

## Two-dimensional Born-Infeld gauge theory: Spectrum, string picture, and large- $N$ phase transition

Alexios P. Polychronakos\*

*Physics Department City College of the CUNY 160 Convent Avenue, New York, New York 10031, USA*  
(Received 13 February 2012; published 22 May 2012)

We analyze  $U(N)$  Born-Infeld gauge theory in two spacetime dimensions. We derive the exact energy spectrum on the circle and show that it reduces to  $N$  relativistic fermions on a dual space. This contrasts to the Yang-Mills case that reduces to nonrelativistic fermions. The theory admits a string theory interpretation, analogous to the one for ordinary Yang-Mills, but with higher-order string interactions. We also demonstrate that the partition function on the sphere exhibits a large- $N$  phase transition in the area and calculate the critical area. The limit in which the dimensionless coupling of the theory goes to zero corresponds to massless fermions, admits a perturbatively exact free string interpretation and exhibits no phase transition.

DOI: [10.1103/PhysRevD.85.105024](https://doi.org/10.1103/PhysRevD.85.105024)

PACS numbers: 11.10.Lm, 11.15.Pg

### I. INTRODUCTION

Two-dimensional gauge theory is special in that local gauge field excitations are absent and only global variables (holonomies) remain as physically relevant degrees of freedom. As such, it is usually amenable to an exact treatment and provides a convenient testing ground for conjectures on the properties of gauge theory. In particular, large- $N$  properties, such as the equivalence of gauge theory and string theory [1–4] and the analyticity of the strong coupling expansion can be directly probed.

The case of ordinary Yang-Mills theory has been analyzed exhaustively. The standard plaquette action has been shown to exhibit a large- $N$  phase transition, leading to a nonanalyticity of the strong coupling expansion below a critical coupling [5,6]. The true fixed point of the plaquette action was shown to be the so-called heat kernel action [7,8], which gives analytic results and allows for the exact calculation of Wilson loop expectation values and of the partition function in terms of infinite sums over representations of the gauge group [9,10].

A particularly attractive feature of two-dimensional Yang-Mills theory is its exact description as a string theory. This was shown both at the level of Wilson loop expectation values, which admit an interpretation as string coverings of the loop area with (taut) string world sheets of various windings [9–11], and the partition function, which can be calculated in terms of wrappings of the spacetime manifold with string world sheets of various windings, including string splitting and joining interactions [12–14]. As usual,  $1/N$  plays the role of the string coupling constant. Nonperturbative effects of order  $\exp(-N)$  also appear, with string winding number being conserved only modulo  $N$ . These results hold without the benefit of supersymmetry and are based on pure group theory of the gauge manifold. Further, the partition function on spacetimes of

spherical topology exhibits a large- $N$  phase transition in the spacetime area, going from a strongly coupled (stringy) phase for large area to a weakly coupled phase for smaller area [15].

Two-dimensional Yang-Mills theory also admits alternative and equivalent descriptions as a collection of free nonrelativistic fermions [16,17], a gauged unitary matrix model [18] and a  $c = 1$  collective field theory [19]. Although the interconnection between these descriptions is known, gauge theory presents a concrete physical realization and allows for a convenient formulation of related string or many-body problems. For instance, two-dimensional Yang-Mills is related to the Sutherland model [20], and generalized interacting Calogero-Sutherland type integrable models of particles with internal degrees of freedom can be obtained and solved in terms of Yang-Mills theory on the cylinder with one or more Wilson loop insertions [21].

Born-Infeld electrodynamics was introduced originally as an attempt to provide a regularization of the short-distance singularity problem of standard electrodynamics while preserving relativistic invariance, at the price of a nonpolynomial action. In more recent contexts, Born-Infeld actions often arise for the gauge fields in string and brane theory (see, e.g., [22,23]). The obvious question is, then, whether such theories in two dimensions are amenable to solution, admit a string interpretation and share the qualitative and quantitative features of standard Yang-Mills theory.

In this paper we analyze two-dimensional Born-Infeld theory and address the above questions. Unlike Yang-Mills theory, which has a unique fixed point, Born-Infeld theory can flow to various inequivalent theories according to the exact renormalization and ordering of its nonpolynomial terms. Under the most “natural” choice, the theory becomes equivalent to a set of *relativistic* fermions, providing a nice generalization of the corresponding Yang-Mills result of nonrelativistic fermions. A string interpretation

\*alexios@sci.cuny.cuny.edu

is still possible but involves higher-order string interactions. On the sphere, a large- $N$  third-order phase transition is also present with the same qualitative features as in Yang-Mills.

The limit in which the dimensionless coupling of the theory goes to zero is particularly interesting: the corresponding fermion picture involves massless particles and admits a perturbatively exact free string interpretation. Further, the would-be phase transition on the sphere disappears, the theory being always in a nominally strongly coupled phase, with a smooth crossover from a true stringy phase to an effectively weakly coupled phase.

## II. CLASSICAL ANALYSIS OF THE SYSTEM

Non-Abelian Born-Infeld actions are, in general, not unique even at the classical level due to ordering ambiguities [24]. Specifically, the determinant over spacetime indices in their Lagrangian involves the product of the components of the field strength tensor, which are matrices and their ordering matters. The unique exception is two-dimensional theory, where the field strength has a unique nonzero component and there is no ambiguity.

We shall consider  $U(N)$  gauge theory on a cylindrical spacetime manifold of spatial period  $L$ . The field strength is

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu + i[A_\mu, A_\nu], \quad \mu, \nu = t, x. \quad (1)$$

The Lagrangian of the theory can be written as

$$\mathcal{L} = -\lambda \text{tr} \sqrt{-\det(\eta_{\mu\nu} + F_{\mu\nu}/T)}, \quad (2)$$

where the trace is over  $U(N)$  matrices while the determinant is over spacetime indices.  $\lambda$  and  $T$  are dimensionful constants, playing the role of brane tension and string tension, respectively. Classically,  $\lambda$  is irrelevant and can be set equal to  $T$ . Quantum mechanically, however, it is a relevant parameter, and the ratio  $\lambda/T$  becomes a dimensionless coupling constant for the theory. In particular, the ‘‘tensionless’’ limit  $\lambda/T \rightarrow 0$  is particularly interesting as will be shown in the sequel.

