and can also be seen from the behavior of the propagator and commutators of fields (Alvarez-Gaume *et al.*, 2001; see also Seiberg *et al.*, 2000a). We now give a simple argument along these lines.

A theory with timelike noncommutativity will be invariant under boosts along the conjugate spatial momentum (e.g., $p_i \propto \theta_{01}$) and therefore, if we set the other momenta to zero, the propagator will be a function of the 1+1 Lorentz invariant $p^2 = -E^2 + (p_i)^2$. Thus we can write a spectral representation

$$\frac{1}{\Gamma^{(2)}(p)} = \int_0^\infty \frac{dm^2}{p^2+m^2}\rho(m^2)$$

with $\rho(m^2) \geq 0$ in a unitary theory. Unitarity requires

$$\lim_{p^2\to 0} \frac{1}{\Gamma^{(2)}(p)} = \int_0^\infty \frac{dm^2}{m^2}\rho(m^2) > 0 \tag{110}$$

assuming this limit exists (it might also diverge). On the other hand, the general behavior produced by UV/IR mixing is

$$\lim_{p^2\to 0} \Gamma^{(2)}(p) = +\infty,$$

which is incompatible with Eq. (110).

We turn to consider purely spatial noncommutativity. This will lead to a dispersion relation of the form

$$E^2 = \vec{p}^{\,2} + m^2 + \Delta M^2\left(\frac{1}{p\,\theta}\right).$$

If $\Delta M^2(\Lambda)$ grows at infinity, we again find low-momentum modes are stiff.

We first discuss the Euclidean theory. The Green's function $G^{(2)}(p) = \Gamma^{(2)}(p)^{-1}$ can be transformed to position space following our discussion in Sec. II.E.4. As discussed there, the long-distance behavior is controlled by the pole in the upper half plane nearest the real axis, as in conventional field theory. Let us consider $\Delta M^2 = 1/4 g^2 \Lambda^2$ for definiteness, then the poles in $G^{(2)}(p)$ can be found by solving a quadratic equation, leading to the following two limiting behaviors. If $g \ll \theta m^2$ (weak coupling or strong noncommutativity), the closer root will be at $p = ig/4\theta m$. This can be interpreted as a new mode with mass $m_2 = g/4\theta m$, and the precondition is equivalent to $m_2 \ll m$, i.e., this is the limit in which the new mode dominates the long-distance behavior. In the other limit, there are two poles at equal distance $(g/8\theta)^{1/2}$, leading to oscillatory-exponential behavior.

Thus, in either case, the IR effects appear to have a sensible description at finite g, in terms of a new mode. One can go on to ask whether the new mode can be described by adding an additional field to the effective Lagrangian. This question is best discussed in the context of the connection to string theory and we postpone it for Sec. VII.E.

We now discuss the quantum noncommutative field theory in Minkowski space. The primary question is whether it is unitary. Formally, the main thing to check is that the $(\partial_0 \phi)^2$ term in the effective action is positive for all spatial momenta (we also assume that higher time derivative terms are not significant). This is necessary for perturbative unitarity of the effective action but almost certainly will be true if the analog conventional theory was renormalizable and unitary. The cutting rules can also be checked and appear compatible with unitarity in this case (Gomis and Mehen, 2000). Since these theories admit a (cutoff) Hamiltonian formulation, any problem with unitarity would almost certainly be tied to instability as well.

However, the position space Green's function in this case is quite bizarre at low momenta. Its main features can be understood by considering wave propagation in a medium with an index of refraction $n(\omega) = k(\omega)/\omega$ which grows as $\omega \to 0$ faster than $1/\omega$.[4] In our example, $\omega \sim g/\theta k$. This leads to a negative group velocity $v_g = \partial\omega(k)/\partial k$. One can still proceed formally to derive a position space Green's function, which exhibits superluminal propagation. In a non-Lorentz invariant theory, this might not be considered a major surprise; however, its short-time propagation is dominated by long wavelengths and it does not satisfy the usual defining property $\lim_{t\to 0} G(x,t) = \delta(x)$. This certainly looks unphysical and at present further interpretation would seem to be a purely academic exercise. On the other hand, perhaps the usual position space interpretation is inappropriate and there is something deep here yet to be discovered.

Finally, we return to our original assumption of the validity of the Wilsonian RG, and discuss the work of Akhmedov *et al.* (2000, 2001) and Girotti *et al.* (2001) on the noncommutative Gross-Neveu model. This is a model of N flavors of interacting massless fermions whose perturbative vacuum is unstable. In the 1+1 conventional theory a nonperturbative condensate forms, as can be shown exactly in the large N limit by solving a gap equation, leading to a vacuum with massive fermionic excitations. This theory exhibits dimensional transmutation; the bare coupling g can be eliminated in favor of the mass gap $M \sim \Lambda \exp(-1/g^2)$.

Following the procedure we just outlined for the noncommutative theory leads to $\Gamma^{(2)}(p) < 0$ at low p (for the condensate) and instability. However, Akhmedov *et al.* argue that instead of following the standard Wilsonian approach and defining the continuum theory as a flow out of a UV fixed point, one should enforce the gap equation. This leads to a stable vacuum but spoils dimensional transmutation and results in a trivial continuum limit. Akhmedov *et al.* also point out the interesting possibility of a nontrivial double scaling limit with fixed $\Lambda M \theta$.

One cannot *a priori* say that one of these definitions is correct; this depends on the underlying microscopic physics and the application. What we would insist on is that one can choose to define the theory as a flow from a UV fixed point and that in this sense the theory is renormalizable; one, however, does not know (at present) whether it has a stable vacuum.

Furthermore, we feel that even the alternate definition can be fit within the conventional renormalization-group picture, allowing for IR physics to determine a condensate which will then react back on the UV physics, determining the couplings through dynamics at vari-

---

[4]Strictly speaking, we are discussing spatial dispersion (Landau *et al.*, 1960), but the difference is not relevant for us.

ous scales. In itself, this phenomenon is not new but has fairly direct analogs in conventional field theory, particularly supersymmetric field theory where IR physics can lift a flat direction, leading to large scalar vevs and large masses for other degrees of freedom.

However, one has opened Pandora's box as the larger point seems to be that strong UV/IR mixing can lead to very different physics for condensates, which is not understood. For example, if the noncommutative Gross-Neveu model defined as a UV fixed point has a stable vacuum, it might not be translationally invariant and perhaps not even be describable by a classical condensate.

The tentative conclusion we shall draw from all of this is that quantum field theories with spatial noncommutativity and strong UV/IR mixing can be consistent, often with rather different physics from their conventional analogs.

D. Gauge theories

The usual discussion of perturbative gauge fixing and the Fadeev-Popov procedure go through essentially unchanged. The Feynman rules for noncommutative gauge theory are the standard ones for Yang-Mills theory, with the Lie algebra structure constants augmented by the phase factors Eq. (100).

Gauge theories can have IR divergences which are stronger than the standard UV/IR relation would suggest (Hayakawa, 1999; Matusis *et al.*, 2000). Consider, for example, the one-loop contribution to the noncommutative QED photon self-energy due to a massless fermion. Standard considerations lead to the amplitude

$$i\Pi_{ij}(p)=-\frac{4g^2}{(2\pi)^4}\int d^4k\,\frac{(2k_ik_j-\eta_{ij}k^2)}{k^4}(1-e^{ik\times p}).$$

The nonplanar part of this is

$$i\Pi_{ij}^{(np)}(p)=\frac{4ig^2}{(2\pi)^4}(\partial_i\partial_j-\eta_{ij}\partial^2)\ln(p\circ p+1/\Lambda^2) \quad (111)$$

$$=-\frac{16ig^2}{(2\pi)^4}\frac{\theta_{ik}p^k\theta_{jl}p^l}{(p\circ p+1/\Lambda^2)^2}. \quad (112)$$

If we remove the cutoff, this diverges as $1/\theta^2p^2$ as $p\rightarrow 0$.

Physically, this term would lead to different dispersion relations for the two photon polarizations. Peculiar as this is, it is consistent with gauge invariance in this non-Lorentz invariant theory. A similar one-loop contribution can be found to the three-point function. It behaves as $(\theta p)_i(\theta p)_j(\theta p)_k/(p\circ p)^2$ at low momenta and diverges as $1/\theta p$ as $p\rightarrow 0$. It is not clear at present what the significance of these effects might be.

In any case, it has been observed that these effects vanish in supersymmetric theories. The logarithmic effects expected from our previous discussion will still be present. A detailed discussion at one loop is given by Khoze and Travaglini (2001); see also Ruiz (2001) and Zanon (2001a).

All of the UV and IR divergences found in any computation done to date are gauge invariant, and it is generally believed that these theories are renormalizable. It has also been argued that spontaneously broken noncommutative gauge theories are renormalizable (Petriello, 2001).[5] Unitarity has recently been discussed in Bassetto *et al.* (2001).

Thus we turn our attention to the structure of the noncommutative effective action. It turns out to be somewhat subtle to write this in a gauge invariant form (Liu and Michelson, 2000; Zanon, 2001b) [this was also seen in related string computations by Garousi (2000)]. One cannot use the usual rule for Yang-Mills theory, according to which one can infer higher-point functions by completing an operator $\partial_iA_j-\partial_jA_i$ to the gauge-invariant F_{ij}. The basic problem is apparent from the analogy with conventional large-N field theory. The rules of Sec. IV.A tell us that nonplanar diagrams will produce contributions of the form

$$\int\prod_i d^dk_i\ \mathrm{Tr}O_1(k_1)\cdots\mathrm{Tr}O_n(k_n)G(k_1,\ldots,k_n).$$

Such terms do not follow the rule from Sec. II.C of associating a single trace with each integral and thus we cannot apply the arguments made there.

Clearly this problem is related to the problems with defining local gauge invariant operators discussed in Sec. II.D.2, and it is believed that the effective action will be gauge invariant if expressed in terms of the open Wilson loop operator (Liu, 2000; Mehen and Wise, 2000) This has been checked in examples (Pernici *et al.*, 2001) and one can also make strong arguments for it from string theory (see Sec. VII.E).

The basic gauge invariant local Green's function then is the two-point function of two open Wilson loops, $\langle W(k)W(-k)\rangle$. A particularly interesting question is the high-energy behavior of this quantity, which was found by Gross *et al.* (2000). Our general arguments that high-energy behavior reduced to planar diagrams applied to naive local observables, but do not apply to the open Wilson line, because its extent grows with energy.

At high energy, the open Wilson loops will become very long and the computation becomes identical to that for the expectation value of a single rectangular Wilson loop, a well-known result in the conventional theory. Since the leading large momentum behavior is given by the sum of planar ladder diagrams, the result can be applied directly. For a rectangular Wilson loop of sides L and r (with $L\gg r$), it is $\exp[-(E+M)L]$ with the Coulomb interaction energy $E=-g^2N/4\pi r$ and M a nonuniversal mass renormalization. This result can then be Fourier transformed, leading to a growing exponential in momentum space, $\exp(+\sqrt{g^2N|k|L/4\pi})$.

[5]There is some controversy about this and about the renormalizability of the related nonlinear sigma model, possibly related to issues discussed at the end of Sec. IV.C; see Campbell and Kaminsky (2001), Girotti *et al.* (2001), and Sarkar and Sathiapalan (2001).

The main adaptation required to compute the open Wilson loop Green's function is just to use Eq. (16) and take $L=|\theta k|$. This produces

$$\langle W(k)W(-k)\rangle \sim \exp\left(+\sqrt{\frac{g^2 N|k\theta||k|}{4\pi}}\right). \tag{113}$$

This exponential growth with momentum is universal and applies to correlators of Wilson loops with operator insertions $W[O](k)$ as well.

This result cannot be properly interpreted without knowing something about the higher point correlators, as the overall normalization of the operator is not a matter of great interest. Thus we consider a ratio of correlation functions in which this cancels,

$$\frac{\langle W(k_1)W(k_2)\cdots W(k_n)\rangle}{\sqrt{\langle W(k_1)W(-k_1)\rangle\cdots\langle W(k_n)W(-k_n)\rangle}}.$$

Gross *et al.* (2000) argue that generic higher point functions do not share the exponential growth, essentially because more than two Wilson lines will not generally be parallel and thus the dependence on L will not exponentiate.

Thus Eq. (113) should be interpreted as leading to a universal exponential falloff of correlation functions at large momentum. This is another strong analogy between noncommutative gauge theory and string theory, as it is a very characteristic feature of string theory [however, see the caution in Gross *et al.* (2000) about this analogy].

Rozali and Van Raamsdonk (2000) have further argued that in a multipoint function in which a pair of momenta become antiparallel (within a critical angle $\phi < \phi_0 \sim 1/|k_i||\theta k_j|$), the exponential growth is restored, which would again be consistent with field theory behavior. See also Dhar and Kitazawa (2001a).

The main motivation for this computation was to check it against the AdS/CFT dual supergravity theory (Aharony *et al.*, 2000), and we discuss this aspect in Sec. VII.F.

Loop equations governing these correlators are discussed in Abou-Zeid and Dorn (2001b) and Dhar and Kitazawa (2001b)

E. Finite temperature

Another limit which will probe noncommutativity is the high-temperature limit $\theta T^2 \gg 1$. We now discuss this regime, following Arcioni and Vazquez-Mozo (2000) and Fischler, Gomis, *et al.* (2000; Fischler, Gorbatov, *et al.* 2000) and using standard techniques of finite-temperature field theory (Kapusta, 1990). The temperature is T and $\beta=1/T$, and we take $\theta_{12}=\theta$ and x^3 and time commutative.

The leading nonplanar diagram contributing to the finite-temperature free energy of ϕ^4 theory is the two-loop diagram obtained by contracting the external legs in Fig. 5. This contributes (for $d=4$)

$$-g^2T^2\sum_{n,l=-\infty}^{\infty}\int\frac{d^3p}{(2\pi)^3}\frac{d^3k}{(2\pi)^3} \times e^{ip\times k}\frac{1}{\left(\frac{4\pi^2}{\beta^2}n^2+p^2+m^2\right)\left(\frac{4\pi^2}{\beta^2}l^2+k^2+m^2\right)}.$$

If we neglect the mass (appropriate if $m \gg T$), this can be reduced using standard techniques to

$$-g^2T\int\frac{d^3p}{(2\pi)^3}\frac{1+2n_\beta(|p|)}{2|p|}\frac{1+2n_{1/\beta}(2\pi|\theta p|)}{4\pi|\theta p|},$$

where $n_\beta(|p|)=(e^{\beta|p|}-1)^{-1}$ is a Bose distribution at temperature $\beta=1/T$.

This expression shows intriguing similarities to string theory, and Fischler, Gomis, *et al.* (2000) interpret the second thermal distribution as describing "winding states"; see also Arcioni *et al.* (2000). It also leads to an IR divergence which must be addressed, either along the lines of the resummation discussed above, or by considering supersymmetric theories with better convergence properties.

In either case, one finds the following fascinating behavior for the the nonplanar contribution to the free energy. While the regime $\theta T^2 \ll 1$ is essentially as for conventional field theory, results for the regime $\theta T^2 \gg 1$ are very much as if the theory had many fewer degrees of freedom in the UV than conventional field theory. For ϕ^4 theory, supersymmetric ϕ^4 theory (the Wess-Zumino model) and $N=4$ NCSYM (all in $D=4$), one finds (at two loops)

$$\frac{F_{nc}}{V}\sim -\frac{g^2}{\theta^2}T^2\theta\ln T^2\theta. \tag{114}$$

In ϕ^3 theory in $D=6$, one finds

$$\frac{F_{nc}}{V}\sim -\frac{g^2}{\theta^2}T^2\theta. \tag{115}$$

This latter result is even more amusing as it can be derived from *classical* statistical mechanics. This is essentially the same as quantum field theory in one lower dimension and leads to an integral

$$F\sim T^2\int d^5k\,d^5p\,\frac{e^{ik\times p}}{k^2p^2(k+p)^2}.$$

While normally such integrals are badly UV divergent (the famous ultraviolet catastrophe of classical physics), this one actually converges and reproduces Eq. (115). This would seem a very concrete demonstration of the idea that the number of UV degrees of freedom has been drastically reduced, presumably to a finite number per $\theta^{d/2}$ volume or Moyal cell in noncommutative space.

However, the more general situation is that the corresponding classical integrals are still UV divergent. In particular, this is true in the cases summarized in Eq. (114). It must also be remembered that the planar contributions are still present and dominate the contributions we are discussing. Thus the full import of these rather striking results is not clear.

See also Landsteiner *et al.* (2001), who argue for a finite-temperature phase transition in noncommutative gauge theory.

F. Canonical formulation

In general terms, canonical quantization of noncommutative field theory and gauge theory with purely spatial noncommutativity can be done by following standard procedures, which we assume are familiar. This leads to commutation relations between free fields, say ϕ and a conjugate momentum π, which in momentum space take the same form as conventional theory. One furthermore has standard expressions for energy and momentum operators associated to commutative dimensions. Possible quantum corrections to their Ward identities have not been much studied.

One can define momentum operators P_i for the noncommutative dimensions by using the restricted energy-momentum tensor $T^0=i[\phi,\pi]$, as in Eq. (26). In noncommutative gauge theory, this makes sense for the reasons discussed in Sec. II.D.3.

A point where interesting differences from the conventional case appear is in discussing the action of gauge transformations and charge quantization. Consider 2 +1 Maxwell theory on a square torus T^2 defined by the identifications $x^i\cong x^i+2\pi$. The total electric flux has two components $\int d^2x E^i$, each conjugate to a Wilson loop $W_i=\exp(i\int dx^i A_i)$:

$$[E^i,W_j]=\delta^i_j W_i .$$

This theory admits large gauge transformations, acting on a charge 1 matter field ψ as $\psi\rightarrow e^{in_ix^i}\psi$, with $n_i\in\mathbb{Z}$ to make the transformation single valued. Thus the zero mode of A_i is a periodic variable, and eigenvalues of E^i must be integrally quantized.

A similar argument can be made in the rank 1 noncommutative theory on $\mathbf{T}^d_\theta$, but now the corresponding gauge transformations $\psi\rightarrow U_i\psi U_i^\dagger$ (notation as in Sec. II.B.2) act on the nonzero modes of ψ as well as in Eq. (15), taking $\psi(x^j)\rightarrow\psi(x^j-\theta^{ji})$. Thus the wave function must be invariant under a simultaneous shift of the zero mode of A_i and an overall translation (Hofman and Verlinde, 1999), leading instead to a quantization law for the quantity

$$\int \mathrm{Tr}E^i-\theta^{ij}P_j . \tag{116}$$

One also finds modified quantization laws for magnetic charges, which we shall discuss in Sec. VI.D.

Taking into account these considerations, one can find the quantum Hamiltonian and its perturbative spectrum along standard lines. We quote as an example the ground-state energy in 2+1 noncommutative Yang-Mills on $\mathrm{Mat}_p(\mathbf{T}^2_\theta)$ at leading order in perturbation theory (this can be shown to be the exact result given enough supersymmetry). Sectors of this theory are labeled by conserved integral quantum numbers n^i [related to electric flux as in Eq. (116)], q (magnetic flux), and m^i (momentum). Finally, we write $N=\mathrm{Tr}1=p-\theta q$; as we explain in Sec. VI, the naive expression $\mathrm{Tr}1=p$ is not valid on $\mathbf{T}_\theta$.

One then has (Connes *et al.*, 1998; Brace and Morariu, 1999; Hofman and Verlinde, 1999; Konechny and Schwartz, 1999)

$$E=\frac{g_Y^2M}{2\sqrt{g}N}g_{ij}(n^i+\theta^{ik}m_k)g_{ij}(n^j+\theta^{jl}m_l) + \frac{\pi^2}{2g_Y^2M\sqrt{g}N}q^2+\frac{2\pi}{N}|m_ip-q\epsilon_{ij}n^j|. \tag{117}$$

The first two terms are the energies associated to zero modes, while the last term is the energy $E=|p|$ associated to a state with massless excitations carrying momentum m_i as well as a contribution from $E\times B$.

G. Other results

Unfortunately, space does not permit us to discuss all of the interesting results obtained in quantum noncommutative field theory, but let us mention a few more.

