SLAC–PUB–PUPT-1547
June 1995
hep-th/9507060

Topological Rigid String Theory and Two Dimensional QCD

Petr Hořava *

Joseph Henry Laboratories
Jadwin Hall
Princeton University
Princeton, NJ 08544, USA

ABSTRACT

We present a string theory that reproduces the large-N expansion of two dimensional Yang-Mills gauge theory on arbitrary surfaces. First, a new class of topological sigma models is introduced, with path integrals localized to the moduli space of harmonic maps. The Lagrangian of these harmonic topological sigma models is of fourth order in worldsheet derivatives. Then we gauge worldsheet diffeomorphisms by introducing the induced worldsheet metric. This leads to a topological string theory, whose Lagrangian coincides in the bose sector with the rigid string Lagrangian discussed some time ago by Polyakov and others as a candidate for QCD string theory. The path integral of this topological rigid string theory is localized to the moduli spaces of minimal-area maps, and calculates their Euler numbers. The dependence of the large-N QCD partition functions on the target area emerges from measuring the volume of the moduli spaces, and can be reproduced by adding a Nambu-Goto term (improved by fermionic terms) to the Lagrangian of the topological rigid string.

⋆ E-mail address: horava@puhep1.princeton.edu

1. Introduction

Quantum string theory emerged from dual models [1,2] as an attempt to formulate a theory of strong interactions directly in terms of degrees of freedom relevant at low energies. Today, as in its early days, the string picture of hadrons and their interactions is based on experimental data (especially, the successes of Regge phenomenology).

Ever since the discovery of quantum chromodynamics, intense efforts have been made to understand the theory better in the infrared, in particular to relate the successes of dual models with the intuitive picture of the color flux tube as the confining stringy force between quarks. In QCD, this string picture of quark confinement has been further reinforced by results of the strong coupling expansion [3] (especially on the lattice, leading to the area law for Wilson loops), and the large-N expansion with N the number of colors [4] (with its suggestive classification of diagrams in terms of surface topologies). For recent reviews, see [5].

Despite these facts, the hypothetical QCD string theory has turned out to be surprisingly elusive. The search for the string theory of quark confinement has produced very attractive and successful spin-offs (such as the string theory of quantum gravity), but a theory that captures adequately the relevant degrees of freedom in the confining regime of gauge theories still remains to be identified.

Short of a theory of confinement that would be equivalent to standard QCD with bosonic gluons and three colors, one might try to first understand the problem in a simplified setting. Two simplifications come to mind: (i) spacetime supersymmetry, and (ii) the large-N limit. The first option has been studied extensively in the past year or so, with some spectacular results [6]. In this paper we follow the latter option. Since the results of this paper suggest the existence of an $\mathcal{N}=2$ supersymmetry on the worldsheet of the QCD string (at least in two target dimensions), one can even speculate that (i) and (ii) might be related, the link between them being the presence of supersymmetry.

In two spacetime dimensions, the large-N theory can be analyzed in detail [7], and does indeed lead to (perturbative) color confinement, as well as an infinite number of resonances when quarks are present in the microscopic theory. These facts, together with the exact solvability of the Yang-Mills theory on arbitrary two-dimensional surfaces [8-11], make the two dimensional theory an excellent starting point for the study of QCD string theory. Since quarks play a relatively minor role in the string formation, we can as well study the pure Yang-Mills theory on arbitrary surfaces and try to reformulate it as a theory of strings. This is exactly the strategy initiated by Gross and Taylor in [12,13] (for an earlier work, see also [14]). While the authors of [13] were able to interpret the large-N expansion in terms of counting specific maps from auxiliary two-dimensional surfaces to the spacetime manifold, they do not present a string Lagrangian that would reproduce the same results by path integral over worldsheet geometries. The main goal of this paper is to present such a path integral framework.

One more remark is in order. The idea behind QCD string theory is to offer an alternative description of strong interactions, dual to the standard description in terms of gluons and quarks. Such a duality can be valid either effectively in the infrared, or exactly at all length scales. The latter scenario is certainly more appealing, and if true, would be much harder to establish. In particular, it would be extremely interesting to see whether string theory can deal with the regimes where the naive string picture apparently breaks down (non-zero thickness of the color flux tube, parton behavior of amplitudes in the UV).

An "exact" duality between strings and Yang-Mills theory could also be interesting for purely mathematical reasons. The results of this paper provide a very simple example of such a mathematical interest. We present an exact relation between the large-N Yang-Mills theory on arbitrary two dimensional surface M_G of genus G, and topological rigid string theory on M_G. Rephrased in a mathematical language, our results establish a relation between cohomological properties of moduli spaces of flat SU(N) connections on M_G as $N \to \infty$, and cohomological properties of moduli spaces of minimal-area maps from auxiliary surfaces Σ_g of genus g to M_G. Similar relations could be expected in higher dimensions if an exact duality between strings and Yang-Mills theory holds.

This paper is organized as follows. In the rest of §1, results of the large-N expansion of SU(N) Yang-Mills theory [13] are reviewed, with a particular emphasis on the stringent constraints they pose on the 2D QCD string theory. We also recall some general aspects of the 2D Yang-Mills theory, to be able to draw some analogies with the worldsheet theory later. In §2 we present a new class of topological sigma models with path integrals localized to moduli spaces of harmonic maps. While §2 is devoted to general aspects of these "harmonic topological sigma models," §3 discusses their partition functions in the case of two-dimensional targets. In §4 we introduce a topological string theory, with path integrals localized to moduli spaces of minimal area maps, and show that this theory is a topological version of the rigid string theory. The partition functions of this topological string theory calculate Euler numbers of the moduli spaces. In §5 the discussion is specialized to two target dimensions, and the Euler numbers of all moduli spaces are computed explicitly. These results are then related in §6 to the large-N expansion of 2D QCD in the zero-tension limit. The area dependence of the large-N QCD partition functions at non-zero tension is shown to come from measuring the volume of the moduli spaces, and a class of deformations of the topological rigid string by cohomology classes of the moduli spaces is discussed which reproduces this area dependence. Since the simplest deforming term is essentially the Nambu-Goto Lagrangian, our string theory can be viewed as an alternative quantization of the Nambu-Goto Lagrangian. We conclude with some remarks in §7.

Results presented here have been reported some time ago [15]. The purpose of this paper is to give their more detailed and systematic exposition, and provide a reference point for our further investigations. Even though the natural extension of the model to higher target dimensions is no longer a topological string theory [16], two dimensions represent the particular case in which topological methods are sufficient for the solution of the model. For this reason, we limit ourselves in this paper to the topological aspects of the theory, and will discuss the extension to higher dimensions elsewhere [16].

1.1 The Large-N Expansion in Two-Dimensional QCD

The path integral of the Yang-Mills gauge theory on an arbitrary two dimensional compact surface M with fixed metric $g_{\mu\nu}$, as defined by the Yang-Mills Lagrangian

$$\mathcal{L}_{\rm YM} = -\frac{1}{4e_0^2}\int_M \mathrm{d}^2x\,\sqrt{g}\,\mathrm{Tr}\left(F^{\mu\nu}F_{\mu\nu}\right), \tag{1.1}$$

can be solved exactly [8-11], for any compact gauge group $\mathcal{G}$. The exact partition function is given by

$$\mathcal{Z}(G,\mathcal{G},e_0^2A) = \sum_R \left(\dim R\right)^{2-2G}\,\mathrm{e}^{-e_0^2A\,C_2(R)}, \tag{1.2}$$

where the sum goes over all irreducible representations R of the gauge group, G is the genus of the spacetime manifold M, e_0 is the gauge coupling constant, A denotes the total area of M in the fixed Riemann metric $g_{\mu\nu}$, and $C_2(R)$ is the quadratic Casimir of R. Several methods have been used to obtain this result, among them an exact lattice formulation [8] and the Duistermaat-Heckman localization in an associated topological Yang-Mills theory [11].