From now on we shall denote by

$$F = F_{tx} = \partial_t A_x - \partial_x A_t + i[A_t, A_x] \quad (3)$$

the unique nonzero component of the field strength, in terms of which the action becomes

$$S = \int d^2x \mathcal{L} = -\lambda \int d^2x \text{tr} \sqrt{1 - F^2/T^2}. \quad (4)$$

In the limit  $|F| \ll T$  the above action becomes (up to an irrelevant additive constant)

$$S \simeq \int d^2x \frac{\lambda}{2T^2} \text{tr} F^2 = \int d^2x \frac{\lambda}{4T^2} \text{tr} F_{\mu\nu}^2 \equiv \int d^2x \frac{1}{4g^2} \text{tr} F_{\mu\nu}^2, \quad (5)$$

which identifies the Yang-Mills coupling in that limit as

$$g^2 = \frac{T^2}{\lambda}. \quad (6)$$

At the large  $N$  limit the 't Hooft scaling of the Yang-Mills coupling is

$$g^2 = \frac{g_o^2}{N}. \quad (7)$$

It is useful to recast the theory in a first-order formalism. To this end, we define the gauge-covariant momentum

$$B = \frac{\delta \mathcal{L}}{\delta F} = \frac{\lambda F}{T^2 \sqrt{1 - F^2/T^2}} \quad (8)$$

in terms of which the Hamiltonian density is

$$\mathcal{H} = \text{tr}(BF - \mathcal{L}) = \text{tr} \sqrt{\lambda^2 + T^2 B^2}, \quad (9)$$

while the action becomes

$$S = \int d^2x \text{tr}(BF - \sqrt{\lambda^2 + T^2 B^2}). \quad (10)$$

The variation in  $B$  of this action yields its defining equation (8), while the variation of  $A_\mu$  gives the Gauss law constraint and equation of motion, respectively

$$D_x B = 0, \quad D_t B = 0. \quad (11)$$

We can use gauge invariance to put  $A_t = 0$ , provided we impose the Gauss law as a constraint. In terms of the unique gauge field  $A_x \equiv A$  the action is

$$S = \int dt \int_0^L dx \text{tr}(BA\dot{A} - \sqrt{\lambda^2 + T^2 B^2}), \quad (12)$$

while the Gauss law constraint remains

$$\partial B + i[A, B] = 0, \quad (13)$$

where overdot and  $\partial$  stand for time and space differentiation, respectively.

The above theory has no local excitations and its only degree of freedom is the nontrivial holonomy (Wilson loop) around the spatial circle. To reduce the theory to its degrees of freedom, we proceed in close analogy to [16]. We define the spatial open Wilson loop

$$W_{a,b} = P e^{i \int_a^b A dx}, \quad W_{b,a} = W_{a,b}^{-1} \quad (14)$$

and consider the dressed momentum

$$\Pi(x) = W_{0,x} B(x) W_{x,0}. \quad (15)$$

Upon differentiating  $\Pi$  with respect to  $x$  and using the Gauss law we get

$$\partial \Pi = W_{0,x} (\partial B(x) + i[A(x), B(x)]) W_{x,0} = 0. \quad (16)$$

So  $\Pi$  is spatially constant and from  $\Pi(x) = \Pi(0)$  we obtain

$$B(x) = W_{x,0} B(0) W_{0,x} = W_{0,x}^{-1} B(0) W_{0,x}. \quad (17)$$

The kinetic term in the action (12) can be expressed as

$$K = \int dt dx \text{tr}(B\dot{A}) = \int dt \text{tr}[B(0) \int dx W_{0,x} \dot{A} W_{x,0}]. \quad (18)$$

The time derivative of the full Wilson loop  $W_{0,L}$ , on the other hand, is

$$\dot{W}_{0,L} = \int dx W_{0,x} i\dot{A}(x) W_{x,L} = i \int dx W_{0,x} \dot{A}(x) W_{x,0} W_{0,L} \quad (19)$$

and comparing with the kinetic term above we find

$$K = -i \int dt \text{tr}(B(0) \dot{W}_{0,L} W_{0,L}^{-1}). \quad (20)$$

Similarly, the Hamiltonian becomes

$$\begin{aligned} H &= \int dx \text{tr} \sqrt{\lambda^2 + T^2 B(x)^2} \\ &= \int dx \text{tr} \sqrt{\lambda^2 + T^2 W_{0,x}^{-1} B(0)^2 W_{0,x}} = L \text{tr} \sqrt{\lambda^2 + T^2 B(0)^2} \end{aligned} \quad (21)$$

due to the cyclicity of trace. So the full action can be expressed in terms of the space-independent fields  $P \equiv B(0)$  and  $U \equiv W_{0,L}$  as

$$S = \int dt \text{tr}(-iP\dot{U}U^{-1} - L\sqrt{\lambda^2 + T^2 P^2}). \quad (22)$$

In addition, there is one residual Gauss law constraint: the periodicity condition  $B(0) = B(L)$  gives

$$P = UPU^{-1} \quad \text{or} \quad [P, U] = 0. \quad (23)$$

In conclusion, we see that the theory reduces to a matrix model for the unitary matrix  $U$  and the Hermitian matrix  $P$  that plays the role of its canonical right-momentum. This model differs from the traditional unitary matrix model in that its kinetic energy is a nontrivial (nonquadratic) function of the canonical momentum.