The Seiberg-Witten solution of $\mathcal{N}=2$ supersymmetric Yang-Mills (Seiberg and Witten, 1994), giving a prepotential which encodes the dependence of BPS masses and low-energy couplings on the choice of vacuum, is a benchmark for nonperturbative studies. It turns out that the solution is the same for noncommutative as for conventional theory; this can be seen from instanton computations (Hollwood *et al.*, 2001) and M theory considerations (Armoni *et al.*, 2001; see also Bellisai *et al.*, 2000).

One may wonder what happened to the UV/IR mixing. The noncommutative theory necessarily includes a $U(1)$ sector, and one can show that only this sector is affected. It is not visible in the prepotential, which is independent of the diagonal component of the scalar vev. This sector (and the rank 1 theory) is nontrivial, and does not reduce to Maxwell theory as $\theta\rightarrow 0$ (Armoni, 2001). It is not understood at present.

Noncommutative sigma models are discussed by Lee *et al.* (2001). The noncommutative Wess-Zumino-Witten model is discussed in Lugo (2000) and Moreno and Schaposnik (2000, 2001).

Anomalies have been studied in Ardalan and Sadooghi (2000), Bonora, Schnabl, and Tomasiello (2000), Gracia-Bondia and Martin (2000), and Martin (2001). Gauge anomalies appear to be directly analogous to those of conventional theory, consistent with their topological origin, and can be described using fairly straightforward noncommutative generalizations of the Wess-Zumino consistency conditions, descent equations, etc. Somewhat surprisingly, however, these formulas appear to lead to more restrictive conditions for anomaly cancellation, because more invariants appear. For example, while the conventional $d=4$ triangle anomaly only involves the invariant $d^{abc}=\mathrm{tr}t^a\{t^b,t^c\}$, which can cancel between different representations, its noncommutative generalization involves $\mathrm{tr}t^at^bt^c$, which cannot. This apparent contradiction to the general similarity we have seen between noncommutative and conventional quantum effects as well as to the topological nature of the anomaly deserves further study.

V. APPLICATIONS TO THE QUANTUM HALL EFFECT

A physical context leading directly to noncommutativity in the Fock space basis is the dynamics of electrons in a constant magnetic field $\vec{B}$, projected to the lowest Landau level (LLL). This is well known in the theory of the quantum Hall effect (QHE) (Prange and Girvin, 1987; Girvin, 1999), and we summarize the basic idea here.

More recently, it has been proposed that a good description of the fractional quantum Hall effect (FQHE) can be obtained using noncommutative rank 1 Chern-Simons theory (Polychronakos, 2001; Susskind, 2001) and we give a brief introduction to these ideas as well.

It seems fair to say that so far, the noncommutative framework has mainly provided a new language for previously known results. However, it also seems fair to say that this formulation connects these problems to a very large body of field theory results whose potential relevance had not been realized, and one can hope that these connections will prove fruitful.

A. The lowest Landau level

The Lagrangian of a system of interacting electrons in two dimensions, subject to a perpendicular magnetic field, is

$$L=\sum_{\mu=1}^{N_e}\frac{1}{2}m_e\dot{\vec{x}}_\mu^2-\frac{ieB}{2c}\varepsilon_{ij}x_\mu^i\dot{x}_\mu^j+V(\vec{x}_\mu) \tag{118}$$

$$+\sum_{\mu<\nu}U(\vec{x}_\mu-\vec{x}_\nu). \tag{119}$$

Defining a projection operator P to the first Landau level for each electron, one finds that the projected coordinates Px_μ^iP do not commute, but instead satisfy

$$[x_\mu^i,x_\nu^j]=i\delta_{\mu\nu}\varepsilon^{ij}\frac{\hbar c}{eB}\equiv i\delta_{\mu\nu}\theta^{ij}. \tag{120}$$

Heuristically, this is because in the limit of strong magnetic field one can neglect the kinetic term, i.e., formally put $m_e=0$. The resulting Lagrangian is first order in time derivatives, turning the original coordinate space into an effective phase space defined by Eq. (120). A more precise argument can be made by taking coordinates $\zeta,\overline{z}$ as in Sec. II.B and showing that

$$P\overline{z}P=\theta\partial/\partial z+z \tag{121}$$

acting on LLL states

$$|k\rangle=\frac{1}{\sqrt{k!}}z^ke^{-zz^*/2\theta}. \tag{122}$$

The resulting single-particle Hamiltonian has been much studied and we refer to Bellissard *et al.* (1993) for a discussion of the uses of noncommutative geometry in this context.

To obtain a field theory, one introduces electron creation and annihilation operators $\psi^\dagger(x)$ and $\psi(x)$, in terms of which the electron density is

$$\rho(\vec{x})=\frac{1}{\sqrt{N_e}}\sum_\mu\ \delta^2(\vec{x}_\mu-\vec{x})\equiv\psi^\dagger(x)\psi(x). \tag{123}$$

The single-particle and pairwise interaction Hamiltonians then become

$$H_V=\int dx V(x)\rho(x), \tag{124}$$

$$H_U=\int dx dx' U(x-x')\rho(x)\rho(x'). \tag{125}$$

The effects of truncation to the LLL are now expressed by noncommutativity, and more specifically enter when we use the star product to compute the commutators of density operators. In momentum space,

$$\rho(\vec{q})=\int e^{iqx}\rho(x),$$

we have

$$[\rho(\vec{q}),\rho(\vec{q}\,')]=\sin i\vec{q}\times_\theta\vec{q}\,'\rho(\vec{q}+\vec{q}\,')$$

which leads to deformed equations of motion, etc.

This type of description has been used by many authors; we cite Sinova *et al.* (2000) on localization in quantum Hall states and Gurarie and Zee (2000) on its classical limit, as the tip of a large iceberg.

B. The fractional quantum Hall effect

The generally accepted explanation of the FQHE is that interactions lead to a state similar to the filled LLL but allowing fractionally charged quasiparticle excitations. A good microscopic description of such a state is provided by the N-electron wave function (Laughlin, 1983)

$$\Psi=\prod_{\mu<\nu}(z_\mu-z_\nu)^m\exp\left(-\sum_\mu\ z_\mu z_\mu^*/2\theta\right), \tag{126}$$

where m is an odd integer. This state has charge density $1/m$ that of a filled Landau level and was argued by Laughlin to be a ground state of Eq. (118), at least for small m.

States with quasiparticles are obtained by simple modifications of this. An operator creating a quasihole at z_0 acts as multiplication by $\Pi_i(z_i-z_0)$, while the conjugate operator (which can create quasiparticles) is $\Pi_i(\overline{z_i}-z_0^*)$ with $\overline{z_i}$ as in Eq. (121). Acting m times with one of these operators adds or subtracts a particle, so the quasiparticles have charge $1/m$.

A more subtle property of the quasiparticles is that they satisfy fractional statistics; a 2π rotation of a state of two quasiparticles produces a phase $\exp(2\pi i/m)$, as can be seen by a Berry phase argument (Arovas *et al.*, 1984).

The low-energy excitations of this ground state can also be described by a Landau-Ginzburg theory of a superfluid density ϕ and a fictitious vector potential A, such as Eq. (82) (Kallin and Halperin, 1984; Girvin *et al.*, 1985; Laughlin, 1985). See also Fradkin (1991). In this picture, the original quasiparticles are vortex solutions, and their fractional statistics is reproduced by an Abelian Chern-Simons term in the action

$$S_{CS}=\frac{im}{2\pi}\int \epsilon^{ijk}A_i\partial_j A_k. \tag{127}$$

With this term, a vortex with unit magnetic charge will also carry electric charge $1/m$, so the Aharonov-Bohm effect will lead to fractional statistics (Wilczek, 1982).

Recently Susskind (2001) has proposed that noncommutative Chern-Simons theory is a better description of fractional quantum Hall states, which can reproduce the detailed properties of these quasiparticles. Perhaps the simplest argument one could give for this is simply to combine the arguments leading to Landau-Ginzburg theory with the arguments leading to noncommutativity in a magnetic field.

One way to make this claim precise has been proposed by Polychronakos (2001). One first observes that Eq. (126) (considered as a function of one-dimensional positions $x_\mu=\mathrm{Re}\,z_\mu$) is the ground state for a Calogero model, defined by the quantum-mechanical Hamiltonian

$$H=\sum_\mu \frac{1}{2}p_\mu^2+\frac{1}{2\theta^2}x_\mu^2+\frac{1}{2}\sum_{\mu<\lambda}\frac{m(m+1)}{(x_\mu-x_\lambda)^2}. \tag{128}$$

One can continue and identify all the excited Calogero states with excited Laughlin wave functions (Hellerman and Van Raamsdonk, 2001).

It is furthermore known (Olshanetsky and Perelomov, 1976) that this model can be obtained from a matrix-vector $U(N)$ gauged quantum mechanics, with action (Polychronakos, 1991)

$$S=\int dt\ \mathrm{Tr}\left(\epsilon_{ij}X^iD_0X^j-\frac{1}{2\theta^2}X_i^2+2mA_0\right) + \psi^\dagger D_0\psi, \tag{129}$$

where X^i are Hermitian matrices with $i=1,2$ and ψ is a complex vector. The gauge field in this action is nondynamical but enforces a constraint which selects the sector with ψ charge m, for which the Hamiltonian is Eq. (128).

Reversing the procedure which led to Eq. (38), we can regard X^i as covariant derivative operators, to obtain an noncommutative gauge theory action, whose kinetic term [coming from the first term of Eq. (129)] is precisely Eq. (127). The other terms are secondary: the X^2 term localizes the state in space, while ψ, although required for consistency at finite N, plays no dynamical role.

In this sense, one has a precise noncommutative field theory description of the fractional Hall state. In particular, the quasiparticles are well-defined excitations of the noncommutative gauge field; for example, the quasihole is rather similar to the fluxon Eq. (83). We refer to the cited references for more details.

VI. MATHEMATICAL ASPECTS

As we mentioned in the Introduction, noncommutative gauge theory was first clearly formulated by mathematicians to address questions in noncommutative geometry. Limitations on length would not permit more than the most cursory introduction to this subject here, and since so many introductions are already available, starting with the excellent Connes (1994), much of which is quite readable by physicists, and including Connes (1995, 2000a, 2000b), Douglas (1999), Gracia-Bondia *et al.* (2001), as well as the reviews cited in the Introduction, we shall content ourselves with a definition: Noncommutative geometry is a branch of mathematics which attempts to generalize the notions of geometry, broadly defined, from spaces M whose function algebras $C(M)$ are commutative, to "spaces" associated to general algebras. The word "space" is in quotes here to emphasize that there is no *a priori* assumption that these spaces are similar to manifolds; all of their attributes emerge through the course of formulating and studying these geometric notions.

The remainder of this section provides an introduction to other examples of noncommutative spaces, and other topics for which a more mathematical point of view is advantageous. We shall only be able to discuss a few aspects of noncommutative geometry with direct relevance for the physics we described; there are many others for which such a role may await, or for which we simply lack the expertise to do them justice. Among them are cyclic cohomology and the related index theorems (Connes, 1982), and the concept of spectral triple (Connes, 1996).

A. Operator algebraic aspects

The origins of noncommutative geometry are in the theory of operator algebras, which grew out of functional analysis. Certain issues in noncommutative field theory, especially the analogs of topological questions in conventional field theory, cannot be understood without these ideas. The following is loosely inspired by a discussion of the meaning of the instanton charge in Schwarz (2001); we also discuss a proposed definition of the noncommutative gauge group (Harvey, 2001b).

A good example of a topological quantity is the total magnetic flux in two spatial dimensions, $\int\mathrm{Tr}F$. In commutative theory, this is the integral of a total derivative $F=dA$. In a pure gauge background, this integral will be quantized; furthermore, it cannot change under variations $A+\delta A$ by functions δA which are continuous, single-valued, and fall off faster than $1/r$ at infinity. Thus one can argue with hardly any dynamical input that the total flux in a sufficiently large region must be conserved.

Let us return to the question raised in Sec. III.B, of whether this argument can be generalized to noncommutative theory. The magnetic flux has the same formal expression. In operator language, it is the trace of a commutator,

$$\mathrm{Tr}F = \mathrm{Tr}[C,\bar{C}] + 1. \tag{130}$$

We need to understand in what sense this is a boundary term, which is preserved under continuous variation of the fields with a suitable falloff condition.

If one does not try to define these conditions, one can easily exhibit counterexamples, such as the path

$$\bar{C} = \lambda z + (1-\lambda) S^\dagger z S \tag{131}$$

which purports to continuously interpolate between the fluxon and the vacuum. Is this a continuous variation? We have to decide; there is no single best definition of continuous in this context.

The minimal criterion we could use for a continuous variation of the fields is to only allow variations by bounded operators. A bounded operator has finite operator norm $|A|$ [this is the largest eigenvalue of $(A^\dagger A)^{1/2}$]. A small variation C' of the connection C then has small $|C'-C|$, so will be a bounded operator times a small coefficient. Boundedness is mathematically a very natural condition as it is the weakest condition we can impose on a class of operators which guarantees that the product of any two will exist, and thus the definition of an operator algebra normally includes this condition.

However, boundedness is not a strong enough condition to force traces of commutators to vanish: one can have $\mathrm{Tr}[A,B] \neq 0$ even if both A and B are bounded operators (e.g., consider $\mathrm{Tr}[S^\dagger,S]$ where S is the shift operator). The operator $O = z - S^\dagger z S$ appearing in Eq. (131) is bounded.

A condition which does guarantee $\mathrm{Tr}[A,B]=0$ is for A to be bounded and B to be trace class, roughly meaning that its eigenvalues form an absolutely convergent series. More precisely, A is trace class if $|A|_1$ is finite, where $|A|_p = [\mathrm{Tr}(A^\dagger A)^{p/2}]^{1/p}$ is the Schatten p norm. More generally, the p-summable operators are those for which $|A|_p$ is finite. This is more or less the direct analog of the conventional condition that a function (or some power of it) be integrable. A related condition which also expresses falloff at infinity is for A to be compact, meaning that the sequence of eigenvalues of $A^\dagger A$ has the limit zero.

Although these are important conditions, they do not solve the problem at hand, because they are not preserved by the derivatives Eq. (3), so do not give strong constraints on Eq. (130). In particular, one can have $\mathrm{Tr}[z,K] \neq 0$ (and even ill defined) for K in any of the classes above (e.g., try $\mathrm{Tr}[O,\bar{z}]$).

Another approach is to adapt a falloff condition on functions on $\mathbb{R}^n$ to $\mathbb{R}^n_\theta$ by placing the same falloff condition on the symbol Eq. (18). An important point to realize about the noncommutative case is that one cannot separately define growth at large radius from growth with large momentum. The only obvious criterion one can use is the asymptotics of matrix elements at large mode number, which does not distinguish the two. Physically, as we saw in our considerations of Gaussians in Sec. II.E.2, interactions easily convert one into the other. Thus questions about whether configurations disperse or go off to large distance are hard to distinguish from questions about the possibility of forming singularities.

Thus, useful falloff conditions must apply to both position and momenta. Such conditions are standard in the theory of pseudodifferential operators, and can be used to define various algebras which are closed under the star product (Schwarz, 2001). Perhaps the simplest is the class of operators $S(\mathbb{R}^d_\theta)$, which can be obtained as the transform Eq. (18) of smooth functions with rapid decrease, i.e., which fall off both in position and momentum space faster than any power. This condition is very much stronger than boundedness (roughly, it requires matrix elements in the Fock basis to fall off exponentially) and is preserved under differentiation, so $\mathrm{Tr}F$ will certainly be preserved by this type of perturbation.

As we discussed in Sec. III.B, the question of physical flux conservation is whether time evolution allows flux to get to infinity at finite time. This is a dynamical question, but might be addressed by finding conditions on the fields which are preserved under time evolution. For example, in the conventional case one can take fields which are pure gauge outside a radius $r=R$, and then causality guarantees that they will stay pure gauge outside of $r=R+t$. Since $S(\mathbb{R}^d_\theta)$ is preserved by both the product and the Laplacian, it is a good candidate for an analogous class of noncommutative fields which is preserved under time evolution. The fluxon and its perturbations by finitely many modes fall into this class, so an argument along these lines should show that flux is conserved, and thus the decay of the fluxon does not lead in finite time to a stable ground state. Of course, conservation of $\mathrm{Tr}F$ (and energy $\mathrm{Tr}F^2$) is surely true for a larger class of initial conditions.

A related question, discussed by Harvey (2001b), is the precise definition of the gauge group $U(\mathcal{H})$. In conventional gauge theory, the topology of the group $\mathcal{G}_0$ of gauge transformations which approach the identity at infinity is directly related to that of the configuration space of gauge fields (connections modulo gauge transformations) by a standard argument: since the space of connections $\mathcal{A}$ is contractible, we have $\pi_n(\mathcal{A}/\mathcal{G}_0) \cong \pi_{n-1}(\mathcal{G}_0)$. One might expect the topology of $U(\mathcal{H})$ to play an analogous role and in fairly direct correspondence to the considerations above, this would imply that we cannot identify $U(\mathcal{H})$ with the group of unitary bounded operators on Hilbert space; this group is contractable (Kuiper, 1965). One can get the expected nontrivial topology, $\pi_k(G) = \lim_{N\to\infty} \pi_k[U(N)] = \mathbb{Z}$ for $k = 2n+1$, by using the unitaries $1+K$ with K compact or p summable (Palais, 1965). Harvey (2001b) suggests using K compact, as the largest of these groups. Alternatively, one might try to use a smaller group defined by imposing conditions involving the derivatives.

A similar open question about which less is known at present is that of what class of fields to integrate over in the functional integral (the functional measure). As we commented in Sec. II.D.1, formally all pure bosonic noncommutative gauge theories have the same action; however, we expect that different quantum theories exist, distinguished by the choice of measure. In particular, this subsumes the choice of the dimension of space-time.

Again, the minimal proposal is that one uses an action such as Eq. (38), expands around a pure gauge configuration $C_i = \overline{z^i}$ and $\overline{C_i} = z^i$ for $1 \leq i \leq r$, and integrates over all variations of this by bounded operators. This type of prescription does make sense in the context of the perturbation theory of Sec. IV with IR and UV cutoffs, and would specify the dimension of space-time there. This dimension can also be inferred from the nature of one-loop divergences; see Connes (1994) and Varilly and Gracia-Bondia (1999)

For the reasons we discussed above, this prescription probably does not make sense beyond perturbation theory. One might address this by proposing a smaller class of variations to integrate over. At our present level of understanding, however, it may be better to work with an explicit cutoff, such as that provided by the large-N limit of matrix approximations as discussed in Sec. II.E.6. We described there the proposal of Ambjorn *et al.* (1999) along these lines; while concrete, this appears to share the problems of conventional lattice definitions of breaking the symmetries of the continuum theory, and of not admitting supersymmetric generalizations. Thus the problem of finding the best regulated form of noncommutative field theory remains open.

Given such a regulated theory, one then wants to study the continuum limit. Of course the analogous questions in conventional quantum field theory are not trivial, and their proper understanding requires the ideas of the renormalization group. We already made such a discussion in Sec. IV, following the usual paradigm in which the momentum-space behavior of Green's functions is central. It is not yet clear that this is the best paradigm for noncommutative quantum theory; perhaps other classes of fields such as those discussed here will turn out to be equally or more useful.

B. Other noncommutative spaces

Our discussion so far was limited to noncommutative field theory on $\mathbb{R}^d_\theta$, mostly because it and $\mathbf{T}^d_\theta$ are the only examples for which field-theoretic physics has been explored sufficiently at present to make it worth writing a review. There are many other examples which allow defining noncommutative field theories, whose physics has been less explored.

For example, let us try to define a noncommutative sphere S^d_θ. This should be a space associated to an algebra of $d+1$ operators x^i, $1 \leq i \leq d+1$, satisfying the relation

$$\sum_{i=1}^{d+1} (x^i)^2 = R^2 \ \mathbf{1}. \tag{132}$$

We next need to postulate some analog of Eq. (1). Although one can define algebras without imposing commutation relations for each pair of variables, these will be much larger than the algebra of functions on S^d. For $d=2$, there is a natural choice to make,

$$[x^i, x^j] = i\theta\epsilon^{ijk} x^k,$$

which preserves $SO(3)$ symmetry.