In [13], the exact formula was further analyzed for $\mathcal{G}=\mathrm{SU}(N)$; in particular, the partition functions were expanded in the powers of $1/N$, and the resulting numerical coefficients were interpreted in terms of rules for counting specific classes of maps from auxiliary two-dimensional surfaces (string worldsheets) to the spacetime manifold M. Schematically, the results of [13] can be summarized in the following formula,[⋆]

$$\mathcal{Z}(G,N,\lambda A) = \sum_{g=1}^{\infty}\frac{1}{N^{2g-2}}\sum_{n,\tilde{n}=0}^{\infty}\mathrm{e}^{-(n+\tilde{n})\lambda A/2}\sum_{i=0}^{2g-2-(n+\tilde{n})(2G-2)}\left(\frac{\lambda A}{2}\right)^i\omega^{g,G}_{i,n,\tilde{n}}. \tag{1.3}$$

$\lambda\equiv e_0^2N$ is the combination of e_0 and N that is kept fixed in the large N limit, and $\omega^{g,G}_{i,n,\tilde{n}}$ are A-independent numbers, whose explicit form can be extracted from the results of [13] (see also §6 below).

⋆ Throughout this paper, we will limit our discussion of large-N QCD (and subsequently the QCD string theory) to $G>0$, in order to avoid the complications that lead to the Douglas-Kazakov phase transition on the sphere.

In the "string" interpretation of (1.3), $1/N$ is the string coupling constant, and g is the genus of the (not necessarily connected) string worldsheet, which covers M outside a finite number of points, where specific singularities (such as branchpoints and collapsed handles) can emerge. In this interpretation, n and $\tilde{n}$ correspond to the number of sheets that cover the target in the orientation preserving and reversing sectors, respectively, and $\omega^{g,G}_{i,n,\tilde{n}}$ are interpreted as specific symmetry factors that essentially count the number of covering maps with the allowed types of singularities. In this section we will only need the following properties of $\omega^{g,G}_{i,n,\tilde{n}}$: (i) they vanish identically unless either n or $\tilde{n}$ is positive, (ii) they are invariant under the reversal of worldsheet orientation, $n \leftrightarrow \tilde{n}$, and (iii) not all of them are positive.

The results of [13] thus provide a wealth of data, and we can learn some useful lessons about the string theory that is supposed to reconstruct these data by a path integral over worldsheet geometries. A brief look at (1.3) indicates that such a string theory should be highly unusual, especially in comparison with the string theories of quantum gravity.† Here is a list of several basic properties of the large-N expansion as summarized in (1.3), which will thus serve as constraints on the QCD string Lagrangian:

1. Folds of the maps from the worldsheet to the target are dynamically suppressed; i.e. they are either gauge artifacts, or their contribution is identically zero.
2. Contributions to (1.3) from trivial homotopy classes of maps are zero. In particular, for targets of genus higher than zero, there is no contribution from worldsheets of spherical topology.
3. Unlike fundamental string theory, the QCD string spectrum contains neither the tachyon (massless in two dimensions) nor the graviton-dilaton multiplet. In fact, the large-N expansion indicates [13] that the physical spectrum of the one-string states in 2D QCD string theory contains exactly one bosonic string state for each non-trivial element of $\pi_1(M)$.
4. The string theory should be invariant under the $\mathbf{Z}_2$ parity transformation that reverses worldsheet orientation while leaving the target intact. We call a theory that respects this $\mathbf{Z}_2$ symmetry "non-chiral."
5. The theory should generate negative weights in the path integral, which suggests the presence of worldsheet fermions.
6. At general values of the gauge coupling λ, the theory only depends on the target metric through its total dimensionless area λA (and the individual areas λA_i in case of a subdivision of the spacetime by Wilson lines). The theory is invariant under a w_∞

† Lacking a better terminology, we will call the string theory of quantum gravity the "fundamental string theory" henceforth.

symmetry of area-preserving target diffeomorphisms, inherited from the w_∞ symmetry of (1.1).

7. The terms exponential in λA behave as $(n+\tilde{n})\lambda A$, where n and $\tilde{n}$ is the number of sheets of the cover that preserve and reverse orientation, respectively. The exponential terms thus measure the total induced area of the worldsheet (as opposed e.g. to the degree of the map, the latter being proportional to $n-\tilde{n}$), and λ is the physical tension of the string. The rest of the area dependence takes the form of finite polynomials in A for a given set of data $(g, G, n, \tilde{n}, \ldots)$, with the top power of A equal to $2g-2-(n+\tilde{n})(2G-2)$.

8. The $\lambda = 0$ limit (i.e. the limit of zero target area/zero gauge coupling constant/zero string tension) should be described by a topological string theory. In this limit, the path integral should not depend on the spacetime metric, and all physical correlation functions (including the $\lambda \to 0$ limit of all the partition functions) should be topological invariants, only depending on such general data as the genus of the worldsheet and the target.

In this paper we present a string theory that fulfills these requirements in a very simple manner. It turns out, in fact, that once we construct a relatively simple theory that satisfies these criteria, direct calculations reproduce the results of the two-dimensional QCD on arbitrary compact surfaces (i.e., the qualitative structure of (1.3) *and* the numerical values of $\omega^{g,G}_{i,n,\tilde{n}}$). Heuristically, this is so because the number of topological invariants that can be defined and subsequently calculated as partition functions of a two dimensional string theory is quite limited; in other words, not so many "string universality classes" exist that respect all the symmetries summarized above. This also suggests that other worldsheet Lagrangians might exist that describe the same "string universality class" (i.e. give the same set of partition and correlation functions),‡ and the question is which realization of the string universality class of large-N 2D QCD will prove most efficient for extensions to higher dimensions.

Several remarks are in order:

1. Extension of the results to higher dimensions is of course the central motivation for the study of two dimensional QCD strings. This program can only be considered successful if it provides some hints about the string description of QCD in higher dimensions. The Lagrangian presented here suggests a particularly natural and non-trivial extension to higher dimensions. This extension goes well beyond topological theory, and will be discussed elsewhere [16].

‡ Since our results were announced in [15], a very interesting alternative approach to the 2D QCD string theory has been suggested and discussed by Cordes, Moore and Ramgoolam [17]. Although different from the topological rigid string theory discussed in [15] and the present paper, the theory proposed in [17] can be probably considered a different realization of the same string universality class, at least in the regimes analyzed in [17] (i.e., in the chiral sector and/or at $\lambda = 0$.)

2. Even though the two dimensional QCD string theory is topological at zero gauge coupling/zero string tension, its extension to non-zero string tension goes beyond topological theory. In two dimensions, we can still add a tension term to the topological rigid string, and evaluate its contribution to the path integral perturbatively in λ using topological methods. A better substantiation of this procedure would however require an apparatus that goes beyond the scope of this paper.
3. Our approach to the problem of QCD string theory is phenomenological, i.e. instead of deriving the worldsheet theory directly from first principles (such as the Yang-Mills path integral), we compare the two theories at the level of their physical correlation functions (i.e., we try to construct another representant of the same string universality class). Any "microscopic" derivation of the string theory from the Yang-Mills theory would be extremely valuable.