The above matrix model can be further reduced to non-interacting particles upon use of the Gauss constraint (23). Classically, upon use of the equations of motion,  $P$  is a function of  $\dot{U}U^{-1}$ , so the constraint implies

$$[P, U] = [U, \dot{U}] = 0. \quad (24)$$

This means that  $P$  and  $U$  can be simultaneously diagonalized with a time-independent unitary transformation, reducing them to their eigenvalues  $p_n$  and  $e^{i\theta_n}$ , respectively. To make this more explicit, we write  $P$  and  $U$  in the basis where  $U$  is diagonal as

$$U = V\Lambda V^{-1}, \quad P = V(p + Q)V^{-1} \quad (25)$$

with  $V$  the diagonalizer of  $U$ ,  $\Lambda$  and  $p$  diagonal and  $Q$  off-diagonal, that is

$$\Lambda = \text{diag}\{e^{i\theta_n}\}, \quad p = \text{diag}\{p_n\}, \quad Q_{nn} = 0 \text{ (no sum in } n). \quad (26)$$

The canonical (time derivative) term in the action in terms of the above variables is

$$-i \text{tr}(P\dot{U}U^{-1}) = \sum_{n=1}^N p_n \dot{\theta}_n + \sum_{n,m} (1 - e^{i(\theta_n - \theta_m)}) Q_{nm} (\dot{V}V^{-1})_{mn}. \quad (27)$$

We see that the eigenvalues  $\theta_n$  are canonically conjugate to the diagonal elements  $p_n$  of  $P$ , while  $Q - \Lambda Q \Lambda^{-1}$  is the right-momentum of the angular part of  $U$  (the diagonalizer  $V$ ). The Gauss law (23) implies  $Q = 0$ . This is a gauge constraint and must therefore be complemented by gauge fixing the coordinates corresponding to  $Q$ , that is  $V$ . Setting  $V = 1$  we are left with a set on  $N$  coordinates on the unit circle and the their canonical momenta. The reduced action becomes

$$S = \int dt \sum_{n=1}^N (p_n \dot{\theta}_n - L\sqrt{\lambda^2 + T^2 p_n^2}) \quad (28)$$

and describes a set of noninteracting particles with a relativistic energy-momentum relation:

$$E = L\sqrt{\lambda^2 + T^2 p^2} \quad (29)$$

with  $LT$  and  $\lambda/LT^2$  playing the role of the speed of light and the particle's mass, respectively.

### III. QUANTIZATION

Quantum mechanically the story is similar, with some additional twists. The wave function  $\Phi(U)$  is a function of the matrix elements of  $U$ . From the canonical structure of the action,  $P$  generates right-multiplications of  $U$  by unitary matrices, while  $-UPU^{-1}$  generates left multiplications. Their sum  $P - UPU^{-1}$  generates the conjugation

$$U \rightarrow V^{-1}UV. \quad (30)$$

The Gauss constraint (23) implies that wave functions are invariant under unitary conjugations of  $U$  and therefore depend only on the eigenvalues of  $U$ . As usual, the change of variables from  $U$  to its eigenvalues  $e^{i\theta_n}$  and the angular variables  $V$  involves the Jacobian of the transformation  $J = |\Delta|^2$ , where  $\Delta$  is the modified Vandermonde factor

$$\Delta(\theta) = \prod_{n < m} \sin \frac{\theta_n - \theta_m}{2}. \quad (31)$$

The Jacobian can be absorbed by incorporating one factor of  $\Delta(\theta)$  in the wave function, rendering the measure in  $\theta_n$  flat. The original wave function  $\Phi(\theta)$  was symmetric under permutation of  $\theta_n$ , so the new wave function

$$\Psi(\theta) = \prod_{n < m} \sin \frac{\theta_n - \theta_m}{2} \Phi(\theta) \quad (32)$$

becomes fermionic. (This is the famous fermionization of the eigenvalues of a matrix model, which holds irrespective of its action.)

The spectrum is evaluated by diagonalizing the Hamiltonian

$$H = L \text{tr} \sqrt{\lambda^2 + T^2 P^2} = L \lambda \sum_{n=0}^{\infty} c_n \left(\frac{T}{\lambda}\right)^{2n} \text{tr} P^{2n} \quad (33)$$

with  $c_n$  the Taylor expansion coefficients of the square root. Since  $P$  generates unitary transformations (left-multiplications of  $U$ ) it satisfies the  $U(N)$  algebra. The quantum commutation relations of its matrix elements read

$$[P_{mn}, P_{kl}] = i(P_{ml} \delta_{kn} - P_{kn} \delta_{ml}) \quad (34)$$

and  $\text{tr} P^n$  is the  $n$ -th Casimir operator for  $U(N)$  ( $\text{tr} P$  being the  $U(1)$  charge). These are diagonalized on states that are irreducible representations (irreps) of the above algebra. Given that states must also be singlets under conjugation of  $U$ , they are the characters of the representations. The fermionic eigenstates are

$$\Psi_R = \Delta(\theta) \chi_R(U) = \Delta(\theta) \text{tr}_R U_R \quad (35)$$

with  $R$  an irrep of  $U(N)$  and  $U_R$  the  $R$ -matrix representation of the element  $U$ . For  $U(N)$ , the Casimirs of order larger than  $N$  are not independent but they are still diagonal on irreps.

The above also leads to the result that the energy states of the theory are simply free states of  $N$  fermions on the unit circle determined by their momenta. Specifically, they are given by the Slater determinant

$$\Psi(\theta) = \det_{kn} (e^{i p_k \theta_n}). \quad (36)$$

The fermion momenta  $p_k$  can be ordered as  $p_{k+1} < p_k$ . Since  $\theta_n$  have period  $2\pi$ , the momenta are quantized to integer steps plus, perhaps, a constant shift. The shift is determined by the properties of Vandermonde factor (31), which is periodic for odd  $N$  and antiperiodic for even  $N$ . So the momenta  $p_k$  are quantized to (half) integers for (even) odd  $N$ . The ground state is in both cases

$$p_{k,o} = \frac{N+1}{2} - k = \left\{ \frac{N-1}{2}, \frac{N-1}{2} - 1, \dots, -\frac{N-1}{2} \right\} \quad (37)$$

representing a Fermi sea symmetric around  $p = 0$  with Fermi momentum  $p_F = (N-1)/2$ . The Slater determinant (36) for the ground state is exactly the Vandermonde factor,  $\Psi_o = \Delta$ , leading to the bosonic ground state  $\Phi_o(U) = 1$ , that is, the singlet.

The relation of the momenta  $p_k$  with the irreps they correspond is standard: the excitation of each momentum from its ground state

$$l_k = p_k - p_{k,o} = p_k + k - \frac{N+1}{2} \quad (38)$$

satisfying  $l_{k+1} \leq l_k$ , represents the length of the  $k$ -th row in the Young tableau of the irrep. The total number of boxes  $\sum l = \sum p$  is the  $U(1)$  charge. Negative lengths correspond to conjugate irreps and can be turned positive by adding a number of columns of length  $N$ , that is, by increasing the  $U(1)$  charge by multiples of  $N$ . Note that for our  $U(N)$  matrix model the  $U(1)$  charge  $Q$  and the  $SU(N)$  irrep are correlated in that the  $Z_N$  charge  $Z$  is common to both and thus  $Z = \exp(iQ)$ .