In fact, there is a very simple way to define such an algebra, called the *fuzzy two-sphere*, discussed in depth by Madore (1992), in Madore (1999), and many other works. It is to consider the $2j+1$-dimensional irreducible representation of the $SU(2)$ Lie algebra, defined by Hermitian generators t^i satisfying the commutation relations $[t^i, t^j] = i\epsilon^{ijk} t^k$. One can then set

$$x^i = \frac{R}{\sqrt{j(j+1)}} t^i$$

from which Eq. (132) follows easily, and one finds $\theta = R/\sqrt{j(j+1)}$.

This algebra can serve as another starting point for defining noncommutative field theory. One cannot use Eq. (3) to define the derivatives, however, as this is inconsistent with Eq. (132). A simple choice which works is to define a linearly dependent set of derivatives $\partial_i f = [X^i, f]$ and use $g_{ij} = \delta_{ij}$ in this basis as a metric. Since these derivatives do not commute, the natural definition of curvature becomes

$$F_{ij} = i[\partial_i + A_i, \partial_j + A_j] - i[\partial_i, \partial_j]$$

[as in Eq. (37)], in terms of which one can again use the Yang-Mills action. This and related theories have been discussed by Grosse, Klimčik, and Presnajder (1996).

Unlike $\mathbb{R}^d_\theta$ and $\mathbf{T}^d_\theta$, this algebra is finite dimensional. If we base our theory on it directly, it will have a finite number of degrees of freedom, and one might question the use of the term field theory to describe it.

One can certainly take the limit $j \to \infty$, but if we take $R \propto j \to \infty$ as well to keep θ finite, we lose the curvature of the sphere, and end up with $\mathbb{R}^2_\theta$.

Finally, if we keep R fixed and take $j \to \infty$, it is a theorem that (with suitable definitions) this algebra goes over in the limit to the algebra of continuous functions on ordinary S^2, so we obtain a conventional field theory. This feature of restoring commutativity in the limit applies to a wide class of constructions, as we shall discuss further in Sec. VI.G. It should be said that this theorem is classical, and there might be a way to quantize the theory which does not commute with the large j limit, leading to a nontrivial noncommutative quantum field theory.

One can still argue that for large but finite j, a theory based on this algebra deserves the name *noncommutative field theory*. We would suggest that the nomenclature be based, first, on the extent to which a theory displays physical characteristics similar to those we have seen for theories on $\mathbb{R}^d_\theta$, and second on the extent to which it shows universality (e.g., has finitely many parameters) analogous to field theory; this is not usually the case for constructions with a finite number of degrees of freedom. These questions have not been settled and we shall not take a position on this here.

We move on and discuss other noncommutative algebras comparable to function algebras which can clearly serve to define field theories. The simplest possibility was of course $\mathrm{Mat}_n[C(M)]$. Interesting variations on this can be obtained by imposing further conditions which respect the product. For example, one could consider an algebra and a $\mathbb{Z}_2$ automorphism preserving the metric, call it g. A simple example would be to take $\mathbb{R}^d_\theta$ and let g be the reflection about some hyperplane. One could then impose $a^\dagger = g(a)$ where g acts on each matrix element. This idea has been used to propose noncommutative gauge theories with analogy to the other classical (SO and Sp) Lie algebras (Bonora, Schnabl, *et al.*, 2000; Bars, Sheikh-Jabbari, and Vasiliev, 2001).

One can go on in this vein, using nonsimple finite algebras and more complicated automorphisms. Indeed, one can obtain the complete action for the standard model by choosing the appropriate algebra (Connes and Lott, 1991). Obviously the significance of this observation is for the future to judge, though any example of a formalism which only describes a subset of all possible gauge theories but can lead to the standard model probably has something important to teach us.

C. Group algebras and noncommutative quotients

A large class of more noncommutative algebras are provided by the group algebras. Given a group G, we define $\mathcal{A}_G$ to be the algebra of all linear combinations of elements of G, with multiplication law inherited from G. For example, consider $G = \mathbb{Z}_2$ with elements 1 and g satisfying $g^2 = 1$. The general element of $\mathcal{A}_G$ is $a + bg$, and $(a+bg)(c+dg) = (ac+bd)+(bc+ad)g$.

It is well known that for a finite group G, $\mathcal{A}_G$ is a direct sum of matrix algebras, one for each irreducible representation of G. One goal of representation theory is to try to make analogous statements for infinite groups. This requires being more precise about the particular linear combinations allowed, and leads one deep into the theory of operator algebras. We shall not go into detail, but these algebras are clearly a good source of noncommutative field theories, as the original definitions of Connes and Rieffel (1987) can be applied directly to this case. One of the main problems in trying to define noncommutative field theories on more general spaces is to either define a concept of metric, or get away without one. Since group spaces are homogeneous spaces, this problem becomes very much simpler.

A variation on this construction is the twisted group algebra $\mathcal{A}_{G,\epsilon}$, which can be defined if $H^2[G,U(1)] \neq 0$. This allows for nontrivial projective representations γ characterized by a two-cocycle ϵ,

$$\gamma(g_1)\gamma(g_2) = \epsilon(g_1,g_2)\gamma(g_1 g_2), \tag{133}$$

and $\mathcal{A}_{G,\epsilon}$ is just the group algebra with this multiplication law. Since the phases Eq. (21) are a two-cocycle for $\mathbb{Z}^d$, $\mathbf{T}^d_\theta$ itself is an example.

A very important source of noncommutative algebras is the *crossed product* construction. One starts with an algebra $\mathcal{A}$, with a group G acting on it, say on the left:

$$a \to U(g)a.$$

One then chooses a representation R of G acting on a Hilbert space H, $g \mapsto \gamma(g) \in \mathcal{B}(H)$, considers the tensor product $\mathcal{A} \otimes \mathcal{B}(H)$, and imposes the condition

$$\gamma^{-1}(g)\, a\, \gamma(g) = U(g)a. \tag{134}$$

The simplest example of this construction would be to take $\mathcal{A}$ finite dimensional. If G is Abelian, we can even take $\mathcal{A} = \mathbb{C}$, and the general solution for a will be some particular solution added to an arbitrary element of $\mathcal{A}_G$. As we discuss below, the particular solution has the interpretation of a connection, and this construction leads directly to gauge theory on $\mathcal{A}_G$.

Suppose we had started with $\mathcal{A} = C(M)$. If we take $\gamma(g)$ to be the trivial representation, Eq. (134) defines the algebra of functions on the quotient space $C(M/G)$. More general choices of $\gamma(g)$ thus lead to a generalized concept of quotient. The striking feature of this is that it provides a way to define quotients by "bad" group actions, those for which the quotient space M/G is pathological, as is discussed in Connes (1994). This definition of quotient also follows from the standard string theory definition of orbifolds, as discussed in Douglas (1999), Konechny and Schwarz (2000a), and Martinec and Moore (2001).

A natural generalization of this construction is the twisted cross product, whose definition is precisely the same except that we take γ to be a projective representation as in Eq. (133). This leads to gauge theory on $\mathcal{A}_{G,\epsilon}$; we shall discuss the toroidal case below.

Another construction with a related geometric picture, the foliation algebra, is discussed in Connes (1994).

D. Gauge theory and topology

A good understanding of the topology of conventional gauge field configurations requires introducing the notions of principal bundle and vector bundle. We recall that conventional gauge fields are connections in some principal G bundle, while matter fields are sections of vector bundles with structure group G. On a compact space such as a torus, the topological classication of these bundles has direct physical implications.

The noncommutative analog of these ideas is a central part of noncommutative geometry. We refer to Konechny and Schwarz (2000b) for a detailed discussion focusing on the example of the noncommutative torus, but we give the basic definitions here. See also Harvey (2001a) for a related discussion.

One aspect of a conventional vector bundle is that one can multiply a section by a function to get another section. This will be taken as the defining feature of a noncommutative vector bundle associated to the algebra $\mathcal{A}$; one considers the (typically infinite dimensional) linear space of sections and requires it to be a module over the algebra $\mathcal{A}$.

A module E over $\mathcal{A}$ is a linear space admitting a multiplicative action of $\mathcal{A}$ which is bilinear and satisfies the rule

$$a\cdot(b\cdot v)=(ab)\cdot v \quad (a,b\in\mathcal{A};\quad v\in E), \tag{135}$$

and carrying whatever other structure $\mathcal{A}$ has (e.g., continuity, smoothness, etc.) The simplest example is just an N-component vector with elements in $\mathcal{A}$, which is called a free module of rank N. Sections of a trivial rank N conventional vector bundle form the free module $C(M)^N$, and this is the obvious generalization.

We chose the multiplication in Eq. (135) to act on the left, defining a left module, but one can equally well let it act on the right, defining a right module, or postulate independent multiplication laws on both sides, defining a bimodule. One can also speak of a ($\mathcal{A}$, $\mathcal{B}$) bimodule, which admits an action of $\mathcal{A}$ on the left and another algebra $\mathcal{B}$ on the right.

A simple example of a bimodule would be the space of $M\times N$ matrices of elements of $\mathcal{A}$, with $\mathcal{A}\otimes\mathbf{1}_M$ acting on the left and $\mathcal{A}\otimes\mathbf{1}_N$ on the right. These sit inside left and right actions of $\mathrm{Mat}_M(\mathcal{A})$ and $\mathrm{Mat}_N(\mathcal{A})$, respectively. An obvious but important point is that these two actions commute, i.e., $(a\cdot v)\cdot b=a\cdot(v\cdot b)$.

We can regard the free module $\mathcal{A}^N$ as a bimodule (the $1\times N$ matrices) and this comment shows that it admits a right action of $\mathrm{Mat}_N(\mathcal{A})$ which commutes with the action Eq. (135). There is a general term for the linear maps acting on a module E which commute with Eq. (135); they are the endomorphisms of the module E, and the space of these is denoted $\mathrm{End}_{\mathcal{A}}E$. In fact $\mathrm{End}_{\mathcal{A}}(\mathcal{A}^N)\cong\mathrm{Mat}_N(\mathcal{A})$ so we know all endomorphisms of the free module. For more general E, $\mathrm{End}_{\mathcal{A}}E$ will always be an algebra, but need not be a matrix algebra.

We can now obtain more examples by starting with free modules and applying a projection. We use the fact that the left module $\mathcal{A}^N$ admits a right action of $\mathrm{Mat}_N(\mathcal{A})$. Given a projection $P\in\mathrm{Mat}_N(\mathcal{A})$, the space of solutions of $v\cdot P=v$ is a module with multiplication law $a\times v=a\cdot v$. These examples are known as finitely generated projective modules and in fact it is only these modules which are natural generalizations of vector bundles, so one normally restricts attention to them. The endomorphisms of these modules also admit a simple description: they are the elements in $\mathrm{Mat}_N(\mathcal{A})$ of the form PaP.

We are now prepared to make a more general definition of connection (Connes, 1980). So far we have been taking a connection to be a set of operators $D_i=\partial_i+A_i$ where the components A_i are taken to be elements of $\mathrm{Mat}_N(\mathcal{A})$, in direct analogy to conventional Yang-Mills theory. We used this in several ways; as Eq. (30) acting on fields in $\mathrm{Mat}_N(\mathcal{A})$ (the *adjoint action*), in the curvatures Eq. (37), and finally acting on *fundamental* matter in Sec. II.D.3. The last of these is the point where vector bundles enter the conventional discussion, and where our generalization will apply most directly.

More generally, a connection on the module E could be any set of linear operators D_i which act on E and satisfy the Leibniz rule,

$$D_i(a\cdot v)=\partial_i(a)\cdot v+a\cdot D_i(v). \tag{136}$$

To see the relation to our previous definition, we first note that the difference between two connections $A_i\equiv D_i-D_i'$ will commute with the action of $\mathcal{A}$, and is thus an endomorphism of E. Thus we can also describe the general connection of this class by choosing a fiducial connection $D_i^{(0)}$ and writing

$$D_i=D_i^{(0)}+A_i,\quad A_i\in\mathrm{End}_{\mathcal{A}}E.$$

Let us consider $E=\mathcal{A}^N$ so we can compare with Sec. II.D.3. One can clearly take $D_i^{(0)}=\partial_i$ for the free module, and as we discussed, the endomorphisms are just $\mathrm{Mat}_N(\mathcal{A})$, so we see that in this case the new and old definitions of connection agree.

All of the modules constructed from projections as above also admit a natural candidate for $D_i^{(0)}$, namely, we apply ∂_i and project back:

$$D_i^{(0)}(v)=(\partial_i v)\cdot P.$$

Thus connections on E can be identified with elements of $\mathrm{End}_{\mathcal{A}}E$, which as we discussed will be some subalgebra of matrices $\mathrm{Mat}_N(\mathcal{A})$, but not itself a matrix algebra.

Finally, the gauge theory action uses one more ingredient, the trace. This can also be defined in terms of the projection P; an endomorphism which can be written PaP as above has trace $\mathrm{Tr}|_{\mathrm{Mat}_N(\mathcal{A})}PaP$. In particular, we can define the dimension of the module as

$$\mathrm{Tr}_E 1=\mathrm{Tr}_{\mathrm{Mat}_N(\mathcal{A})}P.$$

These definitions can be used to obtain all the noncommutative vector bundles, and largely reduces the classification problem to classifying the projections in $\mathrm{Mat}_N(\mathcal{A})$. The next step in such a classification is to find invariants which tell us when two projections are related by a continuous deformation. A natural guess is that one wants to generalize the Chern classes of the conventional theory, and indeed we have been implicitly doing this in claiming that quantities like Eq. (130) are topological quantum numbers. This turns out to be true for the noncommutative torus and indeed there is a well-developed formalism generalizing this to arbitrary algebras, based on cyclic cohomology (Connes, 1982).

E. The noncommutative torus

The noncommutative torus and its associated modules can be obtained using almost any of the constructions we cited, so besides its physical relevance it serves as a good illustration. We follow the definition of $\mathbf{T}_\theta^d$ made in Connes (1980), as the algebra of linear combinations of products $\Pi_i U_i^{n_i}$ of generators with the relations Eq. (23), with coefficients decreasing faster than any power of $|n|$. Our discussion will mostly stick to $\mathbf{T}_\theta^2$.

One can regard commutative $C(\mathbf{T}^2)$ as the group algebra $G_{\mathbb{Z}^2}$, i.e., all linear combinations of products of two commuting generators $U_1=\gamma(g_1)$ and $U_2=\gamma(g_2)$. We can think of γ as the regular representation; it decomposes into the direct sum of all irreducible representations of $\mathbb{Z}^2$, which can be written $U_i=e^{i\sigma_i}$, parametrized by σ_i coordinates on $\mathbf{T}^2$. Each such representation is a one-dimensional module over $G_{\mathbb{Z}^2}$.

Taking instead a projective representation Eq. (133) with $\epsilon(g_1,g_2)=e^{-i\theta/2}$ leads directly to Eq. (23). The regular representation can be written explicitly as

$$U_1=e^{i\sigma_1}S_{\sigma_2,-\theta}, \quad U_2=e^{i\sigma_2}, \tag{137}$$

where $S_{\sigma,a}f(\sigma)\equiv f(\sigma+a)S_{\sigma,a}$. This can again be decomposed into irreducibles, each labeled by a fixed value of σ_1. All of these are equivalent, however (if we take this to act on functions on $\mathbb{R}^2$), under conjugation by $e^{i\alpha\sigma_2}$.

This gives us an example of an projective module E over $\mathbf{T}^2_\theta$: to repeat, we take for E the smooth functions on a real line $\mathcal{S}(\mathbb{R})$, and define the action of the operators U_i on them as

$$(U_1f)(\sigma)=f(\sigma-\theta), \quad (U_2f)(\sigma)=e^{i\sigma}f(\sigma). \tag{138}$$

In terms of the general construction of the previous subsection, we obtained E from $\mathcal{A}^1$ by a projection, which should allow us to compute $\dim E=\mathrm{Tr}_E 1$. Naively $P=\delta(\sigma_1)$; this is too naive as this operator is not bounded. A correct projector can be found using the ansatz

$$P=U_2^\dagger g(U_1)^\dagger+f(U_1)+g(U_1)U_2$$

for functions f and g chosen to satisfy $P^2=P$ (Connes, 1980; Rieffel, 1981; Bars *et al.*, 2001), and this can be used to compute $\dim E=\theta/2\pi$.

Rather than continue by following the general theory, in this case it is easier to write explicit results. For example, the endomorphisms of E are

$$(Z_1f)(\sigma)=f(\sigma+2\pi), \quad (Z_2f)(\sigma)=e^{2\pi i\sigma/\theta}f(\sigma). \tag{139}$$

These operators also satisfy the defining relations of a noncommutative torus, but one with $\theta'=4\pi^2/\theta$.

The fact that two dual tori are involved may seem counterintuitive. However, there is a sense in which the tori $\mathbf{T}^2_\theta$ and $\mathbf{T}^2_{\theta'}$ are equivalent, called Morita equivalence, based on the observation that E is also a module for $\mathbf{T}^2_{\theta'}$ (since we have an action of the Z_i). We return to this below.

A reference gauge connection $D_i^{(0)}$ on E satisfying $D_i^{(0)}(U_j)=2\pi i\,\delta_{i,j}U_j$ can now be defined as

$$D_1^{(0)}f=-\frac{2\pi i\sigma}{\theta}f, \quad D_2^{(0)}f=2\pi\frac{\partial f}{\partial\sigma}. \tag{140}$$

The general connection is a sum $D_i^{(0)}+A_i$ with $A_i\in\mathrm{End}_{\mathbf{T}^2_\theta}E$. In other words, the vector potential naturally lives on a dual noncommutative torus $\mathbf{T}^2_{\theta'}$. Note that $(1/2\pi i)\mathrm{Tr}[D_1,D_2]=1$ is integrally quantized, even though $\dim E$ was not.

This construction can be generalized to produce all modules over $\mathbf{T}^2_\theta$. These modules $E_{p,q}$ are characterized by two integers p and q, and can be produced by tensoring a representation with $\theta_1=p/q$ constructed in Sec. II.E.6, with Eq. (138) with θ_2 chosen to satisfy $\theta=\theta_1+\theta_2$. The computations above then generalize to

$$\dim E_{p,q}=|p-q\,\theta/2\pi|, \tag{141}$$

$$\mathrm{Tr}_{E_{p,q}}F=2\pi iq. \tag{142}$$

The example E of Eq. (138) is $E_{0,1}$, while regarded as a module on $\mathbf{T}^2_{\theta'}$ it turns out to be $E_{1,0}$.

The same construction generalizes to produce all modules on $\mathbf{T}^d_\theta$ (at least for irrational θ) (Connes, 1980; Rieffel, 1988) and thus we can get the topological classification of gauge field configurations on $\mathbf{T}^d_\theta$. This turns out to be the same, in a sense, as for the commutative torus. In both cases, the topological class of a connection is determined by its Chern classes (Connes, 1980), which can be defined as $\int\mathrm{Tr}\,1$, $\int\mathrm{Tr}F_{i_1i_2}\int\mathrm{Tr}F_{i_1i_2}\wedge F_{i_3i_4}$, etc. and which obey additive quantization rules corresponding to the group $\mathbb{Z}^{2^{d-1}}$. This is reasonable as a continuous variation of θ should not change a topological property.

Many properties of these modules are easier to see from other constructions. In particular, one can also regard $\mathbf{T}^2$ as the quotient $\mathbb{R}^2/\mathbb{Z}^2$, and then define this quotient using the crossed product. Let $n^i\in\mathbb{Z}^2$ act on $\mathbb{R}^2$ as $x^j\to x^j+n^i(e_i)^j$, and take $X^i\in\mathbb{R}^2\otimes B(\mathcal{H})$, then Eq. (134) becomes

$$U_i^{-1}X^jU_i=X^j+(e_i)^j. \tag{143}$$

If we take the regular representation as above, and $(e_i)^j=\delta_i^j$ for simplicity, these equations are solved by

$$X^j=-i\frac{\partial}{\partial\sigma^j}+A_j, \tag{144}$$

where A_j commute with the U_i; i.e., are general functions on $\mathbf{T}^2$. It is no coincidence that these solutions look like covariant derivatives; we can rewrite Eq. (143) as

$$X^jU_i=U_iX^j+(e_i)^jU_i \tag{145}$$

which is precisely Eq. (136) with $X^j=D_j$.