1.2 Some Results in Two-Dimensional Yang-Mills Theory

Before proceeding to the definition of our string theory, we first summarize some aspects of two-dimensional Yang-Mills gauge theory, on arbitrary compact surfaces. This summary will prove useful later, since throughout this paper we will be encountering worldsheet phenomena quite reminiscent of the pattern already established in the spacetime Yang-Mills theory. It is by no means a review, and should rather serve as a reminder of those selected features of 2D Yang-Mills theory that are directly relevant to the present paper; for more details, the reader is referred to the original sources [11].

Start with a topological Yang-Mills theory, whose basic BRST multiplet is given by

$$[Q,A_\mu]=\psi_\mu,\qquad \{Q,\psi_\mu\}=D_\mu\phi,\qquad [Q,\phi]=0. \tag{1.4}$$

Here $A\equiv A_\mu\,\mathrm{d}x^\mu$ is the gauge field, $\psi\equiv\psi_\mu\,\mathrm{d}x^\mu$ the fermionic topological ghost, D_μ the covariant derivative defined by A, ϕ a scalar ghost-for-ghost; all fields are in the adjoint representation of the gauge group $\mathcal{G}$. We will also write $F\equiv F_{\mu\nu}\,\mathrm{d}x^\mu\wedge\mathrm{d}x^\nu$ for the field strength of A_μ.

The standard construction of the theory uses the flatness condition $F=0$ as the gauge fixing condition. In order to write a Lagrangian for this theory, we must introduce antighost multiplets (see [11] for details). The Lagrangian is then written as a BRST commutator,

$$\mathcal{L}=\int_M \mathrm{d}^2x\,\sqrt{g}\,\left(F^2+\text{ghost terms}\right), \tag{1.5}$$

and its path integral is localized to the moduli spaces of flat connections on M.

Without spoiling the topological symmetry, one can deform the Lagrangian by terms that are formally BRST commutators. One such term $\mathcal{L}'$ of ghost number minus two has been found in [11]; its addition to the original Lagrangian,

$$\mathcal{L} \to \mathcal{L} + t_0 \mathcal{L}', \tag{1.6}$$

leads to the following dramatic consequences:

1. Even though the new term in the Lagrangian is a BRST commutator, it does change the partition functions of the theory, since it brings in some new components of the moduli spaces from the infinity in the space of all gauge connections.
2. The antighost multiplets can be integrated out, and the whole theory can be written exclusively in terms of the fields contained in the basic BRST multiplet (1.4). After the integration over the antighost multiplets, the Lagrangian is roughly given by

$$\mathcal{L} = \frac{1}{t_0} \int_M \mathrm{d}^2 x \, \sqrt{g} \, \left((D_\mu F^{\mu\nu})^2 + \text{ghost terms} \right). \tag{1.7}$$

3. The deformed theory is now localized to the moduli spaces of all solutions to the Yang-Mills equations $D_\mu F^{\mu\nu} = 0$, rather than the moduli spaces of flat connections. The path integral gets contributions not only from the absolute minima of the Yang-Mills action (1.1), but also from unstable solutions of (1.1).

The partition function of the physical Yang-Mills theory can be evaluated as a correlation function of a specific physical observable in the associated deformed topological Yang-Mills theory. Local physical observables of the topological theory are given by BRST cohomology classes (equivariant with respect to the Yang-Mills gauge symmetry), and can be constructed from invariant polynomials on the Lie algebra of the gauge group, by evaluating the polynomials at the Lie-algebra valued ghost-for-ghost field ϕ of (1.4). These point-like observables $\mathcal{O}^{(0)}$ can be used to construct non-local observables, which are BRST invariant only when integrated over a cycle on the spacetime manifold. The non-local observables are related to the point-like ones by the BRST descent hierarchy,

$$\mathrm{d}\mathcal{O}^{(0)} = \{Q, \mathcal{O}^{(1)}\}, \qquad \mathrm{d}\mathcal{O}^{(1)} = \{Q, \mathcal{O}^{(2)}\} \qquad \mathrm{d}\mathcal{O}^{(2)} = 0. \tag{1.8}$$

The simplest such observables are given by

$$\mathcal{O}_0^{(0)} = \mathrm{Tr}\,(\phi^2), \qquad \mathcal{O}_0^{(1)} = 2\,\mathrm{Tr}\,(\phi\psi), \qquad \mathcal{O}_0^{(2)} = 2\,\mathrm{Tr}\,(\phi F - \psi \wedge \psi). \tag{1.9}$$

The statement that relates the physical Yang-Mills partition function to the correlation function of a physical observable in the topological theory can then be summarized in the following formula:

$$\langle 1 \rangle_{\text{physical YM}} = \left\langle \exp\left\{ -\int_M (\phi F - \psi \wedge \psi) - e_0^2 \int_M \sqrt{g}\, \phi^2 \right\} \right\rangle_{\text{topo. YM}} \tag{1.10}$$

The two observables on the right hand side are indeed the $\mathcal{O}_0^{(0)}$ and $\mathcal{O}_0^{(2)}$ of (1.9). The $\psi \wedge \psi$ term in the exponential plays a relatively minor role, while the remaining part of the exponential term in (1.10) is exactly the Yang-Mills Lagrangian, rewritten in a first-order form with the bosonic ghost-for-ghost ϕ playing the role of an auxiliary field. The path integral on the right hand side of (1.10) is the one of the deformed topological theory, as defined by (1.7). For more details see [11].

The striking analogies between the spacetime theory (i.e. the Yang-Mills path integral) and the worldsheet theory (the path integral of the 2D QCD string) that we will encounter later on provide yet another example of the well-documented "as above, so below" phenomenon of string theory, in which the existence of certain structures in spacetime (such as gravity, gauge invariance, supersymmetry, orbifolds, duality, etc.) is tied to the existence of analogous structures on the string worldsheet. This heuristic principle of string theory works remarkably well, for reasons that still remain mostly mysterious.

2. Harmonic Topological Sigma Models

In §1.1 we argued that in the zero tension limit, the string theory of 2D QCD should be topological. Since apparently the only field that we can use to construct a worldsheet Lagrangian is the map Φ from worldsheet Σ to the target manifold M, the first crude expectation is that the theory may be a certain form of a topological sigma model. The theory must be parity invariant on the worldsheet, though, which invalidates the standard topological sigma model. In that theory, the path integral is dominated by holomorphic maps [18], which of course makes the theory chiral. Moreover, for generic fixed worldsheet and target metrics on two-dimensional manifolds of higher genera, no holomorphic maps exist [19], and the path integral will be either identically zero or may even have problems with topological invariance. One could start with a chiral theory, and worry about the full non-chiral theory later (with such options as a coupling of two sectors with opposite chirality either directly or via an anomaly). It seems much more natural, however, to start with a non-chiral theory from the beginning, which is the approach that we follow in this paper.

Instead of starting with holomorphic maps, we therefore choose a different gauge-fixing codition for the topological sigma model, namely harmonicity of Φ with respect to fixed metrics on Σ and M. This condition has several important properties:

1. It is manifestly non-chiral.

2. The moduli space $\mathcal{M}$ of harmonic maps contains as a subspace the moduli space of holomorphic maps; due to the $\mathbf{Z}_2$ chiral symmetry, $\mathcal{M}$ contains all anti-holomorphic maps as well.
3. In the case of two-dimensional targets, a deep mathematical theory exists [19-21] that guarantees the existence of at least one harmonic map for a generic choice of the worldsheet and target metric, the worldsheet and target genera, and the homotopy class of Φ. (For more details, see §3.)