Since the problem reduces to free particles classically, we expect  $P$  to act essentially as the diagonal momenta  $p_n$  conjugate to  $\theta_n$ , that is  $-i\partial/\partial\theta_n$ . This is, indeed, true for the first two Casimirs.  $\text{tr} P$  reduces to the total momentum of the particles

$$\text{tr} P = \sum_n p_n = -i \sum_n \frac{\partial}{\partial \theta_n}, \quad (39)$$

since the  $U(1)$  charge is just an overall shift of the eigenvalues of  $U$ .  $\text{tr} P^2$  is essentially their quadratic kinetic energy:

$$\text{tr} P^2 = \sum_n p_n^2 = - \sum_n \frac{\partial^2}{\partial \theta_n^2} - C \quad (40)$$

with  $C$  a c-number subtracting the ground-state value of the right hand side operator. In fact,  $\text{tr} P^2$  is the Laplacian on the group  $U(N)$  and it is known to reduce to the above expression when acting on Schur (conjugation-invariant) states.

The situation with higher Casimirs is subtler. In fact,  $\text{tr} P^n$  does *not* reduce to  $\sum p_k^n$  for  $n > 2$ , but involves also polynomials in lower-power sums of  $p_k$  (see, e.g., [25] and references therein). In the classical limit, that is, for  $|p_k| \gg N$ , the two expressions must agree. So we have

$$\text{tr} P^n = \sum_k p_k^n + \text{lower order terms.} \quad (41)$$

This can be viewed as a quantum effect arising from ordering issues in the field theory. Even at the matrix model level, the definition of the quantum Hamiltonian has ordering ambiguities. The first two traces, in terms of the matrix elements of  $P$ ,

$$\text{tr} P = \sum_n P_{nn}, \quad \text{tr} P^2 = \sum_{m,n} P_{mn} P_{nm} \quad (42)$$

are uniquely defined. At the cubic level, however, we already see that there are two possible orderings:

$$\text{tr} P^3 = \sum_{m,n,l} P_{mn} P_{nl} P_{lm} \quad \text{or} \quad \sum_{m,n,l} P_{mn} P_{lm} P_{nl}. \quad (43)$$

The two are classically the same but quantum mechanically inequivalent, differing by lower-order terms. Although the first leads to the conventionally defined

Casimir, there is no reason not to consider the second. In fact, the sum of the two leads to an expression where the first subleading correction to  $\sum p_k^3$  cancels.

We see that the exact definition of the Hamiltonian depends on the ordering of its terms. This is not surprising, since the original field action contained infinitely high powers of time derivatives and such terms require a precise ordering.

We can see this ambiguity at the field theory level before we reduce to the matrix model by a method analogous to the heat kernel in standard two-dimensional Yang Mills: we tessellate spacetime into small plaquettes of arbitrary shape and size and perform the Euclidean path integral over the gauge fields inside each plaquette. From gauge invariance, the result for each plaquette will only depend on the holonomy (Wilson loop)  $W$  of the gauge field around the plaquette. The fixed-point expression must be of the form

$$Z = \sum_R d_R e^{-AE_R} \chi_R(W) \quad (44)$$

with  $A$  the area of the plaquette,  $d_R$  the dimension of irrep  $R$  and  $E_R$  a number depending only on  $R$ . We can then consider two adjacent plaquettes with Wilson element  $U$  on their common boundary and holonomies  $W_1 U$  and  $U^{-1} W_2$  ( $W_1$  and  $W_2$  being the noncommon parts) and integrate their path integrals  $Z_1(U W_1) Z_2(U^{-1} W_2)$  over the common part  $U$  to calculate the path integral for the combined plaquette. Because of the orthogonality property of the irreps

$$\int [dU] \chi_R(WU) \chi_{R'}(U^{-1}V) = \frac{1}{d_R} \delta_{RR'} \chi_R(WV), \quad (45)$$

we see that the result will be of the form

$$Z_{12} = \sum_R d_R e^{-(A_1+A_2)E_R} \chi_R(W_1 W_2), \quad (46)$$

involving their total area and the holonomy  $W_1 W_2$  around the total plaquette, verifying the consistency of the expression (44) for the fixed-point partition function.

The only extra requirement is that the above expression be the quantization of a specific classical action. For this, we need to ensure that for  $A \rightarrow 0$  the expression in the exponent of (44) goes over to the classical action for the fields. For small  $A$ , the sum in (44) is dominated by large irreps, that is, by large values of  $p_k$  in the fermionic description. For such irreps, we must have

$$E_R = \sum_{k=1}^N \sqrt{\lambda^2 + T^2 p_k^2} \quad \text{for } |p_k| \gg N. \quad (47)$$

Any  $E_R$  with the above property provides a consistent gauge invariant quantization of the same Born-Infeld classical field theory. Choosing the expression (33) in terms of the standard Casimirs in just one of many possibilities.

We can, therefore, adopt the simplest definition in which the expression (47) holds for *all* irreps and define the Hamiltonian as

$$H = \sum_{n=1}^N E(p_n) = L \sum_{n=1}^N \sqrt{\lambda^2 + T^2 p_n^2}. \quad (48)$$

The above assigns a positive value to the energy of the ground state, which is inconsequential for expectation value calculations and can easily be removed. In some sense, the above is the most natural definition, since the dynamics of the gauge field reduce to those of a set of uncoupled relativistic particles, admitting the interpretation of points on a relativistic brane in a dual description.

Our final result is that Born-Infeld gauge theory on the cylinder reduces to a set of relativistic fermions on a dual circle. If we incorporate a factor of  $LT$  in the momentum, the energy expression for each particle becomes

$$\tilde{E}(\tilde{p}) = \sqrt{\lambda^2 L^2 + \tilde{p}_n^2}, \quad \tilde{p} = LT p, \quad (49)$$

representing a particle of mass  $\lambda L$  on a circle of radius  $R = (TL)^{-1}$ . This is to be contrasted to regular two-dimensional Yang Mills on the cylinder, which is equivalent to a set of nonrelativistic fermions and in which there is no unique identification of particle mass and radius of the dual circle, the two appearing as one overall coefficient.