To get $\mathbf{T}^2_\theta$, we instead take the twisted crossed product, which amounts to solving Eq. (143) with U_i satisfying Eq. (23). Using the representation Eq. (137) for this, the solutions become

$$X^i=D_i^{(0)}+A_i \tag{146}$$

with $D_i^{(0)}$ as in Eq. (140), and A_i general elements of $\mathbf{T}^2_{\theta'}$.

This generalizes to a procedure for deriving connections on $\mathbf{T}^d_\theta$, which by taking more general U_i produces all constant curvature connections. It corresponds directly to the string and M(atrix) theory definition of quotient space, and thus we shall find in Sec. VII.B that $\mathbf{T}_\theta$ appears naturally in this context.

Although we started with the continuum definition Eq. (137), one could also obtain this as an explicit limit of matrix representations. Using this in the constructions we just discussed leads to the regulated gauge theory discussed in Sec. II.E.6.

F. Morita equivalence

Two algebras $\mathcal{A}$ and $\hat{\mathcal{A}}$ are Morita equivalent if there is a natural one-to-one map carrying all modules and their associated structures from either algebra into the other. On the face of it, this is quite a strong relation, which for commutative algebras $C(M)$ and $C(\hat{M})$ would certainly imply that $M \cong \hat{M}$.

We first note that if such a map exists, we can derive the general map just knowing the counterpart in $\hat{\mathcal{A}}$ of $\mathcal{A}^1$, the free module of rank 1. Call this P; it is an $(\hat{\mathcal{A}},\mathcal{A})$ bimodule, because we know that $\mathcal{A}^1$ also admits a right action of $\mathcal{A}$. There is then a general construction (the mathematical notion of tensor product) which produces the map

$$E \to \hat{E} = P \otimes_{\mathcal{A}} E. \tag{147}$$

An example is that $\hat{\mathcal{A}} = \mathrm{Mat}_N(\mathcal{A})$ is Morita equivalent to $\mathcal{A}$, by taking P to be $\mathcal{A}^N$. Thus we do not want to think of Morita equivalent algebras as literally equivalent; however, their K theory and many other properties are the same.

A more striking example of a Morita equivalence is that $\mathbf{T}^2_\theta$ is Morita equivalent to $\mathbf{T}^2_{\theta'}$ as above. The module E constructed in Eqs. (138) and (139) is the bimodule which provides this equivalence. Thus every module $E_{p,q}$ on $\mathbf{T}^2_\theta$ is associated by Eq. (147) to a module $E_{-q,p}$ on $\mathbf{T}^2_{\theta'}$.

There is a second, simpler equivalence of this type obtained by taking $\theta \to \theta + 2\pi$, which manifestly produces the same algebra. Considering Eq. (141) shows that this acts on the modules as $E_{p,q} \to E_{p+1,q}$.

These two transformations generate the group $SL(2,\mathbb{Z})$ of 2×2 matrices

$$\begin{pmatrix} a & b \\ d & d \end{pmatrix}$$

with integer entries and determinant 1, acting on $\tau \equiv \theta/2\pi$ as $\tau \to (a\tau + b)/(c\tau + d)$ and on (p,q) as a vector. This is an example of a duality group which is a candidate equivalence between gauge theories based on the pair of Morita equivalent modules. This equivalence can be further strengthened by introducing a stronger notion of gauge Morita equivalence (Schwarz, 1998), which produces a map between the spaces of connections on the two modules which preserves the Yang-Mills action.

Although somewhat abstract, in this example these equivalences have a simple geometric origin, which can be understood in terms of the constructions of the previous subsection (Connes, 1994; Douglas, 1999). It is that $\mathbf{T}^2_\theta$ can be obtained as a quotient of $\mathbb{R}$ by the identifications $x \sim x + 2\pi \sim x + \theta$, in other words by a two-dimensional lattice, and as such its moduli space will admit the $SL(2,\mathbb{Z})$ symmetry of redefinitions of the lattice, just as in the construction of $\mathbf{T}^2$ as $\mathbb{R}^2/\mathbb{Z}_2$. Indeed, as this suggests, $\mathbf{T}^2_\theta$ can be regarded is a zero-volume limit of $\mathbf{T}^2$, a picture which will reappear in the string theory discussion.

Morita equivalence between higher-dimensional noncommutative tori is also understood (for irrational θ at least). The basic result is the following (Rieffel and Schwarz, 1998): two tori $\mathbf{T}^d_\theta$ and $\mathbf{T}^d_{\theta'}$ are Morita equivalent if they are related as

$$\theta' = (A\theta + B)(C\theta + D)^{-1},$$

where

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix}$$

is a $2d \times 2d$ matrix belonging to the group $SO(d,d;\mathbb{Z})$. The simplest example is $A = D = 0$, $B = -1$, $C = 1$ which corresponds to the Morita equivalence on $\mathbf{T}^2_\theta$ we described above. The corresponding transformation on the modules is also given by a linear action of $SO(d,d)$ on the Chern class data, which can be regarded as a spinor of this group.

These transformations agree precisely with the action of T duality in toroidal compactifications of string theory (in the limit of a zero volume torus) and this is how they were first conjectured. Conversely, the mathematical proof of these equivalences is part of a new argument for these dualities in M theory.

G. Deformation quantization

Deformation quantization (Bayen *et al.*, 1978; Sternheimer, 1998) is a reformulation of the problem of quantizing a classical mechanical system as follows. One first considers the algebra of observables $\mathcal{A}$ of the classical problem; if one starts with a phase space M (perhaps the cotangent space to some configuration space, but of course it can be more general), this will be the algebra of functions $C(M)$. One then finds a deformation of this algebra in the sense of Sec. II.B.3, i.e., a family of algebras $\mathcal{A}_h$ depending on a parameter h which reduces to $C(M)$ as $h \to 0$, and for which the leading term in the star commutator in a power series expansion in h is the Poisson bracket,

$$f*g - g*f = ih\{f,g\} + \cdots . \tag{148}$$

One can then reinterpret Heisenberg picture equations of motion such as

$$\frac{\partial f}{\partial t} = \frac{i}{h}[H,f]$$

as equations for observables in $\mathcal{A}_h$ involving star commutators.

As discussed in Sec. II.B.3, the Moyal product is precisely such a deformation of the multiplication law of functions, and deformation quantization would appear to be a very direct way to generalize the construction of noncommutative field theory on $\mathbb{R}_\theta$ and $\mathbf{T}_\theta$ to general noncommutative spaces. Not only is it more general, it can be formulated geometrically, without recourse to a specific coordinate system. The primary input, the Poisson bracket, can be specified by the choice of an antisymmetric bivector field θ^{ij},

$$\{f,g\} = \theta^{ij}\partial_i f \partial_j g,$$

such that $\{f,g\}$ satisfies the Jacobi identity.

For the further discussion, it will be useful to recall a bit of the canonical formulation of classical mechanics and its underlying mathematics. A simplifying assumption which often holds is that θ^{ij} is everywhere nondegenerate, in which case one speaks of a manifold with symplectic structure. The symplectic form is $\omega_{ij} = (\theta^{ij})^{-1}$ (the matrix inverse at each point), and the Jacobi identity for the Poisson bracket is equivalent to $d\omega = 0$. More generally, θ^{ij} can degenerate at points or be of less than full rank, in which case one speaks of a manifold with a Poisson structure.

The simplest symplectic structure is constant θ^{ij}, in which case by a linear transformation we can go to canonical coordinates as in Eq. (6). Indeed, one can make a coordinate transformation to canonical coordinates in any contractable region in which θ^{ij} is everywhere nondegenerate (Darboux's theorem). One then distinguishes a special class of coordinate transformations which preserve θ, the canonical transformations, or symplectomorphisms as they are called in mathematics. In infinitesimal form, these are defined by $\delta x = \{S,x\}$ for some generating function S [if $\pi_1(M) \neq 0$, one must allow multivalued generating functions]. One can regard these transformations as generating an infinite dimensional Lie algebra, for which the symplectomorphisms are the corresponding Lie group.

Having defined the deformations of interest, we can now discuss the question of whether they exist. Constructing one involves postulating additional terms in Eq. (148) at all higher orders in h to make the star product associative, $(f*g)*h = f*(g*h)$. One might wonder if this procedure requires more data than just the Poisson bracket, and a little reflection shows that it surely will. After all, we know of a valid star product for constant θ^{ij}, the Moyal product Eq. (21), and if this were the output of a general prescription depending only on θ^{ij}, we would conclude that the Moyal product intertwines with canonical transformations; i.e., if $f \to T(f)$ is a canonical transformation, we would have

$$T(f*g) = T(f)*T(g).$$

However, a little experimentation with nonlinear canonical transformations should convince the reader that this is false. Thus the question arises of what is the additional data required to define a star product, and what is the relation between these different products.

Deformation quantization has been fairly well understood by mathematicians and we briefly summarize the main results, referring to Fedosov (1996) and Kontsevich (1997) for more information. First, deformation quantization always exists. In the symplectic case this was shown by DeWilde and Lecomte (1983) and by Fedosov (1994), who also constructed a trace. For more general Poisson manifolds, it was shown by Kontsevich (1997). As we shall discuss in Sec. VI.C, Kontsevich's construction is in terms of a topological string theory, and has been rather influential in the physical developments already.

In the symplectic case, it is known that the additional data required is a choice of symplectic connection, analogous to the connection of Riemannian geometry but preserving the symplectic form, $\nabla\omega = 0$. Unlike Riemannian geometry, this compatibility condition does not determine the connection uniquely, and acting with a canonical transformation will change the connection. In this language, the special role of the canonical coordinates is analogous to that of normal coordinates; they parametrize geodesics of the symplectic connection, $\nabla x^i = 0$.

Although the precise form of the star product depends on the connection, different choices lead to gauge equivalent star products, in the sense that one can make a transformation

$$G(f*g) = G(f)*' G(g)$$

of the form

$$G(f) = f + h g_1^i \partial_i f + h^2 g_2^{ij} \partial_i \partial_j f + \cdots$$

which relates the two. Thus the two algebras are (formally) equivalent.

We now ask what is the relation between deformation quantization and more conventional physical ideas of quantization, or other mathematical approaches such as geometric quantization. Now the usual intuition is that the dimension of the quantum Hilbert space should be the volume of phase space in units of $(2\pi\hbar)^d$, so a finite-volume phase space should lead to a finite number of states.

This fits with our earlier discussion of fuzzy S^2 but leads us to wonder how we obtained field theory on $\mathbf{T}_\theta$. This appears to be connected to the fact that $\pi_1(T^2) = \mathbb{Z}^2 \neq 0$, and a nice string theory explanation of how this changes the problem can be found in Seiberg and Witten (1999); because a string can wind about π_1, the phase space of an open string on T^2 is in fact noncompact.

In any case, deformation quantization gets around all of these questions in a rather peculiar way: the series expansions in h one obtains are usually formal in the sense that they do not converge, not just when applied to badly behaved functions but for any sufficiently large class of functions. Thus they typically (i.e., for generic values of h) do not define algebras of bounded operators, and do not even admit representations on Hilbert space of the sort which explicitly or implicitly lay behind many of our considerations (Fedosov, 1996; Rieffel, 1998).

Given some understanding of this point, in the general case one also needs to define a noncommutative metric $g_{ij}(x)$ to make sense of Eq. (29); other field theory actions require even more structure.

At this writing, the question of whether and when deformation quantization can be used to define noncommutative field theory is completely open. Some suggestions in this direction have been made in Asakawa and Kishimoto (2000) and Jurco *et al.* (2001).

VII. RELATIONS TO STRING AND M THEORY

We now discuss how noncommutative field theories arise from string theory and M theory, and how they fit into the framework of duality.

Historically, the first use of noncommutative geometry in string theory was in the formulation of open string field theory due to Witten (1986), which uses the Chern-Simons action in a formal setup much like that of Sec. II, with an algebra $\mathcal{A}$ defined using conformal field theory techniques, whose elements are string loop functionals. Noncommutativity is natural in open string theory just because an open string has two ends, and an interaction which involves two strings joining at their end points shares all the formal similarities to matrix multplication which we took advantage of in Sec. II.C.

Although these deceptively simple but deep observations combined with the existence of the string field framework strongly suggested that noncommutative geometry has a deep underlying significance in string theory, it was hard to guess just from this formalism what it might be. Further progress in this direction awaited the discovery of the Dirichlet brane (Dai *et al.*, 1989; Polchinski, 1995), which gave open strings a much more central place in the theory, and allowed making geometric interpretations of much of their physics.

Now contact between string theory and conventional geometry, as epitomized by the emergence of general relativity from string dynamics, relies to a large extent on the curvatures and field strengths in the background being small compared to the string length l_s. Conversely, when these quantities become large in string units, one may (but is not guaranteed to) find some stringy generalization of geometry.

The simplest context in which noncommutative field theory as we described it arises, and by far the best understood, is in a limit in which a large background antisymmetric tensor potential dominates the background metric. In this limit, the world-volume theories of Dirichlet branes become noncommutative (Connes *et al.*, 1998; Douglas and Hull, 1998). This can be seen from many different formal starting points, as elucidated in subsequent work, and it provides very concrete pictures for much of the physics we discussed in Secs. III and IV. It will also lead to new theories: noncommutative string theories, and even more exotic theories such as open membrane (OM) theory.

After reviewing a range of arguments which lead to noncommutative gauge theory, we focus on its origins from the string world-sheet, following Seiberg and Witten (1999), who were the first to precisely state the limits involved. We also describe related arguments in topological string theory, originating in work of Kontsevich (1997). We then give the string theory pictures for the solitonic physics of Sec. III, and other contacts with duality such as the AdS/CFT proposal (Maldacena, 1998). Finally, we discuss some of the limits which have been proposed to lead to new noncommutative theories.

A. Lightning overview of M theory

Obviously space does not permit a real introduction to this subject, but it is possible to summarize the definition of M theory so as to provide a definite starting point for our discussion. Details of the following arguments can be found in Polchinski (1996a, 1996b). Throughout this section, we shall try to state the central ideas at the start of each subsection, though as the discussion progresses we shall reach a point where we must assume prior familiarity on the part of the reader.

A unified way to arrive at the various theories which are now considered part of M theory is to start with one of the various supergravity actions with maximal supersymmetry (32 supercharges in flat space-time), compactify some dimensions consistent with this supersymmetry (the simplest choice is to compactify n dimensions on the torus T^n), find the classical solutions preserving as much supersymmetry as possible (16 supercharges), make arguments using supersymmetry that these are exact solutions of the quantum theory, and then claim that in a given background (say, with specified sizes and shapes of the torus), the lightest object must be fundamental, so some fundamental formulation should exist based on that object.

Thanks to the remarkable uniqueness properties of actions with maximal supersymmetry, in each case only one candidate theory survives even the simplest consistency checks on these ideas, and thus one can be surprisingly specific about these fundamental formulations and see some rather nontrivial properties of the theory even in the absence of detailed dynamical understanding.

The simplest and most symmetric starting point is 11-dimensional supergravity (Cremmer *et al.*, 1978), a theory with no free parameters and a single preferred scale of length, the Planck length l_p. Its fields are the metric, a spin-3/2 gravitino, and a third rank antisymmetric tensor potential, traditionally denoted C_{ijk}. Such a potential can minimally couple to a $2+1$-dimensional extended object, and indeed a solution exists corresponding to the background fields around such an object and preserving 16 supersymmetries. One can also find a supersymmetric solution with magnetic charge emanating from a $5+1$-dimensional hypersurface. We can thus define M theory as the well-defined quantum theory of gravity with the low-energy spectrum of this supergravity, containing solitonic branes, the 2-brane (or supermembrane) and 5-brane, whose long-range fields (at distances large compared to l_p) agree with the solutions just discussed. These branes have tensions (energy per unit volume) c_2/l_p^3 and c_5/l_p^6 as is clear by dimensional analysis; arguments using supersymmetry and charge quantization determine the constants c_n.

This is not a constructive definition and indeed one might doubt that such a theory exists at all, were it not for its connections to superstring theory. The simplest connection is to consider a compactification of the theory on $\mathbb{R}^{9,1}\times S^1$ with a flat metric, circumference $2\pi R$ for the S^1, and no other background fields. One can derive the resulting ten-dimensional supergravity by standard Kaluza-Klein reduction and find IIa supergravity, a theory which can be independently obtained by quantizing the IIa superstring. Indeed, the superstring itself can be identified with a 2-brane with one spatial dimension identified with (or wrapped on) the S^1; in the small R limit this will look like and has precisely the action for a string in ten dimensions. This relation determines the string length l_s, the average extent of a string, and the string tension, $1/l_s^2 = cR/l_p^3$ (again with a known coefficient).

The claim is now that this compactification of M theory exhibits two limits, one with $R \gg l_p$ which can be understood as 11-dimensional supergravity with quantum corrections, and one with $R \ll l_p$ which can be understood as a weakly coupled superstring theory. Considerations involving supersymmetry as well as many nontrivial consistency checks have established this claim beyond reasonable doubt, as part of a much larger web of dualities involving all of the known string theories and many of their compactifications.

The string theory limits are still much better understood than the others, because the string is by far the most tractable fundamental object. One can use them to make a microscopic definition of certain branes, the Dirichlet branes. A Dirichlet brane is simply an allowed end point for open strings. The crucial generalization beyond the original definition of open string theory is that one allows Dirichlet boundary conditions for some of the world-sheet coordinates and this fixes the end point to live on a submanifold in space-time. For a simple choice of submanifold such as a hyperplane, the world-sheet theory is still free, so the physics can be worked out in great detail.

The central result in this direction and the starting point for most further considerations is that the quantization of open strings ending on a set of N Dp-branes, occupying coincident hyperplanes in ten-dimensional Minkowski space-time, leads to $p+1$-dimensional $U(N)$ MSYM. Its field content is a vector field, $9-p$ adjoint scalars and their supersymmetry partner fermions, and its action is the dimensional reduction of Eq. (31).

A crucial point is the interpretation of the adjoint scalars. Let us denote them as Hermitian matrices X^i; the action contains a potential

$$V = -\sum_{i<j} \mathrm{Tr}[X^i, X^j]^2$$

(the sign is there for positivity). A zero energy configuration satisfies $[X^i, X^j] = 0$ and is thus given by a set of $9-p$ diagonal matrices (up to gauge equivalence).

The point now is that the N vectors of eigenvalues X^i_{nn} must be identified as the positions of the N branes in the $9-p$ transverse dimensions. This identification is behind most of the geometric pictures arising from D-brane physics, and the promotion of space-time coordinates to matrices is at the heart of the noncommutativity of open string theory.

As the simplest illustration of this, a generic configuration of adjoint scalars will break $U(N)$ gauge symmetry to $U(1)^N$ by the Higgs effect, giving masses $|X_{mm} - X_{nn}|$ to the (m,n) off-diagonal matrix elements of the fields. This corresponds to the mass of a string stretched between two branes at the positions X_{mm} and X_{nn} and we see that the Higgs effect has a simple picture in terms of the geometry of an extra dimension.

This picture shows some resemblance to pictures from noncommutative gauge theory appearing in Connes (1994), and this observation (Douglas, 1996; Ho and Wu, 1997) led to a search for a more direct connection between noncommutative gauge theory and D-brane physics.

B. Noncommutativity in M(atrix) theory

As we commented in Sec. II.D.1, the simplest derivation of noncommutative gauge theory from a more familiar physical theory is to start with a dimensionally reduced gauge theory (or matrix model) action such as Eq. (38), and find a situation in which the connection operators C_i of Eq. (34) obey the defining relations of a connection on a module over a noncommutative algebra as given in Sec. VI.D, perhaps after specifying appropriate boundary conditions or background fields.

A particularly significant theory of this type is maximally supersymmetric quantum mechanics, the $p=0$ case of MSYM, with action

$$S = \int dt\, \mathrm{Tr} \sum_{i=1}^{9} (D_t X)^2 - \sum_{i<j} [X^i, X^j]^2 + \chi^\dagger (D_t + \Gamma_i X^i) \chi. \tag{149}$$

Here $D_t = \partial/\partial t + iA_0$, and varying A_0 leads to the constraint that physical states be invariant under the action of $U(\mathcal{H})$.