Although the ultimate interest of this paper is in the topological rigid string theory (as the 2D QCD string theory), we discuss the harmonic topological sigma models in some detail first, for the following two reasons:

1. The topological rigid string is directly related to the harmonic topological sigma model, by gauging worldsheet diffeomorphism symmetry. In this section, we will be able to explain and understand some features shared by the topological rigid string theory in the much simpler setting of topological sigma models, i.e. before the formulas become complicated by gauged worldsheet diffeomorphisms and dynamical worldsheet gravity.
2. Harmonic topological sigma models are interesting in their own right. We will see below that they are related in several distinct ways to their holomorphic counterparts. Furthermore, the theory of harmonic maps between manifolds is itself an intensely studied part of mathematics, with many important results and unexpected ramifications (see e.g. [19] and references therein). Hence, harmonic topological sigma models should be relevant to algebraic geometry, mirror symmetry and fundamental string theory.

The topological sigma model [18] is a theory of maps $\Phi : \Sigma \to M$ from a worldsheet Σ with coordinates σ^a, $a = 1, 2$,[⋆] to a target manifold M with coordinates x^μ, $\mu = 1, \ldots, D$. We assume that M carries a fixed Riemann structure, defined by a metric with components $g_{\mu\nu}$. For the purposes of the gauge fixing, we choose a fixed auxiliary metric h_{ab} on Σ. Notice that unlike in standard topological sigma models, we do not pick a complex structure in the target, nor do we use the worldsheet complex structure explicitly.

The Riemann structures on Σ and M allow us to define the Laplacian on the maps Φ from Σ to M. In coordinates, we have

$$\Delta x^\mu \equiv h^{ab}\nabla_a\partial_b x^\mu = h^{ab}\left(\partial_a\partial_b x^\mu + \Gamma^\mu_{\sigma\rho}\,\partial_a x^\sigma\partial_b x^\rho - \Gamma^c_{ab}\,\partial_c x^\mu\right). \tag{2.1}$$

Here Γ^c_{ab} and $\Gamma^\mu_{\sigma\rho}$ are the Christoffel symbols of the metric connection associated with h_{ab} and $g_{\mu\nu}$ respectively, and ∇_a denotes the covariant derivative on $T^*\Sigma \otimes \Phi^{-1}(TM)$.

⋆ Here we assume that Σ is oriented and without a boundary; an extension to worldsheets with boudaries and crosscaps is discussed in §2.5.

To use the harmonicity condition

$$\Delta x^\mu = 0 \tag{2.2}$$

as the gauge fixing condition in the topological sigma model, we first introduce the BRST multiplet that maps x^μ to their ghosts,

$$[Q, x^\mu] = \psi^\mu, \qquad \{Q, \psi^\mu\} = 0. \tag{2.3}$$

While the ghost fields are uniquely determined by the original fields of the theory, in our case x^μ, the antighosts and their auxiliary fields are determined by the gauge fixing condition. In the case of (2.2), the antighosts χ^μ and auxiliaries B^μ are sections of $\Phi^{-1}(TM)$:

$$\{Q, \chi^\mu\} = B^\mu, \qquad [Q, B^\mu] = 0. \tag{2.4}$$

The gauge fixing condition is then implemented by the following choice of the Lagrangian:

$$\mathcal{L} = \{Q, \frac{1}{r_0}\int_\Sigma \mathrm{d}^2\sigma\,\sqrt{h}\,\chi^\mu g_{\mu\nu}(\Delta x^\nu + \frac{1}{2}\Gamma^\nu_{\sigma\rho}\,\chi^\sigma\psi^\rho - \frac{1}{2}B^\nu)\}. \tag{2.5}$$

The apparently non-covariant term with the explicit dependence on $\Gamma^\mu_{\sigma\rho}$ is needed for spacetime diffeomorphism invariance. While this term is standard in topological sigma models, we will see a natural explanation of its existence in this particular model later.

Upon performing the BRST commutator in (2.5), the Lagrangian becomes

$$\begin{aligned}\mathcal{L} = \frac{1}{r_0}\int_\Sigma \mathrm{d}^2\sigma\,\sqrt{h}\,\Big\{&-\frac{1}{2}B^\mu g_{\mu\nu}B^\nu + B^\mu g_{\mu\nu}(\Delta x^\nu + \Gamma^\nu_{\sigma\rho}\,\chi^\sigma\psi^\rho) - \Delta x^\mu g_{\mu\nu}\,\Gamma^\nu_{\sigma\rho}\,\chi^\sigma\psi^\rho\\ &- \chi^\mu g_{\mu\nu}\,\Delta\psi^\nu - R_{\mu\nu\sigma\rho}\,h^{ab}\partial_a x^\nu\partial_b x^\sigma\,\chi^\mu\psi^\rho - \left(\frac{1}{4}R_{\mu\nu\sigma\rho} + \frac{1}{2}g_{\lambda\kappa}\Gamma^\lambda_{\mu\sigma}\,\Gamma^\kappa_{\nu\rho}\right)\chi^\mu\psi^\sigma\chi^\nu\psi^\rho\Big\}.\end{aligned} \tag{2.6}$$

In this formula, the Laplacian Δ acting on ψ^μ is the covariant Laplacian on $\Phi^{-1}(TM)$, i.e.

$$\begin{aligned}\Delta\psi^\mu = h^{ab}\nabla_a\nabla_b\psi^\mu \equiv h^{ab}&\left(\partial_a\partial_b\psi^\mu + 2\Gamma^\mu_{\sigma\rho}\,\partial_a x^\sigma\partial_b\psi^\rho - \Gamma^c_{ab}\,\partial_c\psi^\mu\right)\\ &+ \Gamma^\mu_{\sigma\rho}\,\Delta x^\sigma\psi^\rho + h^{ab}\partial_a x^\sigma\partial_b x^\rho\,\psi^\nu\left(R^\mu{}_{\rho\sigma\nu} + \partial_\nu\Gamma^\mu_{\sigma\rho}\right).\end{aligned} \tag{2.7}$$

While x^μ and B^μ carry ghost number zero, ψ^μ and χ^μ are of ghost number $+1$ and -1, respectively. The ghost number generates a $\mathrm{U}(1)$ symmetry of the theory; as we will se below, this symmetry is even preserved quantum mechanically, unlike in standard topological sigma models.

B^μ is an auxiliary field and can be eliminated from the Lagrangian by using its equation of motion,

$$B^\mu = \Delta x^\mu + \Gamma^\mu_{\sigma\rho}\,\chi^\sigma\psi^\rho, \tag{2.8}$$

which reduces the Lagrangian to

$$\begin{aligned}\mathcal{L} =& \frac{1}{r_0}\int\limits_\Sigma \mathrm{d}^2\sigma\,\sqrt{h}\,\left\{\frac{1}{2}\Delta x^\mu\, g_{\mu\nu}\,\Delta x^\nu - \chi^\mu\, g_{\mu\nu}\,\Delta\psi^\nu\right.\\ &\left. - R_{\mu\sigma\rho\nu}\, h^{ab}\partial_a x^\sigma\partial_b x^\rho\,\chi^\mu\psi^\nu - \frac{1}{4}R_{\mu\nu\sigma\rho}\chi^\mu\psi^\sigma\chi^\nu\psi^\rho\right\}.\end{aligned} \tag{2.9}$$

The simplicity of (2.9) makes it one of the most useful expressions for the Lagrangian of the harmonic topological sigma model. (2.9) is BRST invariant under

$$[Q, x^\mu] = \psi^\mu, \qquad \{Q,\psi^\mu\} = 0 \qquad \{Q,\chi^\mu\} = \Delta x^\mu + \Gamma^\mu_{\sigma\rho}\,\chi^\sigma\psi^\rho. \tag{2.10}$$

To demostrate the nilpotence of this BRST charge and the invariance of the Lagrangian, we have to use the equations of motion, so these properties are only valid on-shell.