#### IV. THE LARGE- $N$ LIMIT AND STRING DESCRIPTION

There are various ways to take the large- $N$  limit in the above theory. The one relevant to the string interpretation is what we can call the conformal field theory limit. In this limit, the low-lying energy excitations of the theory become equally spaced and approach those of a  $c = 1$  conformal field theory, that is, a relativistic fermion.

The excitations of the Born-Infeld theory consist of fermion excitations above the Fermi level. There are two Fermi levels, at  $p = \pm(N-1)/2$ , leading to two left- and a right-moving noninteracting sectors. (Depletion of the Fermi sea corresponds to nonperturbative in  $1/N$  interaction effects.) Concentrating on excitations near the right-moving Fermi level  $p_F = (N-1)/2$ , a fermion excited from  $p = (N-1)/2 - m$  to  $(N-1)/2 + n$ , with  $m, n$  positive and of order 1, has excitation energy

$$\Delta E = E(p_F - m) - E(p_F + n) \simeq \partial_p E(p_F)(n+m) \quad (\text{for } N \gg 1). \quad (50)$$

So the scale of the energy gap is set by the Fermi velocity (velocity of sound on dual space)

$$v_F = f(p_F), \quad f(p) \equiv \frac{\partial E(p)}{\partial p} = \frac{LT^2 p}{\sqrt{\lambda^2 + T^2 p^2}}. \quad (51)$$

For  $p_F = (N-1)/2$  the above will be of order  $N^0$  if  $T$  does not scale but  $\lambda$  scales with  $N$ , that is,

$$\lambda = \lambda_o N. \quad (52)$$

From (6) we see that in the Yang-Mills limit the above scaling is consistent with the standard 't Hooft scaling, with 't Hooft coupling

$$g_o^2 = \frac{T^2}{\lambda_o}. \quad (53)$$

From now on we will always assume the expression (52) for  $\lambda$  and will write  $\lambda$  instead of  $\lambda_o$  to alleviate notation.

The string picture of the gauge theory on the cylinder remains largely as in standard Yang-Mills: the leading- $N$  terms in the excitation energy represent a theory of free strings wrapping around  $S^1$  with string tension

$$T_{st} = \frac{v_F}{L} = \frac{T^2}{\sqrt{4\lambda^2 + T^2}}. \quad (54)$$

The term of order  $n^2$  in the expansion of  $E(p_F + n) - E(p_F)$  is a  $1/N$  correction that introduces a cubic string interaction representing string splitting or joining. The string coupling constant is

$$g_{st} = \frac{1}{2} \frac{\partial^2 E(p_F)}{\partial p^2} = \frac{\lambda^2 T^2}{N(4\lambda^2 + T^2)^{(3/2)}}. \quad (55)$$

The difference from Yang-Mills theory comes from the existence of higher orders in the expansion of  $E(p)$  in  $p$ , which are absent in the Yang-Mills case. The cubic term, of order  $N^{-2}$ , introduces a quartic string interaction that represents a localized double string interaction, that is, two pairs of strings touching and reconnecting at the same point of space and time. Such interactions are not so natural from the world sheet point of view. Higher terms lead to higher yet order string interactions. Overall, we have a nonpolynomial string field theory.

## V. LARGE $N$ PHASE TRANSITION ON THE SPHERE

The partition function on the circle is given by the path integral on a Euclidean torus  $(L, \beta)$ . The result is

$$Z = \sum_{\{p_k\}} e^{-A \sum_n \sqrt{\lambda^2 N^2 + T^2 p_n^2}} \quad (56)$$

with  $A = L\beta$  the area of the world sheet. The summation is over all combinations of fermionic momenta  $p_1 > p_2 > \dots > p_N$  on a (half) integer lattice for (even) odd  $N$ .

For spacetimes of genus  $g$ , the partition function is similar but with an extra measure factor  $d_R^{2-2g}$  in each term. This factor can be understood as a remnant from the plaquette formula (44) as we coalesce the plaquettes on spacetimes of different topologies. For spherical topology, in particular, it can be understood as arising from the insertion of a singular wave function at the north and south pole of the sphere, representing the constraint  $W = 1$  at these points in a canonical formulation [17]. In terms of

fermion momenta the extra factor  $d_R^2$  on the sphere is expressed as a Vandermonde-like product

$$d_R^2 = \frac{\prod_{n < m} (p_n - p_m)^2}{\prod_{n < m} (n - m)^2}. \quad (57)$$

The denominator is the product for the ground state momenta, ensuring  $d_R = 1$  for the singlet, and contributes an overall normalization factor that will be omitted. The resulting partition function on the sphere is

$$Z_{sph} = \sum_{\{p_k\}} \prod_{n < m} (p_n - p_m)^2 e^{-A \sum_n \sqrt{\lambda^2 N^2 + T^2 p_n^2}} = \sum_{\{p_k\}} e^{-S_{\text{eff}}}. \quad (58)$$

The effective action contains the exponentiated measure and reads

$$S_{\text{eff}} = A \sum_n \sqrt{\lambda^2 N^2 + T^2 p_n^2} - 2 \sum_{n < m} \ln |p_n - p_m|. \quad (59)$$

The measure introduces a repulsive logarithmic two-body potential in the momenta.