This action first entered M theory as a regulated form of the action for the supermembrane, which as we discussed one might try to use as a fundamental definition of the theory (de Wit *et al.*, 1988). How this might work was not properly understood until the work of Banks *et al.* (1997), who argued that a simpler and equally valid way to obtain Eq. (149) from string theory was to take the theory of $D0$-branes in IIa string theory and boost it along the x^{11} dimension to the infinite momentum frame. Bound states of these $D0$-branes would be interpreted as the supergravity spectrum, while the original membrane configuration could also be obtained as a nontrivial background. An important feature of this interpretation is that the compact eleventh dimension of our previous subsection does not disappear in taking this limit; one should think of the resulting theory as M theory compactified on a lightlike circle. See Taylor (2001) for a recent review.

In this framework, compactification on the torus T^n is quite simple to understand, in more than one way. One can first compactify the IIa string and take a similar limit to obtain $n+1$-dimensional MSYM, with a similar interpretation. The case $n=1$ reproduces the original string theory in a slightly subtle but convincing way (Dijkgraaf *et al.*, 1997; Motl, 1997) and this is one of the main pieces of confirming evidence for the proposal.

Another approach, spelled out by Taylor (1997), is to define toroidal compactification using the general theory of D-branes on quotient spaces discussed in Sec. VI.C. Letting $U_i = \gamma(g_i)$ for a set of generators of $\mathbb{Z}^n$, and taking $\mathcal{A} = \mathrm{Mat}_n(\mathbb{C})$, this leads to Eqs. (143),

$$U_i^{-1} X^j U_i = X^j + \delta_i^j 2\pi R_i .$$

These are solved by the connection Eq. (144), and substituting into Eq. (149) leads to MSYM on $T^n \times \mathbb{R}$.

This construction admits a natural generalization, namely, one can impose the relations

$$U_i U_j = e^{i\theta_{ij}} U_j U_i .$$

Again as discussed in Sec. VI.C, Eq. (143) now defines a twisted crossed product, and its solutions (146) are connections on the noncommutative torus. Substituting these into Eq. (149) leads directly to noncommutative gauge theory. This was how noncommutativity was first introduced in M theory by Connes *et al.* (1998).

Having seen this possibility, one next must find the physical interpretation of this noncommutativity. Since M theory has no dimensionless parameters, one is not allowed to make arbitrary modifications to its definition but rather must identify all choices made in a particular construction as the values of background fields. Although θ had not been previously noticed as such a choice, it appears naturally given the interpretation of M(atrix) theory as M theory on a lightlike circle, as a background constant value for the components C_{ij-} of the three-form potential, where $-$ denotes the compact lightlike direction. This interpretation was supported by comparing the expected duality properties of the noncommutative gauge theory and of M theory in this background, a subject we return to below.

Having seen how M(atrix) theory can lead to noncommutativity and then to a string, one wants to close the circle and show that string theory can lead to noncommutativity on brane world volumes, from which noncommutative M(atrix) theory can be derived. The *IIa* string interpretation of C_{ij-} is as the "Neveu-Schwarz B field," a field which minimally couples to the string world sheet as in the action

$$S = \frac{1}{4\pi l_s^2} \int_\Sigma (g_{ij} \partial_a x^i \partial^a x^j - 2\pi i l_s^2 B_{ij} \epsilon^{ab} \partial_a x^i \partial_b x^j). \tag{150}$$

We consider space-time $\mathbb{R}^{9-n,1} \times T^n$ where the torus has metric $g_{ij} = R^2 \delta_{ij}$ and constant Neveu-Schwarz B field B_{ij}. In this case, the term in Eq. (150) involving B is an integral of a total derivative, and will be nonzero either because of the nontrivial topology of the torus or in the presence of a world-sheet boundary.

Contact with M(atrix) theory suggests that we study the physics of $D0$-branes in this theory. One way to proceed (Douglas and Hull, 1998) is to apply a T duality along one axis (say x^1) of the torus, which one can show turns the $D0$-branes into $D1$-branes extending along the x^1 axis, and the T^2 into another T^2 with $B=0$ and metric defined by the identifications $(0,0) \sim (l_s, 0) \sim (\theta = B, \epsilon = V/l_s)$. See Fig. 6.

This gets rid of the B field and thus one must be able to understand the physics in conventional geometric terms. The point is that while general arguments lead to a $1+1$-dimensional gauge theory on the world volume of this brane, in the limit of small ϵ, open strings which

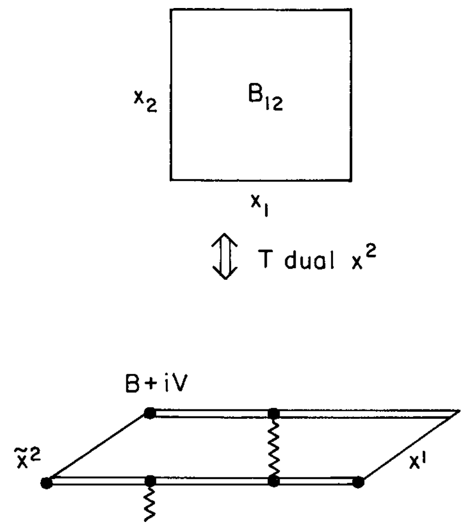

FIG. 6. T duality to an anisotropic torus.

wind about the x^2 dimension will also become light and must be included in the action; their winding number w^2 becomes a new component of momentum in $2+1$-dimensional gauge theory. In this anisotropic geometry, the two ends of a winding open string will have different locations in x^1, with separation θw^2. Thus the fundamental objects turn out to be dipoles in exactly the sense described in Sec. II.B.2, with the corresponding noncommutative interactions. This construction directly reproduces the quotient construction of $\mathbf{T}_\theta$ as $C(S^1)/\mathbb{Z}$.

A world-sheet argument treating $D0$-branes on the original torus leads to the same conclusion (Cheung and Krogh, 1998). Now one must take the size R of both axes of the torus small and keep winding strings in both directions. The point now is that since the term $\int B$ in Eq. (150) is a total derivative, it contributes to the interaction of two strings with winding numbers (w_1, w_2) and (w_1', w_2') about the two axes of the torus by a phase proportional to the product $w_1 w_2' - w_2 w_1'$, directly producing Eq. (21). This argument can be generalized to other string theory situations involving similar phases, such as discrete torsion on orbifolds (Vafa, 1986), and leads to the twisted crossed product with finite groups (Douglas, 1998; Ho and Wu, 1998).

These derivations arose naturally in the consideration of M(atrix) theory, and in this context there is a rather striking test one can make. The original derivation led to an identification of M theory compactified on T^d as a large N limit of $d+1$-dimensional MSYM, and the most basic prediction of this identification is that the two theories share the same duality properties. Connes *et al.* (1998) discussed compactification on T^3, which according to M theory considerations must be invariant under an action of the U-duality group $SL(3,\mathbb{Z}) \times SL(2,\mathbb{Z})$ on the moduli and brane spectrum. The $SL(3,\mathbb{Z})$ can easily be identified with large diffeomorphisms of the T^3, but the $SL(2,\mathbb{Z})$ symmetry is a prediction: in fact it is just the $SL(2,\mathbb{Z})$ duality of $N=4$ SYM proposed by Montonen and Olive (1977).

The generalized compactifications with $C_{-ij} \neq 0$ allow accessing more dualities, namely, those which can be seen in the limit of compactification on a T^{d+1} containing a lightlike dimension. For example, the noncommutative 2+1 MSYM is predicted to have $SL(2,\mathbb{Z})\times SL(2,\mathbb{Z})$ duality symmetry, the subgroup of $SL(3,\mathbb{Z})\times SL(2,\mathbb{Z})$ preserving the choice of lightlike direction. Now the second, nonclassical duality is manifest: it is the $SL(2,\mathbb{Z})$ Morita duality group of $\mathbf{T}^2_\theta$, and one can verify that the spectrum predicted in Eq. (117) is invariant under this change of parameters.

This type of argument has been extended in many directions, and an in-depth treatment would require another full length review, which happily already exists (Konechny and Schwarz, 2000b). In the M(atrix) theory context, it has been argued (Verlinde, 2001) that by careful treatment of the limit, one can extend the duality to a general T^{d+1}, not necessarily preserving the lightlike dimension. In the string theory context, as we discuss below, one can see that the nonclassical duality arises from T duality, leading to the prediction that noncommutative gauge theory on T^d will have $SO(d,d;\mathbb{Z})$ duality, which was the motivation behind the theorem of Rieffel and Schwarz (1998) discussed in Sec. VI.E.

C. Noncommutativity in string theory

The arguments we just gave established that limits of M theory and string theory compactified on a torus will lead to noncommutative gauge theory, realizing the intrinsic noncommutativity of open strings. However, the torus is not the simplest noncommutative space; one is led to ask whether noncommutativity can also emerge without compactification, leading to gauge theory on $\mathbb{R}^d_\theta$. This question gained particular focus after the discovery of noncommutative instantons on $\mathbb{R}^d_\theta$ (Nekrasov and Schwarz, 1998).

While the arguments of the previous section lead directly to this result (to obtain $\mathbb{R}^d_\theta$, one just adjusts the parameters to make the string modes light, which takes the volume of $\mathbf{T}^d_\theta$ to infinity), many other string theory computations on $\mathbb{R}^d$ with general background B had been done previously and noncommutativity had not been seen. Indeed, there are general arguments which lead from open strings to conventional gauge theory, making the new claim appear paradoxical.

On the other hand, Kontsevich (1997) had argued that deformation quantization and the Moyal product in particular could come from open string theory, at least in a mathematical sense. This was turned into an argument in topological string theory by Cattaneo and Felder (2000) and in physical string theory by Schomerus (1999), proving that $\mathbb{R}^d_\theta$ could indeed emerge directly from world-sheet considerations.

The paradox was resolved by (Seiberg and Witten, 1999) who explained how both conventional and noncommutative descriptions could be correct, along the general lines we already indicated in Sec. II.D.4. Their careful treatment of the limit leading to noncommutative field theory has spurred numerous further developments.

The first point to make is that on $\mathbb{R}^d$ and unlike T^d, a constant background potential B_{ij} is pure gauge and thus cannot affect closed string physics. However, this is not true in the presence of a Dirichlet brane extending along the directions i,j, because the gauge transformation also acts on its world-volume gauge potential A as

$$\delta B^{(2)}=d\lambda(x), \quad \delta A=\lambda(x). \tag{151}$$

This can be seen by rewriting the total derivative term in Eq. (150) as a boundary term, as is appropriate in the presence of open strings, and adding the gauge coupling:

$$S_b=i\int_{\partial\Sigma}[A_j(x)-B_{ij}x^i]\partial_t x^j, \tag{152}$$

where ∂_t and ∂_n denote the tangential and normal derivatives along the boundary $\partial\Sigma$ of the world sheet Σ. The transformation Eq. (151) can be undone by an integration by parts.

Equation (151) implies that the open string effective action can only depend on the combination $F+B$, not F or B separately. In particular, one can gauge B to zero, replacing it by a background magnetic field. However, because of stringy effects, in the limit $l_s^2 B\gg 1$, this could lead to physics quite different than that of a magnetic field in the usual Yang-Mills action.

To proceed, we shall need to make use of the standard relation between world-sheet correlation functions of vertex operators, the S matrix for string scattering, and effective actions which can reproduce this physics (Green *et al.*, 1987; Polchinski, 1998). The basic relation is that each local world-sheet operator $V_n(z)$ corresponds to a space-time field Φ_n. Operators in the bulk of the world-sheet correspond to closed strings, while operators on the boundary correspond to open strings and thus fields which propagate on the world volume of a D-brane. A term in the effective Lagrangian

$$\int d^{p+1}x\sqrt{\det G}\,\mathrm{Tr}\Phi_1\Phi_2\cdots\Phi_n \tag{153}$$

is obtained as a correlation function

$$\left\langle \int dz_1 V_1(z_1)\int dz_2 V_2(z_2)\cdots\int dz_n V_n(z_n)\right\rangle \tag{154}$$

on a world sheet Σ with the disk topology, with operators V_i at successive points z_i on the boundary $\partial\Sigma$, integrated over all z_i satisfying the same ordering as in Eq. (153).

Taking only vertex operators for the massless fields, one finds that the leading $l_s\to 0$ limit of the S matrix is reproduced by the MSYM effective action. It turns out that these considerations also lead to a simple universal effective action which describes the physics of a D-brane with arbitrarily large but slowly varying field strength, the Nambu-Born-Infeld action

$$S_{NBI}=\frac{1}{g_s l_s(2\pi l_s)^p}\int d^{p+1}x\sqrt{\det[g+2\pi l_s^2(B+F)]}. \tag{155}$$

Here g_s is the string coupling, and g is the induced metric on the brane world volume. See Schwarz (2001) for its supersymmetrization, and other references on this topic.

The original string computations leading to the Born-Infeld action were done by Fradkin and Tseytlin (1985) and Abouelsaood *et al.* (1987) in the context of type-I open string theory, i.e., the case $p=9$. For $p<9$, the induced metric g contains the information about the embedding of the brane, and substituting $F=0$, Eq. (155) reduces to the Nambu action governing its dynamics. Finally, the prefactor is the D-brane tension as computed from string theory. Besides detailed computation, there are simple physical arguments for the Born-Infeld form (Bachas, 1996; Polchinski, 1998).

This action summarizes essentially all weakly coupled and weakly stringy physics of a single D-brane and is even valid in the large $l_s^2(B+F)$ limit, so one might at first hypothesize that it is valid in the limit of large B we just discussed, without need of noncommutativity. However, it is not in general valid for rapidly varying field strengths $l_s \partial F \sim 1$, nor is its non-Abelian generalization understood. Thus we can reconcile our earlier arguments for noncommutativity with the Nambu-Born-Infeld action if $\theta < l_s^2$, as all of the new physics we observed would be associated with length scales at which the Nambu-Born-Infeld action broke down.

1. Deformation quantization from the world sheet

The key point in arguing that string theory will lead to noncommutative field theory is to see that a correlation function Eq. (154) will obtain the phase factors Eq. (100). Since this is a product of terms for each successive pair of fields, it should also be visible as a phase Eq. (21) in the operator product expansion of two generic boundary operators carrying momenta k and k', say

$$V_k(z_1)V_{k'}(z_2) \rightarrow (z_1-z_2)^{\Delta_{k+k'}-\Delta_k-\Delta_{k'}} e^{-(i/2)\,\theta^{ij}k_i k_j'} V_{k+k'} + \cdots \tag{156}$$

with

$$V_k(z) = :e^{ik\cdot x(z)}:$$

or any operator obtained by multiplying this by conformal fields ∂x, fermions in the superstring, etc.

Since the action (150) is quadratic, the world-sheet physics is entirely determined by the propagator $\langle x^i(z)x^j(w)\rangle$. The boundary conditions which follow from varying the action Eq. (150) are

$$g_{ij}\partial_n x^j + 2\pi i B_{ij} l_s^2 \partial_t x^j|_{\partial\Sigma} = 0. \tag{157}$$

Now, taking Σ to be the upper half plane with the coordinate $z=t+iy$, $y>0$ we find the boundary propagator to be

$$\langle x^i(t)x^j(s)\rangle = -\alpha' G^{ij}\ln(t-s)^2 + \frac{i}{2}\theta^{ij}\epsilon(t-s), \tag{158}$$

where $\epsilon(t) = -1,0,+1$ for $t<0, t=0, t>0$, respectively, and

$$G^{ij} = \left(\frac{1}{g+2\pi l_s^2 B}\right)^{ij}_S, \tag{159}$$

$$\theta^{ij} = 2\pi l_s^2 \left(\frac{1}{g+2\pi l_s^2 B}\right)^{ij}_A, \tag{160}$$

where S and A denote the symmetric and antisymmetric parts, respectively. From this expression we deduce

$$[x^i, x^j] := T[x^i(t-0)x^j(t) - x^i(t+0)x^j(t)] = i\theta^{ij}. \tag{161}$$

Thus the end points of the open strings live on a noncommutative space with

$$[x^i, x^j] = i\theta^{ij}$$

with θ^{ij} a constant antisymmetric matrix. Similarly, Eq. (156) becomes

$$V_p(t)V_q(s) = (t-s)^{2l_s^2 G^{ij} p_i q_j} e^{-(i/2)\,\theta^{ij} p_i q_j} V_{p+q}(s). \tag{162}$$

Indeed, in this free world-sheet theory, it is no harder to compute the analogous phase factor for an n-point function: it is precisely Eq. (100).

In the formal limit $g_{ij} \to 0$, one finds from Eqs. (159) and (160) that $G^{ij}=0$ and $\theta^{ij} = 2\pi l_s^2 (B_{ij})^{-1}$. Thus the dependence on the world-sheet coordinates s and t drops out (the vertex operators have dimension zero), and the operator product expansion (OPE) reduces to a conventional multiplication law. For Eq. (162), this is the Moyal product Eq. (21), and by linearity this extends to the product of two general functions.

We have again found that background B_{ij} leads to noncommutativity. A precise connection to the previous discussion of toroidal compactification can be made by applying T duality to turn the $D2$-branes of the present discussion into $D0$-branes, and then taking the zero-volume limit of the torus. This T duality leads to the relation $\theta^{ij} = 2\pi l_s^2 (B_{ij})^{-1}$. However, the present argument also works for $\mathbb{R}^d_\theta$.

A rather similar argument shows that world volumes of D-branes in the Wess-Zumino-Witten model are described by field theory on fuzzy spheres as described in Sec. VI.B (Alekseev *et al.*, 1999). An interesting difference is that here $dB \neq 0$ and the corresponding algebra is not associative, except in suitable limits.

2. Deformation quantization from topological open string theory

The idea that by considering only vertex operators of dimension zero, a vertex operator algebra such as Eq. (162) will reduce to an associative algebra, is rather general and is best thought of in the framework of topological string theory (Witten, 1988; Dijkgraaf, 1998). For present purposes, these are string theories for which correlation functions depend not on the locations of operators but only on their topological arrangement on the world sheet, which is exactly the property we used of the $g_{ij}\to 0$ limit of Eq. (162).

One can use topological string theory to construct a deformation quantization corresponding to a general Poisson structure, generalizing the discussion we just gave. We summarize this, following the work of Baulieu *et al.* (2001). See also Cattaneo and Felder (2001).

We start with the action (150), not assuming that g_{ij}, B_{ij} are constant, and rewrite it in the first-order form:

$$S=\frac{1}{4\pi l_s^2}\int_\Sigma (g_{ij}\partial_a x^i \partial^a x^j - 2\pi i l_s^2 B_{ij}\epsilon^{ab}\partial_a x^i \partial_b x^j)$$

$$=\int p_i \wedge dx^i - \pi l_s^2 G^{ij} p_i \wedge \star p_j - \frac{1}{2}\theta^{ij} p_i \wedge p_j, \qquad (163)$$

where

$$2\pi l_s^2 G+\theta = 2\pi l_s^2 (g+2\pi i l_s^2 B)^{-1}.$$

Now take the $l_s^2\to 0$ limit keeping θ and G fixed. The remaining part $\int p\wedge dx + \frac{1}{2}\theta p\wedge p$ of the action (163) exhibits an enhanced gauge symmetry (Cattaneo and Felder, 2000),

$$p_i \mapsto p_i - d\lambda_i - \partial_i \theta^{jk} p_j \lambda_k, \quad x^i \mapsto x^i + \theta^{ij}\lambda_j. \qquad (164)$$

To quantize, this symmetry must be gauge fixed, which can be done by standard Batalin-Vilkovisky (BV) procedures, leading to a topological string theory with some similarities both with the type-A and type-B sigma models. Its field content is conveniently described by promoting p_i and x^i to twisted superfields, with an expansion in world-sheet differential forms $d\sigma^i$ with components of all degrees 0,1,2. The original fields are the 0-form part of x^i and the 1-form part of p_i; the other components are the additional ghosts and auxiliary fields of the BV framework.