Sometimes it is convenient to keep the off-shell BRST symmetry by retaining the auxiliary fields B^μ explicitly, and rewrite the Lagrangian in a first order form. Up to a total derivative, we indeed have[⋆]

$$\mathcal{L} = \{Q, \frac{1}{r_0}\int\limits_\Sigma \mathrm{d}^2\sigma\,\sqrt{h}\,\left[-h^{ab}\nabla_a\chi\cdot\partial_b x + \frac{1}{2}\chi^\mu g_{\mu\nu}(\Gamma^\nu_{\sigma\rho}\,\chi^\sigma\psi^\rho - B^\nu)\right]\}. \tag{2.11}$$

The explicit evaluation of the BRST commutator then gives

$$\begin{aligned}\mathcal{L} = \frac{1}{r_0}\int\limits_\Sigma \mathrm{d}^2\sigma\,\Bigg\{& - h^{ab}\nabla_a B\cdot\partial_b x + h^{ab}\nabla_a\chi\cdot\nabla_b\psi - \frac{1}{2}B^2 + R_{\mu\sigma\rho\nu}h^{ab}\partial_a x^\sigma\partial_b x^\rho\chi^\mu\psi^\nu\\ &+ (B^\mu - \Delta x^\mu)\,g_{\mu\nu}\Gamma^\nu_{\sigma\rho}\,\chi^\sigma\psi^\rho - \left(\frac{1}{4}R_{\mu\sigma\nu\rho} + \frac{1}{2}\Gamma^\lambda_{\mu\nu}\,g_{\lambda\kappa}\Gamma^\kappa_{\sigma\rho}\right)\chi^\mu\psi^\nu\chi^\sigma\psi^\rho\Bigg\}.\end{aligned} \tag{2.12}$$

This first-order form is useful for the comparison with the topological rigid string theory, whose Lagrangian is most natural in a similar first-order form.

⋆ From now on, we simplify notation by introducing "$\cdot$" to denote the scalar product in TM; thus, we write $v^\mu g_{\mu\nu}w^\nu \equiv v\cdot w$ for any two vectors v^μ, w^ν from TM.

2.1 Symmetry Between Ghosts and Antighosts

The harmonic topological sigma model enjoys a remarkable property that is by no means generic to all topological field theories: The ghost and antighost fields ψ^μ and χ^μ are both sections of the same bundle $\Phi^{-1}(TM)$ over the worldsheet. This fact suggests the possibility of a hidden symmetry in the Lagrangian. Indeed, it is easy to show that the Lagrangian is indeed symmetric under a bosonic U(1) symmetry that mixes ghosts and antighosts, and is generated by

$$\begin{aligned} &[J,\psi^\mu]=\chi^\mu, \qquad &&[J,\chi^\mu]=-\psi^\mu,\\ &[J,x^\mu]=0, \qquad &&[J,B^\mu]=0. \end{aligned} \tag{2.13}$$

The interchange of ghosts and antighosts is the $\mathbf{Z}_2$ subgroup of this U(1).

The existence of this U(1) symmetry allows us to define another BRST-like supersymmetry charge $\bar{Q}$,

$$\begin{aligned} &[\bar{Q},x^\mu]=\chi^\mu, \qquad &&\{\bar{Q},\chi^\mu\}=0,\\ &\{\bar{Q},\psi^\mu\}=-B^\mu, \qquad &&[\bar{Q},B^\mu]=0. \end{aligned} \tag{2.14}$$

The two supercharges are of course related by $\bar{Q}=[J,Q]$. They anticommute with each other and are both nilpotent,

$$\{Q,\bar{Q}\}=0, \qquad Q^2=0, \qquad \bar{Q}^2=0. \tag{2.15}$$

Notice that all the fields of the theory now fall into one irreducible multiplet of the extended supersymmetry algebra (2.15). This is again reminiscent of the topological Yang-Mills theory, which can also be expressed (after a deformation that leads from the moduli spaces of flat connections to the moduli spaces of all solutions to Yang-Mills equations) in terms of the single BRST multiplet that only contains ghosts and no antighosts.

Because of the symmetry between ghosts and antighosts, Q should no longer play a preferred role in the Lagrangian. We can indeed use the new supercharge $\bar{Q}$ to write the Lagrangian as a double commutator,

$$\mathcal{L}=\{Q,[\bar{Q},F]\}. \tag{2.16}$$

Because $\bar{Q}$ carries ghost number minus one, F is a bosonic function of ghost number zero. It is easy to demonstrate that the following choice of F,

$$F=-\frac{1}{2r_0}\int_\Sigma \mathrm{d}^2\sigma\,\sqrt{h}\,\left(h^{ab}\partial_a x\cdot\partial_b x-\chi\cdot\psi\right), \tag{2.17}$$

when substituted in (2.16), reproduces the Lagrangian of the harmonic topological sigma model, (2.12).

The extremely simple form of F has important consequences. We will use (2.17) elsewhere [16] to draw some useful analogies between the harmonic topological sigma models and topological rigid strings on one hand, and some models studied in other areas of physics, such as the physics of polymers, disordered systems and stochastic quantization, on the other. Already at this stage, however, we can use the double commutator expression for the Lagrangian to explain the existence of the peculiar $\Gamma^{\mu}_{\sigma\rho}$-dependent term in the gauge fixing fermion of (2.5), which now comes from the manifestly covariant gauge fixing boson F as a variation of $g_{\mu\nu}$ under $\bar{Q}$.

2.2 Relations to Holomorphic Topological Sigma Models

Harmonic topological sigma models are related in several distinct ways to their holomorphic counterparts. Since this line of thought is not directly related to the main aim of this paper, we will discuss this issue only briefly. These relations are important because they place the theory of harmonic topological sigma models into a wider and much better understood context, and provide a different perspective of the models.

The key to one such correspondence is the analogy between the formal structure of the harmonic topological sigma models and the properties of the topological Yang-Mills theory (as summarized in §1.2). In the previous subsection we have seen that the harmonic topological sigma model exhibits a symmetry between ghosts and antighosts, and all fields of the theory form an irreducible multiplet of an extended supersymmetry algebra. In analogy with the Yang-Mills theory (cf. §1.2), we can interpret this multiplet as a multiplet of fields and ghosts in another topological field theory (with a double-topological symmetry [22], cf. also [23]), and the Lagrangian of the harmonic topological sigma model as an effective Lagrangian in which all auxiliaries and antighosts of the double-topological field theory have already been integrated out. The associated topological field theory is a double-topological holomorphic sigma model.