For large  $N$  the effective action is of order  $N^2$ , since  $p_n$  are of order  $N$ , and the partition will be dominated by the classical minimal effective energy configuration in a saddle-point approximation. By differentiating  $S_{\text{eff}}$  with respect to  $p_n$  we obtain the minimal energy condition

$$A \frac{T^2 p_n}{\sqrt{\lambda^2 N^2 + T^2 p_n^2}} - \sum_{m(\neq n)} \frac{2}{p_n - p_m} = 0. \quad (60)$$

For large  $N$  we can approximate the distribution of momenta with a continuous density  $\tilde{\rho}(p)$ . The minimum energy condition becomes

$$\frac{A}{2} \frac{T^2 p}{\sqrt{\lambda^2 N^2 + T^2 p^2}} = \int \frac{\tilde{\rho}(p')}{p - p'} dp' \quad \text{for } \tilde{\rho}(p) \neq 0 \quad (61)$$

with  $\tilde{\rho}$  satisfying

$$\tilde{\rho} \geq 0, \quad \int \tilde{\rho}(p) dp = N. \quad (62)$$

We can define a rescaled variable  $x = p/N$  and a corresponding density

$$\rho(x) = \tilde{\rho}(Nx), \quad \int \rho(x) dx = 1. \quad (63)$$

In terms of the new variable and density the condition (61) becomes

$$\frac{A}{2} \frac{T^2 x}{\sqrt{\lambda^2 + T^2 x^2}} = \int \frac{\rho(y)}{x - y} dy \quad \text{for } \rho(x) \neq 0, \quad (64)$$

eliminating all reference to  $N$  in the large- $N$  limit.

The solution of the above equation is well known. The function

$$u(z) = \int \frac{\rho(y)}{z-y} dy \quad (65)$$

is analytic on the upper half plane and behaves as  $\rho(z) \sim 1/z$  at  $z \rightarrow \infty$ . Near the real axis it becomes

$$u(x+i0) = -i\pi\rho(x) + \int \frac{\rho(y)}{x-y} dy. \quad (66)$$

For a symmetric distribution  $\rho(-x) = \rho(x)$  that vanishes outside of an interval  $(-a, a)$  the solution for  $u(z)$  is

$$u(z) = \frac{1}{2\pi i} \sqrt{a^2 - z^2} \oint \frac{AT^2 s}{2(s-z)\sqrt{\lambda^2 + T^2 s^2} \sqrt{a^2 - s^2}} ds. \quad (67)$$

In the above the square roots are defined with a cut along  $(-a, a)$  and the integration contour is clockwise around the cut but not including the pole at  $s = z$ . It is easy to see from the above formula that  $u(x+i0)$  is real for  $|x| > a$ , while its imaginary part is the left-hand side of (64) for  $|x| < a$ . Therefore it satisfies (64), provided it also has the proper asymptotic behavior for large  $z$ . This will be ensured if

$$\int_{-a}^a \frac{AT^2 s^2}{\sqrt{\lambda^2 + T^2 s^2} \sqrt{a^2 - s^2}} ds = 2\pi. \quad (68)$$

The density  $\rho(x)$  is then recovered as

$$\rho(x) = \frac{1}{\pi^2} \sqrt{a^2 - x^2} \int_{-a}^a \frac{AT^2 s}{2(s-x)\sqrt{\lambda^2 + T^2 s^2} \sqrt{a^2 - s^2}} ds. \quad (69)$$

For the case of standard Yang-Mills, where  $f(p) = Ag^2 p$ , the corresponding integral (67) can be easily calculated by blowing up the contour to infinity, and leads to the Wigner semicircle distribution. In our case this is not so easy, since the contour encounters the cut of the square root in the denominator on the imaginary axis  $(i\lambda/T, \infty)$ . The integrals can be expressed in terms of elliptic functions and implicitly define  $a$  and  $\rho(x)$ .

The above solution is valid as long as  $A$  is not too big, in the so-called weak coupling phase. As  $A$  increases,  $a$  decreases and the momentum distribution becomes denser. The momenta, however, are fermionic and lie on a lattice of spacing 1, so their density cannot exceed 1. Correspondingly,  $\rho(x) = \bar{\rho}(Nx)$  cannot exceed 1. The maximum of  $\rho(x)$  occurs at  $x = 0$ . Therefore, when  $\rho(0)$  reaches the value 1, fermionic momenta will start condensing and the above solution will not be valid any more, signaling a phase transition. To find the critical area we put  $\rho(0) = 1$ :

$$\rho(0) = \frac{aA_{cr}T^2}{2\pi^2} \int_{-a}^a \frac{ds}{\sqrt{\lambda^2 + T^2 s^2} \sqrt{a^2 - s^2}} = 1. \quad (70)$$

Combining (68) and (70) we obtain an equation for the width  $a$  at critical area:

$$\int_{-a}^a \frac{a - \pi s^2}{\sqrt{\lambda^2 + T^2 s^2} \sqrt{a^2 - s^2}} ds = 0, \quad (71)$$

which fixes  $a$  in terms of  $T/\lambda$  and, upon inserting in (70), it determines  $A_{cr}$ .

For  $A > A_{cr}$  the solution develops a flat central part where the fermion momenta condense, and an outer tail part:

$$\rho(x) = 1|x| < b = \bar{\rho}(x)b < |x| < a = 0|x| > a. \quad (72)$$

The contribution from the flat central part can be taken explicitly into account in Eq. (64) producing an extra logarithmic potential [15]. The remaining density  $\bar{\rho}(x)$  vanishing outside  $(-a, a)$  and inside  $(-b, b)$  can be found in a way similar to  $\rho(x)$ . The solution for its analytic extension  $\bar{u}(z)$  becomes a two-cut integral with an additional logarithm cut between  $(-b, b)$ :

$$\begin{aligned} \bar{u}(z) = & \frac{1}{2\pi i} \sqrt{(a^2 - z^2)(b^2 - z^2)} \\ & \times \oint \frac{\frac{AT^2 s}{2\sqrt{\lambda^2 + T^2 s^2}} + \ln \frac{s-b}{s+b}}{(s-z)\sqrt{(a^2 - s^2)(b^2 - s^2)}} ds \end{aligned} \quad (73)$$

with the contour encircling the square root cuts between  $(-a, -b)$  and  $(b, a)$  but not the log cut and the pole at  $z$ . The second part of the above integral, involving the logarithm, can be explicitly evaluated by deforming the contour around the log cut (we encounter the pole at  $s = z$  and no other cuts). The result is

$$\begin{aligned} & \frac{1}{2\pi i} \sqrt{(a^2 - z^2)(b^2 - z^2)} \oint \frac{\ln \frac{s-b}{s+b}}{(s-z)\sqrt{(a^2 - s^2)(b^2 - s^2)}} ds \\ & = \ln \frac{z-b}{z+b} - \sqrt{(a^2 - z^2)(b^2 - z^2)} \\ & \times \int_{-b}^b \frac{ds}{(s-z)\sqrt{(a^2 - s^2)(b^2 - s^2)}}. \end{aligned} \quad (74)$$