A basic observable in this theory is the three-point function on the disk,

$$\langle f(x(0)) g(x(1)) [h(x)\chi\cdots\chi](\infty)\rangle_{\theta,\Sigma=\text{disk}} = \int_X f\star g\, h \qquad (165)$$

for $f,g\in C(X)$, $h\in\Omega^{\dim X}(X)$. As the notation indicates, this will be identified with the star product of a deformation quantization associated to θ. This can be seen by developing the perturbation series in powers of θ; each term in the expansion can be expressed as an explicit sum over Feynman diagrams, producing Kontsevich's construction of deformation quantization.

An important advantage of this approach is that many properties of the formalism have simple arguments from string theory. In particular, the associativity of the $\star$ product defined by Eq. (165) follows in a sense from associativity of the OPE. This is best expressed in terms of a more general set of Ward identities obeyed by the string amplitudes in the theory, which allow making contact with and generalizing the discussion in Sec. VI.G.

As an example of a new result derived from Eq. (163), we give an expression for the θ-deformed action of infinitesimal diffeomorphisms $\delta x^i = v^i(x)$ on functions and covariant derivatives (Baulieu *et al.*, 2001), which can be used to make Eq. (24) generally covariant. One defines these in terms of the conventional action on functions $\delta f = v^i \partial_i f$, with θ transforming as a conventional tensor, $\delta\theta^{ij} = v^k \partial_k \theta^{ij} - \theta^{ki}\partial_k v^j + \theta^{kj}\partial_k v^i$. In the topological field theory (TFT), this becomes

$$\delta f = U_v f(x) = \left\langle f(x(0)) \oint [p_i v^i(x)]^{(1)} [\delta(x(\infty)-x)]\right\rangle_{\theta,\Sigma},$$

where the integral is taken over an arc surrounding the point 0 and ending on the boundary of the disk Σ. This leads to

$$U_v f(x) = v^i \partial_i f(x) + \theta^{kl}\partial_k v^i \partial_{li}^2 f(x) + \cdots .$$

3. The decoupling limit

Many of the deeper aspects of the connection to string theory require a more careful treatment of the decoupling limit leading to noncommutative field theory, as was first made by Seiberg and Witten (1999).

String theories have many more perturbative degrees of freedom than field theories, so to get field theories one looks for controlled limits in which almost all of these degrees of freedom go away, usually by sending their masses to infinity. In the context of D-branes, the generic such limit takes $l_s\to 0$ and thus the string tension $T=1/l_s^2$ to infinity, and simultaneously takes the transverse distance L between branes to zero holding LT fixed. This is the energy scale of the lightest open strings stretched between branes, so in this limit we keep these degrees of freedom while sending excited string energies to infinity as $1/l_s$. Finally, one rescales the coupling constant to keep it fixed in the limit. This leads to a field theory of the lightest open strings, which for flat D-branes will be MSYM. Similar limits in other brane theories can lead to more exotic results, as we shall mention below.

However, there are other massless states in the string theory as well, the closed strings which lead to the gravitational sector, and we need to argue that these decouple. A naive but often correct argument for this is that their couplings are gravitational and are suppressed by a factor $G_N E$ which will also go to zero in this limit. This requires detailed consideration in examples, however. It will turn out to be true in the case at hand at least up to $3+1$ dimensions; we will discuss a potential subtlety below.

The key new point in the present context is that the masses of open strings in this limit are determined by the metric G^{ij} defined in Eq. (159), which has nontrivial B dependence. This follows from the standard string theory relation between the mass of a state and the world-sheet dimension of the corresponding vertex operator, which [as is visible in Eq. (162)] is controlled by G^{ij}. Thus the decoupling limit, sometimes referred to as the Seiberg-Witten limit in this context, takes $l_s\to 0$ while holding G and θ fixed.

The nontrivial relation between the original metric g_{ij} and G^{ij} will show up at many points in the subsequent discussion. One refers to g_{ij} as the closed string metric and G^{ij} as the open string metric as they each govern the kinetic terms and the energies of the lightest states in their respective sectors.

Once this realization is made, the subsequent steps in the derivation of the D-brane world-volume action go through without major changes from the conventional case, leading to noncommutative MSYM and even the noncommutative Nambu-Born-Infeld action, defined by taking Eq. (155), replacing products with star products, and setting the metric to G^{ij} and $B=0$.

The remaining step is to determine the prefactor and thus the gauge coupling. This follows directly if we accept that the Seiberg-Witten map of Sec. II.D.4 between conventional and noncommutative gauge theories maps the conventional action Eq. (155) into the noncommutative action Eq. (155), as we can just specialize to $F=\hat{F}=0$. This leads to the relation

$$\frac{\sqrt{\det(G)}}{G_s l_s(2\pi l_s)^p}=\frac{\sqrt{\det(g+2\pi l_s^2 B)}}{g_s l_s(2\pi l_s)^p} \tag{166}$$

which determines an open string coupling constant G_s and the corresponding noncommutative gauge coupling.

The decoupling limit now will take $l_s\to 0$ in Eqs. (159), (160), and (166), scaling the original string coupling as $g_s \sim l_s^{(3-p+r)/4}$ to end up with noncommutative Yang-Mills theory with finite parameters,

$$\theta=B^{-1},\quad G_{ij}=(2\pi l_s^2)^2 B_{ik}B_{jl}g^{kl}. \tag{167}$$

4. Gauge invariance and the Seiberg-Witten map

An important point which can be seen by carrying out this discussion more explicitly is the precise point at which conventional gauge invariance is replaced by noncommutative gauge invariance. As we mentioned in Sec. II.D.4, there is a very general world-sheet argument for conventional gauge invariance, and indeed this argument is not incorrect; rather, one obtains noncommutative gauge invariance by choosing different conventions, and there is a formal equivalence between conventional and noncommutative gauge theories (Seiberg and Witten, 1999; Andreev and Dorn, 2000b; Seiberg, 2000).

The origins of gauge invariance can be seen in open bosonic string theory. In this theory, the vertex operator for a gauge boson is

$$e_i(p){:}\partial_t x^i e^{ip\cdot x}{:}\leftrightarrow A=A_i(x)dx^i, \tag{168}$$

as was already implicit in Eq. (152).

The Abelian gauge invariance,

$$\delta A_j=\partial_j\varepsilon, \tag{169}$$

then follows by varying A_j in Eq. (152), by taking the integral of the total derivative $\int dz\,\partial_t(\varepsilon)$ as zero. Extra terms can appear in Eq. (169) if there are divergences when the operator under consideration (say V_i) coincides with its neighbors V_{i-1} and V_{i+1}. In Yang-Mills theory, this leads to contact terms $V_{i-1}\varepsilon-\varepsilon V_{i+1}$ which become the nonlinear terms in the gauge transformation law.

The previous discussion makes it very plausible that such an argument will carry through in the noncommutative case with the star product appearing as in Eq. (162). However, making this argument precise requires choosing a cutoff on the world sheet, and different choices at this step can lead to different results. If one point splits and uses the propagator (158), one obtains the star product, but one can also find prescriptions in which the term proportional to θ^{ij} goes away in the coincidence limit, leading to conventional gauge invariance instead. In particular, this will be the result if one defines a propagator and cutoff at $B=0$, treating the B term in Eq. (150) using world-sheet perturbation theory.

Physics cannot depend on this choice and in general a change of renormalization prescription on the string world sheet corresponds to a field redefinition in spacetime. By the preceding arguments, this field redefinition must be the Seiberg-Witten map of Sec. II.D.4. The two descriptions are therefore equivalent, at least in some formal sense, but they are each adapted to different regimes, with conventional gauge theory simpler for small B and noncommutative gauge theory simpler for large B.

Once we understand that the resulting gauge invariance depends on the choice of world-sheet regularization, we can consider choices that lead to different star products. A simple choice to consider is one which leads to the same star product Eq. (21), but defined using a parameter θ which does not satisfy Eq. (160), which would be obtained by treating part of the B term in Eq. (150) as a perturbation. Denote this part as Φ; it will enter Eq. (155) as did B, so we would obtain a description in terms of a noncommutative action depending on $\hat{F}+\Phi$.

This description can be obtained from string theory by a simple generalization of the discussion above. The result is again Eqs. (159) and (160) but now with an antisymmetric term in the metric on both sides [analogous to the B term in Eq. (150)]. Combining this with Eq. (166) to determine the gauge coupling, we obtain

$$\left(\frac{1}{G+2\pi l_s^2\Phi}\right)_S=\left(\frac{1}{g+2\pi l_s^2 B}\right)_S, \tag{170}$$

$$\left(\frac{1}{G+2\pi l_s^2\Phi}\right)_A+\frac{1}{2\pi l_s^2}\theta=\left(\frac{1}{g+2\pi l_s^2 B}\right)_A, \tag{171}$$

$$G_s=g_s\left(\frac{\det(G+2\pi l_s^2\Phi)}{\det(g+2\pi l_s^2 B)}\right)^{1/2} \tag{172}$$

which determines the parameters in the noncommutative form of Eq. (155),

$$S_{NCNBI}=\frac{1}{G_s l_s(2\pi l_s)^p}\int \mathrm{Tr}\sqrt{\det[G+2\pi l_s^2(\hat{F}+\Phi)]}. \tag{173}$$

Using Eq. (42), one can show that for slowly varying background fields $l_s^2 F \ll 1$, the descriptions with different values of θ are equivalent.

In the special case of $\Phi=-\theta^{-1}$, these relations simplify even further to Eq. (167), even at finite l_s. As mentioned in Sec. II.B.1, this is what one gets if one defines the derivatives as inner derivations, so this is what emerges naturally from matrix model arguments (Seiberg, 2000).

More generally, it has been shown that the techniques of Kontsevich's deformation quantization can be used to write a Seiberg-Witten map for arbitrary (nonconstant) θ (Jurco *et al.*, 2000).

D. Stringy explanations of the solitonic solutions

So far we have only discussed the simplest arrangement of Dirichlet branes, N parallel branes of the same dimension. There are a bewildering variety of more complicated possibilities, with branes of lower dimension sitting inside of those of higher dimension, intersecting branes, curved branes, and so forth.

Not all of the possibilities are actually distinct, however. For example, a configuration of lower-dimensional D-branes sitting inside a higher-dimensional D-brane, is topologically equivalent to and can often be realized as a limit of a smooth configuration of gauge fields on the higher-dimensional brane. Another example is that an intersecting brane configuration, say of branes A and B, can often be described as a single object, a nontrivial embedding of brane A in the higher-dimensional space and thus by some configuration of scalar fields on brane A. Thus the solitons and instantons in D-brane world-volume gauge theory themselves have interpretations as D-branes. We refer to Polchinski (1996b, 1998) for an introduction and overview of this subject.

A great deal of this structure survives the limit taking string theory to noncommutative field theory, leading to stringy pictures for the gauge theory solutions of Sec. III. What is quite striking is that the agreement is not just qualitative but even quantitative, with brane tensions and other properties agreeing, apparently for reasons other than supersymmetry (Dasgupta *et al.*, 2000; Gross and Nekrasov, 2000b; Harvey *et al.*, 2000). Establishing this detailed agreement of course requires a good deal of string theory input and we refer the reader to Nekrasov (2000) and Harvey (2001a) for detailed reviews of this class of results, but provide an introduction here.

The simplest examples are the $D(p-k)$-branes embedded in Dp-branes. There is a general result to the effect that a collection of N Dp-branes carrying gauge fields with Chern character,

$$\mathrm{ch}(F)=\mathrm{Tr}\, e^F = N+\mathrm{ch}^{(1)}_{i_1 i_2}+\mathrm{ch}^{(2)}_{i_1 i_2 i_3 i_4}+\cdots,$$

carries the same charges as (and is topologically equivalent to) a collection with no gauge fields, but with additional D-branes whose number and orientation is given by the quantized values of the Chern characters. For example, $\mathrm{ch}^{(1)}_{ij}$ counts $D(p-2)$-branes which are localized in the i and j dimension, and so forth.

Carrying this over to the noncommutative theory in the obvious way, the fluxon of Sec. III.B is an embedded $D(p-2)$-brane, and the instanton of Sec. III.C is an embedded $D(p-4)$-brane. According to this identification and the results of Sec. III, the $D(p-2)$ should be unstable to decay, while the $D(p-4)$ should be stable, and this is borne out by computation in string theory; the destabilizing mode in Eq. (87) is a tachyonic open string.

The agreement is much more detailed and striking than this. Another qualitative point is that m coincident $D(p-2)$-branes must themselves carry $U(m)$ gauge symmetry; this is manifestly true of the m-fluxon solution Eq. (83) with $U=S^m$. A similar discussion can be made for higher codimension branes and even for processes such as the annihilation of a D-brane with an anti-D-brane, which are quite difficult to study in the original string theory. A very general relation between the topology of gauge field configurations in noncommutative gauge theory, and the classification of D-branes in string theory, has been found in Harvey and Moore (2000), based on the K theory of operator algebras.

Even better, paying careful attention to conventions, one finds that Eq. (84) exactly reproduces the tension of the $D(p-2)$-brane in string theory. It is somewhat surprising that the noncommutative field theory limit would preserve any quantitative properties of the solutions. The agreement of the tension is related to a deep conjecture of Sen (1998), as explained by Harvey *et al.* (2000). Witten (2000) has shown how the noncommutative field theory and these arguments can be embedded in the framework of string field theory, and much recent progress has been made in this direction (Gross and Taylor, 2001; Rastelli *et al.*, 2001; Shatashvili, 2001).

The monopole in Sec. III provides another example, which is now related to intersecting brane configurations. One can start from the $B=0$ description of a monopole in the $3+1$ $U(2)$ MSYM gauge theory of 2 $D3$-branes, which is a $D1$-brane suspended between the branes, i.e., extending in a transverse dimension perpendicular to the $D3$-branes, and with one end on each brane. One can show that this configuration not only reproduces the magnetic charge of the solution, but even obtain the full Nahm formalism from D-brane world-volume considerations (Diaconescu, 1997).

If one turns on B along the $D3$-branes, the essential point is now that the $D1$-brane is no longer perpendicular, it tilts. The end points of the $D1$-brane are magnetically charged with respect to the $U(1)$ gauge fields on the $D3$-branes. The background B field acts as a magnetic field, which pulls these charges apart. Tilting of the $D1$-brane makes its tension work in the opposite direction and stabilizes the configuration. This qualitative picture is confirmed by the profile of the explicit analytic solution (Gross and Nekrasov, 2000b).

Finally, let us mention the work of Myers (1999), which shows that Dp-branes in background Ramond-Ramond fields form bound states which can be thought of as $D(p+2)$-branes whose world volume is a product space with a fuzzy S^2. This noncommutativity plays a number of interesting roles in M theory (Myers, 2001) and its slightly different origin poses the challenge to find a broader picture which would incorporate it into our previous discussion.

E. Quantum effects and closed strings

The string theory derivation of noncommutative gauge theory allows one to compute its coupling to a background metric, at least for small variations of the flat background, by adding closed string graviton (and other) vertex operators in the world-sheet computations discussed above. Results of this type are given by many authors; we mention Das and Trivedi (2001), Liu and Michelson (2001b), and Okawa and Ooguri (2001a, 2001c) as representative. Analysis of these couplings to Ramond-Ramond closed string fields (Liu and Michelson, 2001a; Mukhi and Suryanarayana, 2001; Okawa and Ooguri, 2001b) was an important input into the result (45).

As a particularly simple example, the coupling to the graviton defines a stress-energy tensor in noncommutative gauge theory, which turns out to be precisely Eq. (40).

The general type of UV/IR relation we discussed at length in Sec. IV is very common in string theory. Perhaps its simplest form is visible in the computation of the annulus world sheet with one boundary on each of a pair of D-branes. This amplitude admits two pictures and two corresponding field theoretic interpretations: it can be thought of as describing emission of a closed string by one D-brane and its absorption by the other, a purely classical interpretation, and it can equally well be thought of as the sum of one-loop diagrams over all modes of the open string stretched between the pair of branes, a purely quantum interpretation. World-sheet duality implies that the two descriptions must be equal.

As discussed by Douglas *et al.* (1997), this leads very generally to the idea that for D-branes at substringy distances (i.e., with separation L as above satisfying $L \ll l_s$), conventional gravity is replaced by quantum effects in the world-volume gauge theory. In general this leads to different predictions from Einstein gravity or supergravity, but in special circumstances (e.g., with enough supersymmetry or in the large N limit) the substringy predictions can agree with gravity. Conversely, since L/T is a world-volume energy scale, taking $L \gg l_s$ accesses the UV limit of gauge theory amplitudes, and one sees that these are replaced by the IR limit of the gravitational description.

Although the decoupling limit takes $L \to 0$ and is thus in the substringy regime, one can ask whether nevertheless this potential connection to gravity can shed light on the nature of UV/IR mixing.

A strong sense in which this could be true would be if the new IR divergences could be described by adding additional light degrees of freedom in the effective theory, which would be directly analogous to the closed strings. This picture was explored in Minwalla *et al.* (2000) and Van Raamsdonk and Seiberg (2000). For example, if we introduce a field χ describing the new mode, we can then reproduce a singularity such as Eq. (107) by adding

$$\int \chi(\theta^{ik}\theta^{jl}G_{kl}\partial_i\partial_j)^{d/2-1}\chi + \chi\phi$$

to the effective Lagrangian. This particular example looks quite natural in $d=4$; one can also produce different power laws by postulating that χ propagates in a different number of dimensions than the original gauge theory, as is true of the closed strings in the analogy.

The main observation is then that, comparing with Eq. (167), the kinetic term for χ contains precisely the closed string metric g^{ij}, which is compatible with the idea. Indeed, there are cases in which this interpretation seems to be valid (Rajaraman and Rozali, 2000), namely, those in which the divergence is produced by a finite number of closed string modes because of supersymmetric cancellations, e.g., as in $\mathcal{N}=2$ SYM (Douglas and Li, 1996).

In general, however, the closed string picture is more complicated than this and one cannot identify a simple set of massless modes which reproduce the new IR effects (Andreev and Dorn, 2000a; Bilal *et al.*, 2000; Gomis *et al.*, 2000; Kiem and Lee, 2000). Furthermore the effective field theory required to reproduce higher loop effects does not look natural (Van Raamsdonk and Seiberg, 2000).

One can also argue that if this had worked in a more complicated situation, it would signal the breakdown of the decoupling limit we used to derive the theory from string theory (Gomis *et al.*, 2000). This is because exchange of a finite number of closed string modes would correspond to exchange of an infinite number of open string modes, including the massive open strings we dropped in the limit. Explicit consideration of the annulus diagram, however, shows that these massive open strings do not contribute in the limit.

F. AdS duals of noncommutative theories

One of the beautiful outcomes of string theory is the description of strongly coupled large N gauge theories by the AdS/CFT correspondence (Maldacena, 1998; Aharony *et al.*, 2000). In particular, one expects that the $D3$-brane realization of noncommutative gauge theory has a supergravity dual in the large-N, strong 't Hooft coupling limit. This was found in Hashimoto and Itzhaki (1999) and Maldacena and Russo (1999) and takes the form

$$ds^2 = l_s^2\sqrt{\lambda}\left[U^2(-dt^2+dx_1^2) + \frac{U^2}{1+\lambda\,\Delta^4 U^4}(dx_2^2+dx_3^2) + \frac{dU^2}{U^2} + d\Omega_5^2\right],$$

$$U^2 = \frac{1}{\lambda l_s^{22}}(x_4^2+\cdots+x_9^2),$$

$$e^{\phi} = \frac{\lambda}{4\pi N}\frac{1}{\sqrt{1+\lambda\Delta^4 U^4}},$$

$$B = -l_s^2\frac{\lambda\Delta^2 U^4}{1+\lambda\Delta^4 U^4}dx_2\wedge dx_3. \quad (174)$$

It is dual to the $\mathcal{N}=4$ $U(N)$ noncommutative gauge theory, with the noncommutativity $[x_2,x_3]\sim\Delta^2$, and 't Hooft coupling $\lambda = g_{YM}^2 N$. The transverse geometry is a five-sphere of radius R, $R^2 = l_s^2\sqrt{\lambda}$. For $U \ll 1/(\Delta\lambda^{1/4})$ the solution (174) approaches the $AdS_5\times S^5$ supergravity background, dual to ordinary large N $\mathcal{N}=4$ super-Yang-Mills theory. However, for large U, corresponding to the large energies in the gauge theory, the solution differs considerably: the dilaton flows, the B field approaches a constant in 23 directions, and the 23 directions collapse.