The construction can be outlined as follows. Consider $\mathcal{N}=4$ supersymmetric sigma model on a hyper-Kähler manifold, for which we choose T^*M where M is complex. (After a relation to the harmonic topological sigma model on M is established, nothing depends on the auxiliary complex structure of M.) A topological twist exists that turns two fermionic partners of the target coordinates into worldsheet scalars. These two fermionic scalars play the role of topological ghosts of the double-topological BRST algebra, which contains two fermionic scalar nilpotent charges. The two remaining fermions of the original $\mathcal{N}=4$ supersymmetry now are one-forms on the worldsheet, and can be interpreted as antighost fields of the double-topological symmetry. We can eliminate them (toghether with their auxiliaries) from the Lagrangian by deforming the Lagrangian of the twisted $\mathcal{N}=4$ supersymmetric sigma model $\mathcal{L}$ by a term which violates the ghost number of the double-topological theory by two,

$$\mathcal{L}\to\mathcal{L}+r_0\mathcal{L}'. \tag{2.18}$$

Here $\mathcal{L}'$ is chosen such that the anti-ghosts and auxiliaries of the double-topological sigma model can be eliminated by their equations of motion, which set them equal to worldsheet derivatives of the remaining fields. Schematically, if χ_z^I denotes one of the anti-ghost fields of the double-topological sigma model, its equation of motion will relate it to one of the ghost fields ψ^I:

$$\chi_z^I \propto \frac{1}{r_0}\partial_z \psi^I. \tag{2.19}$$

After the elimination of all auxiliaries and anti-ghosts (i.e. all fields that carry a non-trivial worldsheet tensor structure), we are left with an effective Lagrangian which only contains the target coordinates x^μ, another bosonic field B^μ, and the ghost fields ψ^μ, χ^μ of the double-topological symmetry, and turns out to be the Lagrangian of the harmonic topological sigma model.

Notice several facts:

1. This interpretation of $\mathcal{L}$ as a deformed double-topological sigma model allows us to take the limit of $r_0 \to 0$ in the harmonic topological sigma model, which is otherwise not well defined in (2.9).

2. Since the double-topological holomorphic sigma model is a twisted version of a $\mathcal{N}=4$ supersymmetric sigma model, the construction outlined above indicates that in the theory of harmonic topological sigma models we are in fact dealing with a twisted $\mathcal{N}=4$ theory in disguise (cf. [23]). An analogous observation can be made in the case of the topological rigid string theory discussed in §4.

3. There is a remarkable similarity between the double-topological sigma model and the topological theory used in [17] for the construction of an alternative approach to 2D QCD string theory. Hence, the relation between the double-holomorphic topological sigma models and harmonic topological sigma models might shed some light on the relation between the string theory of [17] on one hand, and the topological rigid string theory on the other.

Another relation between harmonic topological sigma models and their more thoroughly studied holomorphic counterparts arises as follows. The theory of harmonic maps between manifolds is one of the simplest mathematical theories that allows for a twistor description [21]. This twistor correspondence canonically identifies the moduli space of harmonic maps from Σ to M, with a moduli space of holomorphic maps from Σ to a different target M', associated with M by the twistor correspondence. (In many cases, the natural almost complex structure of M' is not integrable, which is not an obstacle for the construction of holomorphic topological sigma models [18].) Typically, M' is fibered over M, and holomorphic maps to M' project to harmonic maps to M. For example, a typical pair of manifolds related by this twistor correspondence is

$$M = S^4, \qquad M' = \mathbf{C}P^3. \tag{2.20}$$

If extendable to the full topological sigma model, the twistor correspondence between M and M' would give a map between the harmonic topological sigma model on M and its holomorphic counterpart with a different target, M'. This is somewhat reminiscent of mirror symmetry, which also relates two topological sigma models with different target manifolds.

2.3 Partition Functions

By arguments standard in topological theories, the path integral of the harmonic topological sigma model is independent of the coupling constant r_0, and can be exactly calculated in the semiclassical approximation. In the weak coupling limit of $r_0 \ll 1$, the whole integral is localized to the locus of solutions to the gauge-fixing condition, i.e. to the moduli space $\mathcal{M}$ of harmonic maps. (Notice that $\mathcal{M}$ is also the moduli space of all classical solutions to an associated bosonic problem, in this case the bosonic non-linear sigma model. This is again quite analogous to a similar phenomenon in the Yang-Mills theory, cf. §1.2.)

Given a harmonic map Φ_0, the next step is to evaluate the one-loop determinants of quantum fluctuations around Φ_0. In topological field theories, the bosonic part of this determinant cancels against its fermionic counterpart (at least up to a sign). In our case, this cancellation is a consequence of the following BRST formula,

$$[Q, \Delta x^\mu] = \Delta\psi^\mu - R^\mu{}_{\sigma\rho\nu}\, h^{ab}\partial_a x^\sigma \partial_b x^\rho \psi^\nu - \Gamma^\mu_{\sigma\rho}\, \Delta x^\sigma\, \psi^\rho. \tag{2.21}$$

When evaluated on a harmonic map, the last term on the right hand side vanishes, and we are left with the linearized equation of motion for the ghost. The non-zero modes of the corresponding operators are thus related by the BRST transformation, and the fermionic one-loop determinant (almost) exactly cancels the bosonic determinant:

$$\frac{\det'\left(\delta^\mu_\nu \Delta - R^\mu{}_{\sigma\rho\nu}\, h^{ab}\partial_a x^\sigma \partial_b x^\rho\right)}{\det'^{1/2}\left\{\left(\delta^\mu_\nu \Delta - R^\mu{}_{\sigma\rho\nu}\, h^{ab}\partial_a x^\sigma \partial_b x^\rho\right)^2\right\}} = (-1)^{\#}. \tag{2.22}$$

In this formula, Δ is the Laplacian on $\Phi^{-1}(TM)$ (since both ψ^μ and the quantum fluctuations δx^μ of x^μ are sections of $\Phi^{-1}(TM)$), the prime means that zero modes are omitted, and # denotes the number of negative eigenvalues of $\delta^\mu_\nu \Delta - R^\mu{}_{\sigma\rho\nu}\, h^{ab}\partial_a x^\sigma \partial_b x^\rho$.

The only remaining calculation is the integration over the zero modes, in particular the integral over the moduli spaces of harmonic maps. The fermionic zero modes satisfy the linearized equation of motion in the harmonic background,

$$\Delta\psi_0^\mu - R^\mu{}_{\sigma\rho\nu}\, h^{ab}\partial_a x^\sigma \partial_b x^\rho\, \psi_0^\nu = 0. \tag{2.23}$$

In the theory of harmonic maps, this equation is known as the Jacobi equation. The number of integrable solutions to this equation measures the dimension of the moduli space of harmonic maps, and the zero modes of ψ^μ form a basis of one-forms on the moduli spaces. It is useful to decompose a given fermionic zero mode ψ_0^μ in a normalized basis $f_I^\mu(\sigma)$ (with $I = 1, \ldots, \dim\mathcal{M}$) of solutions to (2.23), as follows:

$$\psi_0^\mu = \sum_I a^I f_I^\mu(\sigma). \tag{2.24}$$

The anticommuting coefficients a^I can be identified with one-forms on the moduli space, $a^I \propto \mathrm{d}m^I$, with m^I a coordinate system on $\mathcal{M}$.

Since the anti-ghost zero modes χ_0^μ satisfy exactly the same equation (2.23) as the ghost zero modes ψ_0^μ, they can be similarly written as

$$\chi_0^\mu = \sum_I b^I f_I^\mu(\sigma). \tag{2.25}$$

Thus, the ghost number is preserved quantum mechanically, without any anomaly, and the partition functions can be non-zero without any insertions of BRST invariant observables.