For  $z = x + i0$  the imaginary part of the logarithm above vanishes for  $|x| > b$  and equals  $i\pi$  for  $|x| < b$ . It thus contributes  $-1$  to the density  $\bar{\rho}(x)$  in the interval  $(-b, b)$  and zero outside. Therefore, removing it restores the density to its full value  $\rho(x)$  (equal to 1 between  $-b$  and  $b$ ). The density is reproduced by the above integral plus the first part (nonlogarithm) of the integral in (73). Taking also into account the even nature of  $\rho(x)$  we obtain

$$\begin{aligned} \rho(x) = & \frac{|x|}{\pi^2} \sqrt{(a^2 - x^2)(x^2 - b^2)} \left[ \int_b^a \frac{AT^2 s}{\sqrt{\lambda^2 + T^2 s^2}} + \int_{-b}^b \pi \right] \\ & \times \frac{ds}{(s^2 - x^2)\sqrt{(a^2 - s^2)|b^2 - s^2}}. \end{aligned} \quad (75)$$

(Note that in the case of standard Yang-Mills the first integral above vanishes, as can be shown by contour integration, and only the second term, arising from the logarithm integral, survives.) We must also ensure the proper asymptotic behavior of  $\bar{u}(z)$  at infinity, that is,

$$\bar{u}(z) = 0 \cdot z + 0 \cdot 1 + \frac{1 - 2b}{z} + O(z^{-2}). \quad (76)$$

The vanishing of the constant term above is an identity, but the terms of order  $z$  and  $z^{-1}$  give two conditions that fix, in principle,  $a$  and  $b$  in terms of  $A$ .

The results for standard Yang-Mills theory can be recovered as the limit  $\lambda \gg T$  of the above results. The 't Hooft coupling  $g$  in that limit is

$$g^2 = \frac{T^2}{\lambda} \ll T \ll \lambda. \quad (77)$$

This is essentially the nonrelativistic limit of the fermionic system of momenta, in which the Fermi momentum is much smaller than the rest mass of the particles:  $p_F \ll m$  or  $TN \ll \lambda N$ . In that limit the factors  $\sqrt{\lambda^2 + T^2 s^2}$  in all the previous integrands become simply  $\lambda$  and the integrals simplify. The condition (68) gives  $gA_{cr}a^2 = 4$ , while the saturation condition (70) gives  $gA_{cr}a = 2\pi$ . These two conditions give  $a = 2/\pi$  and

$$gA_{cr} = \pi^2, \quad (78)$$

which is the standard Yang-Mills result [15]. As mentioned before, the weak coupling eigenvalue distribution ( $A < A_{cr}$ ) is a standard Wigner semicircle.

## VI. THE LIMIT $\lambda/T \rightarrow 0$

The case  $\lambda \ll T$  is of particular interest: in terms of fermions, the dispersion relation becomes linear and the fermions become massless. The Fermi velocity is constant and the corresponding string theory contains no higher-order terms and becomes free. In a sense, this is the “stringiest” version of gauge theory and does not even require a large- $N$  limit to manifest a perturbative free string behavior. Finite- $N$  effects arise only as nonperturbative corrections.

It is interesting that in this case the eigenvalue distribution can be calculated exactly. Putting  $\lambda = 0$  and substituting  $\sqrt{\lambda^2 + T^2 x^2}$  by  $T|x|$  we obtain integrals with a cut along the entire imaginary axis that can be explicitly evaluated. The normalization condition (68) gives

$$a = \frac{\pi}{AT}, \quad (79)$$

while the expression for  $\rho$  gives

$$\rho(x) = \frac{AT}{\pi^2} \ln \frac{\pi + \sqrt{\pi^2 - A^2 T^2 x^2}}{AT|x|}. \quad (80)$$

We see that  $\rho(0) = \infty$  for all  $A$ , and so the model is always in the strong coupling (stringy) phase. The above solution, therefore, is not really valid but we must instead calculate the two-cut solution with a flat central region for  $\rho(x)$ . It is still a good approximation to the exact solution for small enough  $A$ , that is,  $AT \ll 1$ . In that case the solution for the density is the above, for  $|x| > b$ , and 1 for  $|x| < b$ , with  $b$

the value for which the above function reaches the value 1, that is,

$$\frac{1}{b} = \frac{AT}{\pi} \operatorname{ch} \frac{\pi^2}{AT}. \quad (81)$$

The above is clearly nonperturbative in  $A$ . For small areas, the solution for  $\rho(x)$  differs very little from the would-be weak coupling solution (80). For  $AT \gg 1$ , on the other hand, the solution approaches a true “stringy” state of a fully filled Fermi sea with few momenta spreading above the Fermi levels and becomes identical to the corresponding Yang-Mills solution upon identifying the Fermi velocities, or

$$2T = g^2 \quad (82)$$

with  $g$  the 't Hooft Yang-Mills coupling [ $g_o$  in (7)].

For small nonzero values of  $\lambda$  ( $\lambda \ll T$ ) we can estimate the critical area  $A_{cr}$ . The integral in (68) is of order  $\lambda^0$  while the integral (70) has a logarithmic divergence in  $\lambda$ . To leading order we obtain

$$\pi a = \ln \frac{2Ta}{\lambda}, \quad (83)$$

which has as leading log solution

$$a = \frac{1}{\pi} \ln \frac{2T}{\pi\lambda}. \quad (84)$$

Altogether this gives the critical area

$$A_{cr} = \frac{\pi^2}{T \ln \frac{2T}{\pi\lambda}}, \quad (85)$$

which is, again, nonperturbative in  $\lambda$ . The same conclusion can be reached by putting  $b = \lambda/T$  in formula (81), which is the value of  $s$  for which the two terms in the expression  $\sqrt{\lambda^2 + T^2 s^2}$  become comparable and thus  $\lambda$  starts regulating the behavior at  $s = 0$ . For  $b$  less than that we do not expect  $\rho(0)$  to reach 1, so at this value of  $b$  in (81) we expect a phase transition.