The large-U limit of anti–de Sitter space is a timelike boundary, and in the usual AdS/CFT correspondence the boundary values of fields are related to couplings of local operators in the gauge theory. The drastic modifications to Eq. (174) in this region have been argued by many authors to be associated with the lack of conventional local gauge invariant observables. Das and Rey (2000) have argued that the proportionality Eq. (16) between length and momenta, characteristic of the open Wilson loop and exploited in Eq. (113), emerges naturally from this picture.

See Berman *et al.* (2001), Danielsson *et al.* (2000), Li and Wu (2000), and Russo and Sheikh-Jabbari (2001) for further physics of this correspondence, and Elitzur *et al.* (2000) for a discussion of duality and Morita equivalence.

G. Timelike θ and exotic theories

So far we discussed the theories with spatial noncommutativity, which arise from Dirichlet branes with a B field along the spatial directions. There are also limits with timelike B field, leading to exotic noncommutative string and membrane theories, the noncommutative open string theory (Gopakumar, Maldacena, *et al.*, 2000; Seiberg *et al.*, 2000b), the open membrane theory (Bergshoeff *et al.*, 2000; Gopakumar, Minwalla, *et al.*, 2000), and the open Dp-brane theories (Gopakumar, Minwalla, *et al.*, 2000; Harmark, 2001). These appear to evade the arguments against timelike noncommutativity in field theory.

Let us start with the $D3$-brane in $\mathbb{I}b$ string theory in a large spatial B field. This theory exhibits S duality, which maps electric field to magnetic field and vice versa. Since constant spatial B-field is gauge equivalent to a constant magnetic field on the brane, S duality must map this background into one with large electric field. This should, in turn, lead to space-time noncommutativity, with $\theta^{i0}\neq 0$.

Thus, combining accepted elements of M/string duality, one concludes that our previous arguments against timelike noncommutativity must have some loophole. This is correct and one can in fact take the large electric-field limit for any of the Dp-brane theories, not just $p=3$. However, the details are rather different from the spacelike case. An electric field in open string theory cannot be taken to be larger than a critical value $E_c = 1/2\pi l_s^2$ (Burgess, 1987; Bachas, 1996). As one approaches this limit, since the ends of the open string carry opposite electric charges, its effective tension goes to zero, and any attempt to reach $E>E_c$ will be screened by string formation in the vacuum.

So, one takes the limit $E\to E_c$ while keeping the effective open string tension $l_{so}^2 = l_s^2\sqrt{E^2/(E_c^2-E^2)}$ finite. It turns out that while one manages to decouple closed string modes, the open string excitations remain in the spectrum. The resulting theory, noncommutative open string theory, is apparently a true string theory with an infinite number of particlelike degrees of freedom. This is consistent both with the earlier arguments and with the general idea that an action with timelike noncommutativity effectively contains an infinite number of time derivatives, thereby enhancing the number of degrees of freedom.

The noncommutative open string theory contradicts standard arguments from world-sheet duality that the open string theory must contain closed strings. How this works can be seen explicitly in the annulus diagram; extra phases present in the nonplanar open string diagrams make the would-be closed string poles vanish. These effects apparently resolve the unitarity problems of Sec. IV.C as well.

Since this string theory decouples from gravity, it provides a system in which the Hagedorn transition of string theory can be analyzed in a clean situation, free from black-hole thermal effects and other complications of gravitating systems (Gubser *et al.*, 2000; Barbon and Rabinovici, 2001).

A similar limit can be taken starting with the M-theory five-brane (the M5-brane). Its world-volume theory is also a gauge theory, but now involving a rank two antisymmetric tensor potential. The membrane is allowed to end on an M5-brane and thus parallel M5-branes come with light degrees of freedom (open membranes) directly analogous to the open strings which end on a Dirichlet brane. In the limit that the branes coincide, the resulting light degrees of freedom are governed by a nontrivial fixed-point theory in $5+1$ dimensions, usually called the (2,0) theory after its supersymmetry algebra. A similar limit in IIa theory leads to little string theory, these theories have recently been reviewed by Aharony (2000).

To get a noncommutative version of the open membrane theory, we start with N coincident M5-branes with the background 3-form strength and the metric:

$$H_{012}=M_p^3\tanh\beta,\quad H_{345}=-\sqrt{8}M_p^3\frac{\sinh\beta}{(e^{\beta}\cosh\beta)^{3/2}},$$

$$g_{\mu\nu}=\eta_{\mu\nu},\quad \mu,\nu=0,1,2,$$

$$g_{ij}=\frac{1}{e^{\beta}\cosh\beta}\delta_{ij},\quad i,j=3,4,5,$$

$$g_{MN}=\frac{1}{e^{\beta}\cosh\beta}\delta_{MN}\quad M,N=6,7,8,9,11, \tag{175}$$

where M_p is the gravitational mass scale, and β is the parameter to tune. The critical limit is achieved by taking $\beta\to\infty$, while keeping

$$M_{\rm eff}=M_p\left(\frac{1}{e^{\beta}\cosh\beta}\right)^{1/3}$$

(the mass scale for the membrane stretched spatially in the 1,2 directions) finite. In this limit $M_p/M_{\rm eff}\to\infty$ and so the open membranes propagating along the M5-branes will decouple from gravity.

Finally, one can start with the NS five-brane in either of the type-II string theories. Dirichlet p-branes can end on the NS five-brane, with all even p in IIa and all odd p in IIb, leading to open Dp-brane degrees of freedom. Each of these is charged under a specific Ramond-Ramond gauge field, and by taking a near critical electric background for one of these fields, one can again reach a decoupling limit in which only the corresponding open brane degrees of freedom remain, discussed in Gopakumar, Minwalla, *et al.* (2000) and Harmark (2001).

All of these theories are connected by a web of dualities analogous to those connecting the conventional decoupled brane theories and the bulk theories which contain them. For example, compactifying OM theory on a circle leads to the noncommutative open string $4+1$ theory; conversely the strong coupling limit of this noncommutative open string theory has a geometric description (OM theory) just as did the strong-coupling limit of the IIa string (M theory).

VIII. CONCLUSIONS

Field theory can be generalized to space-time with noncommuting coordinates. Much of the formalism is very parallel to that of conventional field theory and especially with the large-N limit of conventional field theory. Although not proven, it appears that quantum noncommutative field theories under certain restrictions (say with spacelike noncommutativity and some supersymmetry) are renormalizable and sensible.

Their physics is similar enough to conventional field theory to make comparisons possible, and different enough to make them interesting. To repeat some of the highlights, we found that noncommutative gauge symmetry includes space-time symmetries, that nonsingular soliton solutions exist in higher-dimensional scalar field theory and in noncommutative Maxwell theory, that UV divergences can be transmuted into new IR effects, and that noncommutative gauge theories can have more dualities than their conventional counterparts.

Much of our knowledge of conventional field theory still awaits a noncommutative counterpart. Throughout the review many questions were left open, such as the meaning of the IR divergences found in Sec. IV.C, the potential nonperturbative role of the solitons and instantons of Sec. III, the meaning of the high-energy behavior discussed in Sec. IV.D, and the high-temperature behavior.

One central problem is to properly understand the renormalization group. Even if one can directly adapt existing renormalization-group technology, it seems very likely that theories with such a different underlying concept of space and time will admit other and perhaps more suitable formulations of the renormalization group. This might lead to insights into nonlocality of the sort hoped for in the Introduction. Questions about the existence of quantized noncommutative theories could then be settled by using the renormalization group starting with a good regulated nonperturbative definition of the theory, perhaps that of Ambjorn *et al.* (1999) or perhaps along other lines as discussed in Sec. VI.A.

The techniques of exactly solvable field theory, which are so fruitful in two dimensions, await possible noncommutative generalization. These might be particularly relevant for the quantum Hall application.

It is not impossible that noncommutative field theory has some direct relevance for particle physics phenomenology, or possible relevance in the early universe. Possible signatures of noncommutativity in QED and the standard model are discussed by several authors[6] who work with a general extension of the standard model allowing for Lorentz violation (see Kostelecky, 2001, and references therein) and argue that atomic clock-comparison experiments lead to a bound in the QED sector of $|\theta|<(10\ \text{TeV})^{-2}$. A noncommutative brane world scenario is developed by Pilo and Riotto (2001), and cosmological applications are discussed by Alexander and Magueijo (2001) and Chu *et al.* (2000).

This motivation as well as the motivation mentioned in the Introduction of modeling position-space uncertainty in quantum gravity might be better served by Lorentz-invariant theories, and in pursuing the second of these motivations it has been suggested by Doplicher (2001; Doplicher *et al.*, 1994) that such theories could be defined by treating the noncommutativity parameter θ as a dynamical variable. The space-time stringy uncertainty principle of Yoneya (1987) leads to related considerations (Yoneya, 2001).

While we hope that our discussion has demonstrated that noncommutative field theory is a subject of intrinsic interest, at present its primary physical application stems from the fact that it emerges from limits of M theory and

[6]See, for example, Arfaei and Yavartanoo, 2000; Hewett *et al.*, 2000; Mazumdar and Sheikh-Jabbari, 2000; Mocioiu *et al.*, 2000; Baek *et al.*, 2001; Carroll *et al.*, 2001; and Mathews, 2001.

string theory, and it seems clear at this point that the subject will have lasting importance in this context. So far its most fruitful applications have been to duality and to the understanding of solitons and branes in string theory. It is quite striking how much structure which had been considered essentially stringy is captured by these much simpler theories.

Noncommutativity enters into open string theory essentially because open strings interact by joining at their ends, and the choice of one or the other of the two ends corresponds formally to acting on the corresponding field by multiplication on the left or on the right; these are different. This is such a fundamental level that it has long been thought that noncommutativity should be central to the subject. So far, the developments we discussed look like a very promising start towards realizing this idea. Progress is also being made on the direct approach, through string field theory based on noncommutative geometry, and we believe that many of the ideas we have discussed will reappear in this context.

Whether noncommutativity is a central concept in the full string or M theory is less clear. Perhaps the best reason to think this is that it appears so naturally from definition of the M(atrix) theory definition, which can include all of M theory in certain backgrounds. On the other hand, this also points to the weakness of our present understanding: these are very special backgrounds. We do not now have formulations of M theory in general backgrounds; this includes the backgrounds of primary physical interest with four observable dimensions. A related point is that, in string theory, one thinks of the background as defined within the gravitational or closed string sector, and the role of noncommutativity in this sector is less clear.

An important question in noncommutative field theory is to what extent the definitions can be generalized to spaces besides Minkowski space and the torus, which are not flat. D-brane constructions in other backgrounds analogous to what we have discussed for flat space seem to lead to theories with finitely many degrees of freedom, as in Alekseev *et al.* (1999). It might be that noncommutative field theories can arise as large-N limits of these models, but at present this is not clear.

Even for group manifolds and homogeneous spaces, where mathematical definitions exist, the physics of these theories is not clear and deserves more study. As we discussed in Sec. VI, there are many more interesting noncommutative algebras arising from geometric constructions, which would be interesting test cases as well.

At the present state of knowledge, it is conceivable that, contrary to our intuition from the study of both gravity and perturbative string theory, special backgrounds such as flat space, anti–de Sitter space, orbifolds, and perhaps others, which correspond in M theory to simple gauge theories and noncommutative gauge theories, play a preferred role in the theory, and that all others will be derived from these. In this picture, the gravitational or closed string degrees of freedom would be derived from the gauge theory or open string theory, as has been argued to happen at substringy distances, in M(atrix) theory and in the AdS/CFT correspondence.

If physically realistic backgrounds could be derived this way, then this might be a satisfactory outcome. It would radically change our viewpoint on space-time and might predict that many backgrounds that would be acceptable solutions of gravity are in fact not allowed in M theory. It is far too early to judge this point, however, and it seems to us that at present such hopes are founded more on our lack of understanding of M theory in general backgrounds than on anything else. Perhaps noncommutative field theories in more general backgrounds, or in a more background-independent formulation, will serve as useful analogs to M/string theory for this question as well.

In any case, our general conclusion has to be that the study of noncommutative field theory, as well as the more mysterious theories which have emerged from the study of superstring duality (a few of which we mentioned in Sec. VII.G), has shown that field theory is a much broader concept than had been dreamed of even a few years ago. It surely has many more surprises in store for us, and we hope this review will stimulate the reader to pick up and continue the story.

ACKNOWLEDGMENTS

We would like to thank Alain Connes, Jeff Harvey, Albert Schwarz, and Nathan Seiberg for comments on the manuscript, and the ITP, Santa Barbara for hospitality. This work was supported in part by DOE Grant No. DE-FG05-90ER40559, and by RFFI Grants Nos. 01-01-00549 and 00-15-96557.

REFERENCES

Ablowitz, M. J., and P. A. Clarkson, 1991, *Solitons, Nonlinear Evolution Equations and Inverse Scattering*, LMS Lecture Note Series Vol. 149 (Cambridge University Press, Cambridge).

Abou-Zeid, M., and H. Dorn, 2001a, e-print hep-th/0104244.

Abou-Zeid, M., and H. Dorn, 2000b, Phys. Lett. B **504**, 165.

Abouelsaood, A., C. G. Callan, C. R. Nappi, and S. A. Yost, 1987, Nucl. Phys. B **280**, 599.

Aganagic, M., R. Gopakumar, S. Minwalla, and A. Strominger, 2001, J. High Energy Phys. **04**, 001.

Aharony, O., 2000, Class. Quantum Grav. **17**, 929.

Aharony, O., S. S. Gubser, J. Maldacena, H. Ooguri, and Y. Oz, 2000, Phys. Rep. **323**, 183.

Akhmedov, E. T., P. DeBoer, and G. W. Semenoff, 2000, e-print hep-th/0010003.

Akhmedov, E. T., P. DeBoer, and G. W. Semenoff, 2001, e-print hep-th/0103199.

Alekseev, A. Y., A. Recknagel, and V. Schomerus, 1999, J. High Energy Phys. **09**, 023.

Alexander, S., and J. Magueijo, 2001, e-print hep-th/0104093.

Alvarez-Gaume, L., J. L. F. Barbon, and R. Zwicky, 2001, e-print hep-th/0103069.

Ambjorn, J., Y. M. Makeenko, J. Nishimura, and R. J. Szabo, 1999, J. High Energy Phys. **11**, 029.

Ambjorn, J., Y. M. Makeenko, J. Nishimura, and R. J. Szabo, 2000, Phys. Lett. B **480**, 399.
Andreev, O., and H. Dorn, 2000a, Nucl. Phys. B **583**, 145.
Andreev, O., and H. Dorn, 2000b, Phys. Lett. B **476**, 402.
Arcioni, G., J. L. F. Barbon, J. Gomis, and M. A. Vazquez-Mozo, 2000, J. High Energy Phys. **06**, 038.
Arcioni, G., and M. A. Vazquez-Mozo, 2000, J. High Energy Phys. **01**, 028.
Ardalan, F., and N. Sadooghi, 2000, e-print hep-th/0009233.
Arfaei, H., and M. H. Yavartanoo, 2000, e-print hep-th/0010244.
Armoni, A., 2001, Nucl. Phys. B **593**, 229.
Armoni, A., R. Minasian, and S. Theisen, 2001, e-print hep-th/0102007.
Arovas, D., J. R. Schrieffer, and F. Wilczek, 1984, Phys. Rev. Lett. **53**, 722.
Asakawa, T., and I. Kishimoto, 2000, Nucl. Phys. B **591**, 611.
Atiyah, M., V. Drinfeld, N. Hitchin, and Y. Manin, 1978, Phys. Lett. **65A**, 185.
Bachas, C., 1996, Phys. Lett. B **374**, 37.
Baek, S., D. K. Ghosh, X.-G. He, and W. Y. P. Hwang, 2001, e-print hep-ph/0103068.
Bak, D., 2000, Phys. Lett. B **495**, 251.
Bak, D., K. Lee, and J.-H. Park, 2000, e-print hep-th/0011099.
Banks, T., W. Fischler, S. H. Shenker, and L. Susskind, 1997, Phys. Rev. D **55**, 5112.
Barbon, J. L. F., and E. Rabinovici, 2001, e-print hep-th/0104169.
Bars, I., H. Kajiura, Y. Matsuo, and T. Takayanagi, 2001, Phys. Rev. D **63**, 086001.
Bars, I., and D. Minic, 2000, Phys. Rev. D **62**, 105018.
Bars, I., M. M. Sheikh-Jabbari, and M. A. Vasiliev, 2001, e-print hep-th/0103209.
Bassetto, A., L. Griguolo, G. Nardelli, and F. Vian, 2001, e-print hep-th/0105257.
Bateman, H., 1953, *Higher Transcendental Functions* (McGraw-Hill, New York).
Baulieu, L., A. Losev, and N. Nekrasov, 2001, e-print hep-th/0106042.
Bayen, F., M. Flato, C. Fronsdal, A. Lichnerowicz, and D. Sternheimer, 1978, Ann. Phys. (N.Y.) **111**, 61.
Bellisai, D., J. M. Isidro, and M. Matone, 2000, J. High Energy Phys. **10**, 026.
Bellissard, J., A. van Elst, and H. Schulz-Baldes, 1993, e-print cond-mat/9301005.
Bergman, A., O. J. Ganor, and J. L. Karczmarek, 2001, e-print hep-th/0101095.
Bergshoeff, E., D. S. Berman, J. P. van der Schaar, and P. Sundell, 2000, Phys. Lett. B **492**, 193.
Berman, D. S., *et al.*, 2001, J. High Energy Phys. **05**, 002.
Bigatti, D., and L. Susskind, 2000, Phys. Rev. D **62**, 066004.
Bilal, A., C.-S. Chu, and R. Russo, 2000, Nucl. Phys. B **582**, 65.
Bogomolny, E. B., 1976, Sov. J. Nucl. Phys. **24**, 449.
Bonora, L., M. Schnabl, M. M. Sheikh-Jabbari, and A. Tomasiello, 2000, Nucl. Phys. B **589**, 461.
Bonora, L., M. Schnabl, and A. Tomasiello, 2000, Phys. Lett. B **485**, 311.
Bordemann, M., and J. Hoppe, 1993, Phys. Lett. B **317**, 315.
Brace, D., and B. Morariu, 1999, J. High Energy Phys. **02**, 004.
Burgess, C. P., 1987, Nucl. Phys. B **294**, 427.
Campbell, B. A., and K. Kaminsky, 2001, e-print hep-th/0102022.
Carroll, S. M., J. A. Harvey, V. A. Kostelecky, C. D. Lane, and T. Okamoto, 2001, e-print hep-th/0105082.
Cattaneo, A. S., and G. Felder, 2000, Commun. Math. Phys. **212**, 591.
Cattaneo, A. S., and G. Felder, 2001, Mod. Phys. Lett. A **16**, 179.
Chen, G.-H., and Y.-S. Wu, 2001, e-print hep-th/0103020.
Chepelev, I., and R. Roiban, 2000, J. High Energy Phys. **05**, 037.
Cheung, Y.-K. E., and M. Krogh, 1998, Nucl. Phys. B **528**, 185.
Chu, C.-S., B. R. Greene, and G. Shiu, 2000, e-print hep-th/0011241.
Coleman, S., 1985, *Aspects of Symmetry* (Cambridge University Press, Cambridge).
Connes, A., 1980, C.R. Seances Acad. Sci., Ser. A **290**, 599.
Connes, A., 1982, pre-prints I.H.E.S. M/82/53, M/83/19.
Connes, A., 1994, *Noncommutative Geometry* (Academic Press, San Diego).
Connes, A., 1995, in *Quantum Symmetries*, edited by A. Connes, K. Gawedzki, and J. Zinn-Justin, *NATO Advanced Study Institute*, 64th Session, Les Houches (Elsevier, Amsterdam).
Connes, A., 1996, Commun. Math. Phys. **182**, 155.
Connes, A., 2000a, e-print math.QA/0011193.
Connes, A., 2000b, J. Math. Phys. **41**, 3832.
Connes, A., M. R. Douglas, and A. Schwarz, 1998, J. High Energy Phys. **02**, 003.
Connes, A., and J. Lott, 1991, Nucl. Phys. B, Proc. Suppl. **18**, 29.
Connes, A., and M. Rieffel, 1987, *Operator Algebras and Mathematical Physics (Iowa City, 1985)* (AMS, Providence), Vol. 62, p. 237.
Cornalba, L., 1999, e-print hep-th/9909081.
Cremmer, E., B. Julia, and J. Scherk, 1978, Phys. Lett. **76B**, 409.
Curtright, T., T. Uematsu, and C. Zachos, 2000, e-print hep-th/0011137.
Dai, J., R. G. Leigh, and J. Polchinski, 1989, Mod. Phys. Lett. A **4**, 2073.
Danielsson, U. H., A. Guijosa, M. Kruczenski, and B. Sundborg, 2000, J. High Energy Phys. **05**, 028.
Das, S. R., and S.-J. Rey, 2000, Nucl. Phys. B **590**, 453.
Das, S. R., and S. P. Trivedi, 2001, J. High Energy Phys. **02**, 046.
Dasgupta, K., S. Mukhi, and G. Rajesh, 2000, J. High Energy Phys. **06**, 022.
Derrick, G. M., 1964, J. Math. Phys. **5**, 1252.
de Wit, B., J. Hoppe, and H. Nicolai, 1988, Nucl. Phys. B **305**, 545.
DeWilde, M., and P. Lecomte, 1983, Lett. Math. Phys. **7**, 487.
DeWitt, B., 1962, in *Gravitation*, edited by L. Witten (Wiley, New York), pp. 266–381.
Dhar, A., and Y. Kitazawa, 2001a, J. High Energy Phys. **02**, 004.
Dhar, A., and Y. Kitazawa, 2001b, Phys. Rev. D **63**, 125005.
Diaconescu, D.-E., 1997, Nucl. Phys. B **503**, 220.
Dijkgraaf, R., 1998, in *Quantum Symmetries*, edited by A. Connes, K. Gawedzki, and J. Zinn-Justin, *NATO Advanced Study Institute*, 64th Session, Les Houches (Elsevier, Amsterdam).
Dijkgraaf, R., E. Verlinde, and H. Verlinde, 1997, Nucl. Phys. B **500**, 43.
Doplicher, S., 2001, e-print hep-th/0105251.