Since the curvature-dependent two-fermi term in $\mathcal{L}$ was actually a part of the kinetic term of the fermions, the only term left in the zero-mode integration is the curvature-dependent four-fermi term, whose form is identical to the analogous four-fermi term in topological mechanics [24]. Hence, the whole path integral reduces to

$$\int_{\mathcal{M}} (-1)^{\#} \int \prod \mathrm{d}a^I\, \mathrm{d}b^I \exp\left\{-\frac{1}{4} a^I a^J b^K b^L R_{IJKL}(m)\right\}, \tag{2.26}$$

where R_{IJKL} is a tensor on the moduli space, induced from the target curvature tensor by

$$R_{IJKL} \equiv \int_\Sigma \mathrm{d}^2\sigma\, \sqrt{h}\, R_{\mu\nu\sigma\rho} f_I^\mu f_J^\nu f_K^\sigma f_L^\rho. \tag{2.27}$$

The integral over the anti-ghost zero modes b^I in (2.26) gives the Euler density on the moduli space, with the zero modes of ψ^μ playing the role of one-forms on $\mathcal{M}$, $a^I \propto \mathrm{d}m^I$. The remaining integral over $\mathcal{M}$ thus gives the Euler number $\chi(\mathcal{M})$.

In the homotopically trivial sector, the calculation is identical to that of topological mechanics. The moduli space is the target manifold itself, the fermionic zero-mode integral gives the Euler character density on the target, and the integral over the bosonic moduli is equal to the Euler number $\chi(M)$. Path integrals in non-trivial homotopy classes yield interesting stringy corrections to $\chi(M)$; in the simplest cases, they count the number of harmonic maps in the given homotopy class (cf. §3 below).

2.4 Observables and Correlation Functions

Observables in topological field theories are defined as cohomology classes of the BRST charge. In topological sigma models, the simplest observables are point-like on the world-sheet, and are in one-to-one correspondence with the cohomology classes of the target manifold. Given a differential s-form $A \equiv \sum A_{\mu_1 \ldots \mu_s} \mathrm{d}x^{\mu_1} \wedge \ldots \wedge \mathrm{d}x^{\mu_s}$ on M, we define

$$\mathcal{O}_A = \sum A_{\mu_1 \ldots \mu_s} \psi^{\mu_1} \ldots \psi^{\mu_s}. \tag{2.28}$$

$\mathcal{O}_A$ is of course BRST invariant for A a closed form, and BRST exact for A an exact form. Since this argument depends neither on the choice of the gauge fixing condition nor on the antighost multiplet, these "homology observables" exist in any topological sigma model, independently of the specific choice of the gauge fixing condition. [$\mathcal{O}^{(1)}, \mathcal{O}^{(2)}$ not easy to define – another indication that the theory is an effective theory, with some fields already integrated out.]

When the fundamental group of the target is non-trivial, another class of observables can exist. For every element γ of the fundamental group $\pi_1(M)$, consider the vacuum state $\mathcal{O}_\gamma$ of the string in the winding sector γ. This state is a non-trivial BRST cohomology class of the theory, as can be seen from the following argument. The on-shell BRST transformation rules

$$[Q, x^\mu] = \psi^\mu, \qquad \{Q, \psi^\mu\} = 0, \qquad \{Q, \chi^\mu\} = \Delta x^\mu + \Gamma^\mu_{\sigma\rho} \chi^\sigma \psi^\rho \tag{2.29}$$

indicate that states with $\chi^\mu = \psi^\mu = 0$ are BRST invariant if they are annihilated by Δx^μ, which means that the linear part of x^μ is BRST closed but not exact, and $\mathcal{O}_\gamma$ are physical. Since these new observables are parametrized by the elements of the first homotopy group of the target, it is natural to call them "homotopy observables." (Similar observables have been discussed in [25].) The full space of physical observables is roughly a tensor product of the homotopy and homology sectors. One must be careful, however, since the products of homology and homotopy observables can in some cases be singular, and the space of physical states is then a subspace in the direct product.

Although the structure of the point-like observables in the harmonic and holomorphic topological sigma models is very similar, their correlation functions differ dramatically. The symmetry between ghosts and antighosts, not shared by holomorphic topological sigma models, leads to the quantum mechanical conservation of the ghost number. This absence of ghost number anomaly shows up in correlation functions as a special selection rule. Indeed, the correlation functions are zero, unless the total ghost number of all observables under the correlator vanishes:

$$\langle \mathcal{O}_1 \ldots \mathcal{O}_s \rangle = 0 \qquad \text{if} \qquad \sum_{i=1}^{s} \text{ghost } \# \, (\mathcal{O}_i) \neq 0. \tag{2.30}$$

Since all point-like cohomology observables (2.28) with $s > 0$ carry a positive ghost number, all their correlations vanish identically:

$$\langle \mathcal{O}_{A_1} \ldots \mathcal{O}_{A_s} \rangle = 0 \qquad \text{if } \deg{(A_i)} \neq 0 \text{ for any } A_i. \tag{2.31}$$

Unless we compensate for the ghost number of the non-trivial cohomology classes by additional insertions of (non-local) observables with negative ghost numbers, the physical observables corresponding to the cohomology classes of non-zero degree effectively decouple from the correlation functions, and we are left with the bosonic homotopy observables parametrized by the elements of $\pi_1(M)$. This is qualitatively in a very good agreement with the results of the large-N expansion in 2D QCD, where we have one physical string state for each non-trivial element of $\pi_1(M)$. In the topological rigid string, the trivial homotopy class becomes unphysical by virtue of worldsheet diffeomorphism invariance.

2.5 Equivariant Harmonic Topological Sigma Models

So far we have assumed that the worldsheet manifold is oriented and without a boundary. Topological sigma models can be naturally extended to worldsheets with boundaries and/or crosscaps [26]. While this extension may be interesting for pure mathematical reasons (since it refines the topological invariants already calculated by the theory on oriented surfaces), our motivation here is different. Indeed, we conjecture in §7 that the equivariant theory is related to the Yang-Mills theory with alternative gauge groups ($\mathrm{SO}(N), \mathrm{Sp}(N)$) in the same way the original theory is related to Yang-Mills theory with gauge group $\mathrm{SU}(N)$.

The framework that allows us to construct this extension systematically is the theory of equivariant topological sigma models. The central idea is to consider the $\mathbf{Z}_2$ symmetry I_0 that reverses worldsheet orientation, i.e. acts in a suitable coordinate system $(z, \bar{z})$ on Σ by

$$I_0 : \Sigma \to \Sigma, \qquad I_0(z, \bar{z}) = (\bar{z}, z), \tag{2.32}$$

and treat I_0 as an orbifold symmetry (i.e., a discrete gauge symmetry) of the theory on closed oriented surfaces. The resulting orbifold model is a topological sigma model on worldsheets with boundaries/crosscaps.

In order to construct an orbifold theory wit I_0 as a generator of the orbifold group, we must extend I_0 to a $\mathbf{Z}_2$ symmetry of the full quantum theory. Since all fields of a topological sigma model are tensors of a specific degree on the worldsheet, a canonical extension of I_0 to the fields exists, which leaves the target intact. If this canonical extension of I_0 were a symmetry of the model, it could be used to define a canonical orbifold theory (essentially, with standard, Neumann boudary conditions on the open strings).

In holomorphic topological sigma models, this canonical extension of I_0 fails to be a symmetry of the quantum theory, since it transforms holomorphic maps into anti-holomorphic ones. Consequently, there is no canonical orbifold theory parity associated with a given holomorphic topological sigma model (unless there are no non-trivial instantons, a case which is not of great interest). To promote I_0 to a $\mathbf{Z}_2$ symmetry, we must pick an anti-holomorphic involution of the target,

$$I : M \to M \tag{2.33}$$

and gauge the diagonal $\mathbf{Z}_2$,

$$\mathbf{Z}_2 = \{1, I_0 I\}\,. \tag{2.34}$$

Twisted states of such an orbifold theory are open strings with both ends fixed to the submanifold of fixed points of I. This class of models has been discussed in detail in [26] (see also [27]). Physical correlation functions of this theory calculate an equivariant extension of the quantum cohomology algebra on M.