Finally, we can calculate the free energy  $\mathcal{F}(A)$  in the weakly coupled case of small area. It will be given by the value of the effective action for the saddle point distribution for  $p_n$ . Since  $\partial S_{\text{eff}}/\partial p_n = 0$  at the classical saddle point, we have

$$\frac{\partial \mathcal{F}}{\partial A} = \frac{\partial}{\partial A} S_{\text{eff}} = \sum_n T |p_n| = T \int |p| \bar{\rho}(p) dp = N^2 T \int |x| \rho(x) dx. \quad (86)$$

The density  $\rho(x)$  is given by (80) up to nonperturbative corrections in  $A$ . An explicit calculation gives

$$\frac{\partial \mathcal{F}}{\partial A} = \frac{1}{A} \rightarrow \mathcal{F} = \ln A \quad (87)$$

up to a constant. We should also subtract the ground state energy of the fermions, such that the vacuum have zero



energy. For the ground state momenta (37) and the large  $N$  limit we have

$$\beta E_o = AT \sum_n |p_{n,o}| = \frac{1}{4} ATN^2, \quad (88)$$

so overall the free energy is

$$\mathcal{F} = N^2 \left( \ln A - \frac{1}{4} TA \right) \quad (89)$$

up to an  $A$ -independent constant. The  $\ln A$  part is essentially fixed by the scaling properties of the fermion particle energy. An expression  $E(p) \sim p^\alpha$  would contribute a term  $\alpha^{-1} \ln A$ . In the case of Yang-Mills we have  $\alpha = 2$ , while in our case  $\alpha = 1$ .

## VII. CONCLUSIONS

The properties of Born-Infeld two-dimensional gauge theory in general parallel those of standard Yang-Mills, with some interesting twists. The disparity between the two becomes apparent for large values of the gauge field, as expected. On the cylinder, large energy excitations tend to preserve their linear dispersion relation over a wider range, although the deviations are nonpolynomial. On the sphere, a phase transition also occurs, but the critical area decreases as the Born-Infeld theory becomes more relativistic and vanishes in the tensionless limit  $\lambda \rightarrow 0$ . In that limit, the theory

on the cylinder becomes a free string theory, receiving only nonperturbative corrections in the large- $N$  limit.

There are many issues that remain to be investigated. The expansion of the free energy as a function of the area and its nonanalyticity near the transition point on the sphere could be examined with a view to clarify the stringy nature of the strong coupling phase. The question of  $U(1)$  sectors is also an open one: a global momentum shift of the fermions is in principle a low energy excitation but, in the large- $N$  limit, it becomes nonperturbative. The evaluation of the partition function can be performed around an isolated  $U(1)$  (total momentum) sector, similar to the Yang-Mills case [17].

Finally, the calculation and behavior of Wilson loop expectation values is a very interesting issue. In the case of the sphere they would probe the nature of the phase transition, and of the validity of the string description. On the cylinder, insertion of (one or several) timelike Wilson loops would promote the fermion system into an interacting one with internal degrees of freedom, which would constitute integrable and solvable many-body systems [26,27]. The obvious conjecture would be that these systems are generalizations of the Ruijsenaars-Schneider system of “relativistic” fermions including internal degrees of freedom, but the exact form of the Hamiltonian has to be worked out.

- 
- [1] G. 't Hooft, *Nucl. Phys.* **B75**, 461 (1974).
  - [2] K. G. Wilson, *Phys. Rev. D* **10**, 2445 (1974).
  - [3] B. Sakita, *Phys. Rev. D* **21**, 1067 (1980).
  - [4] W. A. Bardeen, I. Bars, A. J. Hanson, and R. D. Peccei, *Phys. Rev. D* **13**, 2364 (1976).
  - [5] D. J. Gross and E. Witten, *Phys. Rev. D* **21**, 446 (1980).
  - [6] S. R. Wadia, *Phys. Lett. B* **93**, 403 (1980).
  - [7] A. A. Migdal, *Zh. Eksp. Teor. Fiz.* **69**, 810 (1975) [*Sov. Phys. JETP* **42**, 413 (1975)].
  - [8] B. E. Rusakov, *Mod. Phys. Lett. A* **5**, 693 (1990).
  - [9] V. A. Kazakov and I. K. Kostov, *Nucl. Phys.* **B176**, 199 (1980); **B105**, 453 (1981).
  - [10] V. A. Kazakov, *Nucl. Phys.* **B179**, 283 (1981).
  - [11] I. K. Kostov, *Phys. Lett.* **138B**, 191 (1984).
  - [12] D. J. Gross, *Nucl. Phys.* **B400**, 161 (1993).
  - [13] J. A. Minahan, *Phys. Rev. D* **47**, 3430 (1993).
  - [14] D. J. Gross and W. Taylor, *Nucl. Phys.* **B400**, 181 (1993).
  - [15] M. R. Douglas and V. A. Kazakov, *Phys. Lett. B* **319**, 219 (1993).
  - [16] J. A. Minahan and A. P. Polychronakos, *Phys. Lett. B* **312**, 155 (1993).
  - [17] J. A. Minahan and A. P. Polychronakos, *Nucl. Phys.* **B422**, 172 (1994).
  - [18] A. P. Polychronakos, *Phys. Lett. B* **266**, 29 (1991).
  - [19] S. R. Das and A. Jevicki, *Mod. Phys. Lett. A* **5**, 1639 (1990).
  - [20] A. Gorsky and N. Nekrasov, *Nucl. Phys.* **B414**, 213 (1994).
  - [21] J. A. Minahan and A. P. Polychronakos, *Phys. Lett. B* **326**, 288 (1994).
  - [22] E. S. Fradkin and A. A. Tseytlin, *Phys. Lett.* **163B**, 123 (1985).
  - [23] R. G. Leigh, *Mod. Phys. Lett. A* **4**, 2767 (1989).
  - [24] A. A. Tseytlin, *Nucl. Phys.* **B501**, 41 (1997).
  - [25] M. R. Douglas, *arXiv:hep-th/9303159*; *arXiv:hep-th/9311130*.
  - [26] J. Blom and E. Langmann, *Phys. Lett. B* **429**, 336 (1998).
  - [27] A. P. Polychronakos, *Nucl. Phys.* **B546**, 495 (1999); **B543**, 485 (1999).