Doplicher, S., K. Fredenhagen, and J. E. Roberts, 1994, Phys. Lett. B **331**, 39.

Dorey, N., T. J. Hollowood, and V. V. Khoze, 2000, e-print hep-th/0010015.

Douglas, M. R., 1996, e-print hep-th/9610041.

Douglas, M. R., 1998, e-print hep-th/9807235.

Douglas, M. R., 1999, e-print hep-th/9901146.

Douglas, M. R., and C. Hull, 1998, J. High Energy Phys. **02**, 008.

Douglas, M. R., D. Kabat, P. Pouliot, and S. H. Shenker, 1997, Nucl. Phys. B **485**, 85.

Douglas, M. R., and M. Li, 1996, e-print hep-th/9604041.

Durhuus, B., T. Jonsson, and R. Nest, 2001, Phys. Lett. B **500**, 320.

Eguchi, T., and R. Nakayama, 1983, Phys. Lett. **122B**, 59.

Elitzur, S., B. Pioline, and E. Rabinovici, 2000, J. High Energy Phys. **10**, 011.

Fairlie, D., 1964, Proc. Cambridge Philos. Soc. **60**, 581.

Fedosov, B., 1996, *Deformation Quantization and Index Theory* (Akademie Verlag, Berlin).

Fedosov, B. V., 1994, J. Diff. Geom. **40**, 213.

Ferrara, S., and M. A. Lledo, 2000, J. High Energy Phys. **05**, 008.

Filk, T., 1996, Phys. Lett. B **376**, 53.

Fischler, W., J. Gomis, E. Gorbatov, A. Kashani-Poor, S. Paban, and P. Pouliot, 2000, J. High Energy Phys. **05**, 024.

Fischler, W., E. Gorbatov, A. Kashani-Poor, R. McNees, S. Paban, and P. Pouliot, 2000, J. High Energy Phys. **06**, 032.

Fradkin, E., 1991, *Field Theories of Condensed Matter Systems* (Addison-Wesley, Redwood City, CA).

Fradkin, E. S., and A. A. Tseytlin, 1985, Phys. Lett. **163B**, 123.

Furuuchi, K., 2000, e-print hep-th/0010006.

Garousi, M. R., 2000, Nucl. Phys. B **579**, 209.

Gerhold, A., J. Grimstrup, H. Grosse, L. Popp, M. Schweda, and R. Wulkenhaar, 2000, e-print hep-th/0012112.

Girotti, H. O., M. Gomes, V. O. Rivelles, and A. J. da Silva, 2000, Nucl. Phys. B **587**, 299.

Girotti, H. O., M. Gomes, V. O. Rivelles, and A. J. da Silva, 2001, e-print hep-th/0102101.

Girvin, S., 1999, proceedings, Ecole d'Ete Les Houches, July 1998, e-print cond-mat/9907002.

Girvin, S. M., A. H. MacDonald, and P. M. Platzman, 1985, Phys. Rev. Lett. **54**, 581.

Gomis, J., K. Kamimura, and J. Llosa, 2001, Phys. Rev. D **63**, 045003.

Gomis, J., M. Kleban, T. Mehen, M. Rangamani, and S. Shenker, 2000, J. High Energy Phys. **08**, 011.

Gomis, J., and T. Mehen, 2000, Nucl. Phys. B **591**, 265.

Gonzalez-Arroyo, A., and C. P. Korthals Altes, 1983, Phys. Lett. **131B**, 396.

Gonzalez-Arroyo, A., and M. Okawa, 1983, Phys. Rev. D **27**, 2397.

Gopakumar, R., M. Headrick, and M. Spradlin, 2001, e-print hep-th/0103256.

Gopakumar, R., J. Maldacena, S. Minwalla, and A. Strominger, 2000, J. High Energy Phys. **06**, 036.

Gopakumar, R., S. Minwalla, N. Seiberg, and A. Strominger, 2000, J. High Energy Phys. **08**, 008.

Gopakumar, R., S. Minwalla, and A. Strominger, 2000, J. High Energy Phys. **05**, 020.

Gracia-Bondia, J. M., and C. P. Martin, 2000, Phys. Lett. B **479**, 321.

Gracia-Bondia, J. M., J. C. Varilly, and H. Figueroa, 2001, *Elements of Noncommutative Geometry* (Birkhaeuser, Boston).

Green, M. B., J. H. Schwarz, and E. Witten, 1987, *Superstring Theory* (Cambridge University Press, Cambridge).

Griguolo, L., and M. Pietroni, 2001a, e-print hep-th/0102070.

Griguolo, L., and M. Pietroni, 2001b, e-print hep-th/0104217.

Gross, D. J., A. Hashimoto, and N. Itzhaki, 2000, e-print hep-th/0008075.

Gross, D. J., and N. A. Nekrasov, 2000a, J. High Energy Phys. **10**, 021.

Gross, D. J., and N. A. Nekrasov, 2000b, J. High Energy Phys. **07**, 034.

Gross, D. J., and N. A. Nekrasov, 2001, J. High Energy Phys. **03**, 044.

Gross, D. J., and W. Taylor, 2001, e-print hep-th/0105059.

Grosse, H., C. Klimčik, and P. Presnajder, 1996, Int. J. Theor. Phys. **35**, 231.

Gubser, S. S., S. Gukov, I. R. Klebanov, M. Rangamani, and E. Witten, 2000, e-print hep-th/0009140.

Gubser, S. S., and S. L. Sondhi, 2000, e-print hep-th/0006119.

Gurarie, V., and A. Zee, 2000, e-print cond-mat/0008163.

Hadasz, L., U. Lindstrom, M. Rocek, and R. von Unge, 2001, e-print hep-th/0104017.

Hamanaka, M., and S. Terashima, 2001, J. High Energy Phys. **03**, 034.

Harmark, T., 2001, Nucl. Phys. B **593**, 76.

Harvey, J. A., 2001a, e-print hep-th/0102076.

Harvey, J. A., 2001b, e-print hep-th/0105242.

Harvey, J. A., P. Kraus, and F. Larsen, 2000, J. High Energy Phys. **12**, 024.

Harvey, J. A., P. Kraus, F. Larsen, and E. J. Martinec, 2000, J. High Energy Phys. **07**, 042.

Harvey, J. A., and G. Moore, 2000, e-print hep-th/0009030.

Hashimoto, A., and N. Itzhaki, 1999, Phys. Lett. B **465**, 142.

Hashimoto, A., and N. Itzhaki, 2000, e-print hep-th/0012093.

Hashimoto, K., 2000, J. High Energy Phys. **12**, 023.

Hayakawa, M., 1999, e-print hep-th/9912167.

Hellerman, S., and M. Van Raamsdonk, 2001, e-print hep-th/0103179.

Hewett, J. L., F. J. Petriello, and T. G. Rizzo, 2000, e-print hep-th/0010354.

Ho, P.-M., and Y.-S. Wu, 1997, Phys. Lett. B **398**, 52.

Ho, P.-M., and Y.-S. Wu, 1998, Phys. Rev. D **58**, 066003.

Hofman, C., and E. Verlinde, 1999, Nucl. Phys. B **547**, 157.

Hollowood, T. J., V. V. Khoze, and G. Travaglini, 2001, e-print hep-th/0102045.

Hoppe, J., 1982, Ph.D. thesis, MIT.

Hoppe, J., 1989, Int. J. Mod. Phys. A **4**, 5235.

Ishibashi, N., S. Iso, H. Kawai, and Y. Kitazawa, 2000, Nucl. Phys. B **573**, 573.

Ishibashi, N., H. Kawai, Y. Kitazawa, and A. Tsuchiya, 1997, Nucl. Phys. B **498**, 467.

Jackson, M. G., 2001, e-print hep-th/0103217.

Jatkar, D. P., G. Mandal, and S. R. Wadia, 2000, J. High Energy Phys. **09**, 018.

Jurco, B., L. Moller, S. Schraml, P. Schupp, and J. Wess, 2001, e-print hep-th/0104153.

Jurco, B., P. Schupp, and J. Wess, 2000, Nucl. Phys. B **584**, 784.

Kallin, C., and B. I. Halperin, 1984, Phys. Rev. B **30**, 5655.

Kapusta, J., 1990, *Finite Temperature Field Theory* (Cambridge University Press, Cambridge).

Khoze, V. V., and G. Travaglini, 2001, J. High Energy Phys. **01**, 026.

Kiem, Y., and S. Lee, 2000, Nucl. Phys. B **586**, 303.
Kinar, Y., G. Lifschytz, and J. Sonnenschein, 2001, e-print hep-th/0105089.
Klauder, J. R., and E. C. G. Sudarshan, 1968, *Fundamentals of Quantum Optics* (Benjamin, New York).
Konechny, A., and A. Schwarz, 1999, Nucl. Phys. B **550**, 561.
Konechny, A., and A. Schwarz, 2000a, Nucl. Phys. B **591**, 667.
Konechny, A., and A. Schwarz, 2000b, e-print hep-th/0012145.
Kontsevich, M., 1997, e-print q-alg/9709040.
Kostelecky, V. A., 2001, e-print hep-th/0104227.
Kuiper, N. H., 1965, Topology **3**, 19.
Landau, L. D., E. M. Lifshitz, and A. Pitaevskii, 1960, *Field Theory of Continuous Media* (Butterworth Heinemann, Oxford).
Landsteiner, K., E. Lopez, and M. H. G. Tytgat, 2001, e-print hep-th/0104133.
Laughlin, R. B., 1983, Phys. Rev. Lett. **50**, 1395.
Laughlin, R. B., 1985, Physica B **126**, 254.
Lee, B.-H., K. Lee, and H. S. Yang, 2001, Phys. Lett. B **498**, 277.
Li, M., 2000, e-print hep-th/0011170.
Li, M., and Y.-S. Wu, 2000, Phys. Rev. Lett. **84**, 2084.
Lindstrom, U., M. Rocek, and R. von Unge, 2000, J. High Energy Phys. **12**, 004.
Liu, H., 2000, e-print hep-th/0011125.
Liu, H., and J. Michelson, 2000, e-print hep-th/0008205.
Liu, H., and J. Michelson, 2001a, e-print hep-th/0104139.
Liu, H., and J. Michelson, 2001b, e-print hep-th/0101016.
Lozano, G. S., E. F. Moreno, and F. A. Schaposnik, 2001, Phys. Lett. B **504**, 117.
Lugo, A. R., 2000, e-print hep-th/0012268.
Madore, J., 1992, Class. Quantum Grav. **9**, 69.
Madore, J., 1999, *An Introduction to Noncommutative Geometry and its Physical Applications*, LMS Lecture Note Series Vol. 257, 2nd ed. (Cambridge University Press, Cambridge).
Maldacena, J., 1998, Adv. Theor. Math. Phys. **2**, 231.
Maldacena, J. M., and J. G. Russo, 1999, J. High Energy Phys. **09**, 025.
Martin, C. P., 2001, Mod. Phys. Lett. A **16**, 311.
Martinec, E. J., and G. Moore, 2001, e-print hep-th/0101199.
Mathews, P., 2001, Phys. Rev. D **63**, 075007.
Matusis, A., L. Susskind, and N. Toumbas, 2000, J. High Energy Phys. **12**, 002.
Mazumdar, A., and M. M. Sheikh-Jabbari, 2000, e-print hep-ph/0012363.
Mehen, T., and M. B. Wise, 2000, J. High Energy Phys. **12**, 008.
Minwalla, S., M. Van Raamsdonk, and N. Seiberg, 2000, J. High Energy Phys. **02**, 020.
Mocioiu, I., M. Pospelov, and R. Roiban, 2000, Phys. Lett. B **489**, 390.
Montonen, C., and D. Olive, 1977, Phys. Lett. **72B**, 117.
Moreno, E. F., and F. A. Schaposnik, 2000, J. High Energy Phys. **03**, 032.
Moreno, E. F., and F. A. Schaposnik, 2001, Nucl. Phys. B **596**, 439.
Motl, L., 1997, e-print hep-th/9701025.
Mukhi, S., and N. V. Suryanarayana, 2001, e-print hep-th/0104045.
Myers, R. C., 1999, J. High Energy Phys. **12**, 022.
Myers, R. C., 2001, e-print hep-th/0106178.
Nahm, W., 1980, Phys. Lett. **90B**, 413.
Nekrasov, N., and A. Schwarz, 1998, Commun. Math. Phys. **198**, 689.
Nekrasov, N. A., 2000, e-print hep-th/0011095.
Nekrasov, N. A., 2001, I.H.E.S. preprint.
Nielsen, H. B., and P. Olesen, 1973, Nucl. Phys. B **61**, 45.
Okawa, Y., and H. Ooguri, 2001a, e-print hep-th/0103124.
Okawa, Y., and H. Ooguri, 2001b, e-print hep-th/0104036.
Okawa, Y., and H. Ooguri, 2001c, Nucl. Phys. B **599**, 55.
Olshanetsky, M., and A. Perelomov, 1976, Invent. Math. **31**, 93.
Palais, R. S., 1965, Topology **3**, 271.
Pernici, M., A. Santambrogio, and D. Zanon, 2001, Phys. Lett. B **504**, 131.
Petriello, F. J., 2001, Nucl. Phys. B **601**, 169.
Pilo, L., and A. Riotto, 2001, J. High Energy Phys. **03**, 015.
Polchinski, J., 1984, Nucl. Phys. B **231**, 269.
Polchinski, J., 1995, Phys. Rev. Lett. **75**, 4724.
Polchinski, J., 1996a, Rev. Mod. Phys. **68**, 1245.
Polchinski, J., 1996b, e-print hep-th/9611050.
Polchinski, J., 1998, *Superstring Theory* (Cambridge University Press, Cambridge).
Polyakov, A. M., 1987, *Gauge Fields and Strings* (Harwood, Chur, Switzerland).
Polychronakos, A. P., 1991, Phys. Lett. B **266**, 29.
Polychronakos, A. P., 2000, Phys. Lett. B **495**, 407.
Polychronakos, A. P., 2001, J. High Energy Phys. **04**, 011.
Prange, R., and S. Girvin, 1987, *The Quantum Hall Effect* (Springer, New York).
Rajaraman, A., and M. Rozali, 2000, J. High Energy Phys. **04**, 033.
Rastelli, L., A. Sen, and B. Zwiebach, 2001, e-print hep-th/0105058.
Rieffel, M., 1981, Pac. J. Math. **93**, 415.
Rieffel, M. A., 1988, Can. J. Math. **40**, 257.
Rieffel, M. A., 1998, Contemp. Math. **228**, 315.
Rieffel, M. A., and A. Schwarz, 1998, e-print math.QA/9803057.
Rozali, M., and M. Van Raamsdonk, 2000, e-print hep-th/0012065.
Ruiz, F. R., 2001, Phys. Lett. B **502**, 274.
Russo, J. G., and M. M. Sheikh-Jabbari, 2001, Nucl. Phys. B **600**, 62.
Sarkar, S., and B. Sathiapalan, 2001, e-print hep-th/0104106.
Schafer, T., and E. V. Shuryak, 1998, Rev. Mod. Phys. **70**, 323.
Schnabl, M., 2000, J. High Energy Phys. **11**, 031.
Schomerus, V., 1999, J. High Energy Phys. **06**, 030.
Schwarz, A., 1998, Nucl. Phys. B **534**, 720.
Schwarz, A., 2001, e-print hep-th/0102182.
Schwarz, J. H., 2001, e-print hep-th/0103165.
Seiberg, N., 2000, J. High Energy Phys. **09**, 003.
Seiberg, N., L. Susskind, and N. Toumbas, 2000a, J. High Energy Phys. **06**, 044.
Seiberg, N., L. Susskind, and N. Toumbas, 2000b, J. High Energy Phys. **06**, 021.
Seiberg, N., and E. Witten, 1994, Nucl. Phys. B **426**, 19.
Seiberg, N., and E. Witten, 1999, J. High Energy Phys. **09**, 032.
Sen, A., 1998, J. High Energy Phys. **9808**, 012.
Shatashvili, S. L., 2001, e-print hep-th/0105076.
Sheikh-Jabbari, M. M., 1999, Phys. Lett. B **455**, 129.
Sinova, J., V. Meden, and S. Girvin, 2000, Phys. Rev. B **62**, 2008.
Snyder, H. S., 1947, Phys. Rev. **71**, 38.
Sternheimer, D., 1998, in *Particles, Fields, and Gravitation*, edited by Jakub Rembieliński, AIP Conf. Proc. No. 453 (AIP, Woodbury, NY), pp. 107–145.
Susskind, L., 2001, e-print hep-th/0101029.

Taylor, W., 1997, Phys. Lett. B **394**, 283.
Taylor, W., 2001, e-print hep-th/0101126.
't Hooft, G., 1974, Nucl. Phys. B **72**, 461.
Tseng, L.-S., 2000, e-print hep-th/0101125.
Vafa, C., 1986, Nucl. Phys. B **273**, 592.
Van Raamsdonk, M., and N. Seiberg, 2000, J. High Energy Phys. **03**, 035.
Varilly, J. C., 1997, e-print physics/9709045.
Varilly, J. C., and J. M. Gracia-Bondia, 1999, Int. J. Mod. Phys. A **14**, 1305.
Verlinde, H., 2001, talk at Davidfest, ITP Santa Barbara, Feb 2001.
Wilczek, F., 1982, Phys. Rev. Lett. **48**, 1144.
Witten, E., 1986, Nucl. Phys. B **268**, 253.
Witten, E., 1988, Commun. Math. Phys. **117**, 353.
Witten, E., 1996, Nucl. Phys. B **460**, 541.
Witten, E., 2000, e-print hep-th/0006071.
Yoneya, T., 1987, in *Wandering in the Fields*, edited by K. Kawarabayashi and A. Ukawa (World Scientific, Singapore), p. 419.
Yoneya, T., 2001, Int. J. Mod. Phys. A **16**, 945.
Zanon, D., 2001a, Phys. Lett. B **502**, 265.
Zanon, D., 2001b, Phys. Lett. B **504**, 101.
Zhou, C.-G., 2000, e-print hep-th/0007255.