Harmonic topological sigma models, on the other hand, are by construction parity invariant on the worldsheet, and the canonical extension of I_0 is a symmetry of the theory and can be gauged. Hence, a canonical orbifold model exists, with standard Neumann boundary conditions at both ends of the open strings. Since these boudary conditions do not spoil the conservation of the ghost number, all homology observables with non-zero degree decouple from the theory by the same argument as in the closed string model. Notice that the open-string sector does not produce any non-trivial homotopy observables, since every open string is homotopically trivial.

3. Harmonic Topological Sigma Models in Two Dimensions

Since the definition of harmonic topological sigma models has not required a choice of a complex structure on the target, this class of topological sigma models can be studied in arbitrary target dimensions. With our primary motivation in mind, however, we will now restrict our attention to two-dimensional targets.

3.1 Targets of Genus $G > 1$

Consider the path integral of the harmonic topological sigma model in a fixed homotopy class $[\Sigma_g, M_G]$ of maps from a two-dimensional worldsheet Σ_g of genus g to a two-dimensional target M_G of genus G with G, g greater than one, and assume that the degree of the map is non-zero. In the set of all homotopy classes, such choice is generic. Given a fixed Riemannian metrics h_{ab} on the worldsheet Σ_g and $g_{\mu\nu}$ on the target M_G, a deep mathematical theorem [19] proves the existence of exactly one harmonic map in $[\Sigma_g, M_G]$; moreover, this harmonic map minimizes the action of the bosonic sigma model within $[\Sigma_g, M_G]$. Consequently, the one-loop determinants (2.22) cancel exactly, and we obtain an extremely simple answer for the partition function of the harmonic topological sigma model in such a (generic) homotopy sector:

$$\mathcal{Z}([\Sigma_g, M_G]) = 1. \tag{3.1}$$

Since the moduli space consists of only one point, the path integral trivially confirms the expectation that $\mathcal{Z}$ is equal to the Euler number of the moduli space.

Even in the homotopy classes of degree zero, harmonic maps always exist, and the worldsheet is mapped by any harmonic map either to a simple geodesic curve in M_G (in case the homotopy class is non-trivial), or to a point in M_G (in the trivial homotopy class). In the latter case, the moduli space coincides with the target itself, and the partition function equals the Euler number of M_G.

3.2 Targets of Genus $G \leq 1$

Since for targets of genus greater than one we are always guaranteed the existence of a harmonic map in any homotopy class, the path integral is always well-defined and easy to calculate. General theory of harmonic maps between surfaces shows, however, that the situation is much more complicated for targets of low genera.

Targets of genus one (with a flat metric $g_{\mu\nu}$) are easy to deal with, since the existence theorem is still valid. A given harmonic map is no longer unique in its homotopy class, however, since we can use the target isometry group to generate a two-parameter class of harmonic maps parametrized by the target itself. The evaluation of the path integral is then proportional to the Euler number of the target, and the partition function is always zero.

When the target is either the sphere or the projective plane, there are known homotopy classes of maps with no harmonic representative, which makes the formal semiclassical evaluation of the path integral in such homotopy classes ill-defined. Thus, for example, there is no harmonic map of degree ± 1 from the torus to the sphere, no matter what metrics we choose on Σ_g and M_G. This singular behavior on the sphere is reminiscent of some properties of the large-N QCD on the sphere (such as the Douglas-Kazakov phase transition). A closer examination of the harmonic topological sigma model in these singular cases should certainly be interesting.

3.3 From Sigma Models to String Theory

Before we go on and construct a topological string theory using the harmonic topological sigma model as worldsheet matter, let us briefly consider our options.

In the topological sigma models, the topology of Σ and the homotopy class of Φ have been fixed. In string theory, we must be able to sum over all worldsheet genera and all homotopy classes of maps. While trying to define this sum, we must deal with worldsheet diffeomorphisms. Even though the partition functions of the harmonic topological sigma model in a given homotopy class are $\mathrm{Diff}_0(\Sigma_g)$ invariant, the naive sum of the partition functions over all worldsheet genera and homotopy classes of maps,

$$\sum_g g_{\mathrm{string}}^{2g-2} \sum_{[\Sigma_g, M_G]} \mathrm{e}^{-c \deg \Phi} \langle 1 \rangle_{[\Sigma_g, M_G]} \,, \tag{3.2}$$

is infinite. (In (3.2), g_{string} is the string coupling constant, $\langle 1 \rangle_{[\Sigma_g, M_G]}$ is the partition function of the harmonic topological sigma model in homotopy class $[\Sigma_g, M_G]$, $\deg \Phi$ is the degree of a map $\Phi \in [\Sigma_g, M_G]$, and c is a "chemical potential" introduced here to regularize the sum over all values of $\deg \Phi$.) This infinity can be easily traced back to the symmetry under global worldsheet diffeomorphisms, given by the mapping class group $\Gamma_{\Sigma_g} \equiv \mathrm{Diff}(\Sigma_g)/\mathrm{Diff}_0(\Sigma_g)$.

There are several possible remedies for this infinity. The minimal way which makes the sum over all homotopy classes finite is to simply divide the infinite sum by the (infinite) volume of the mapping class group, and define the string partition function by

$$\mathcal{Z} = \sum_g g_{\mathrm{string}}^{2g-2} \frac{1}{\mathrm{Vol}\,(\Gamma_{\Sigma_g})} \sum_{[\Sigma_g, M_G]} \mathrm{e}^{-c \deg \Phi} \langle 1 \rangle_{[\Sigma_g, M_G]} \,, \tag{3.3}$$

Since we have already calculated earlier in this section most of the ingredients of the right hand side of (3.3) (most of the contributions are actually equal to one), direct comparison with (1.3) shows that the final result for $\mathcal{Z}$ is indeed very different from the results of the large-N expansion in 2D QCD.

The second option is a little more sophisticated, since it makes the theory diffeomorphism invariant by coupling the matter theory to topological gravity. While such theory in general depends on precise details of this coupling, the most straightforward approach is to consider harmonic maps and allow the worldsheet metric to vary. This "minimal" coupling leads to a theory localized to moduli spaces which are canonically fibered over the moduli spaces of Riemann surfaces. The minimal coupling leads to a further ramification, depending on how we treat the symmetry between ghosts and antighosts found in the harmonic topological sigma model. We can either treat this symmetry as accidental and couple the matter to usual topological gravity, or we can interpret the harmonic topological sigma model as a theory with double topological symmetry and couple it to double topological gravity (a theory which calculates the Euler numbers of the moduli spaces of Riemann surfaces). Although both of these conservative approaches might be of some independent interest and apparently lead to self-consistent theories, in this paper we follow a different route, explained in the following section.

4. Topological Rigid String Theory

In the standard setting, the Lagrangian of a topological (string) theory is constructed as an exact BRST commutator,

$$\mathcal{L} = \{Q, \Psi\}, \tag{4.1}$$

where Ψ is a suitably chosen gauge-fixing fermion. In our case, the only fields that describe the string are the coordinates x^μ of the map Φ from the worldsheet Σ to the spacetime M. In particular, we do not introduce an independent worldsheet metric, and will use the induced one whenever a metric is needed. The basic BRST multiplet then consists of x^μ and their ghost partners ψ^μ,

$$[Q, x^\mu] = \psi^\mu, \qquad \{Q, \psi^\mu\} = 0. \tag{4.2}$$

Of course, ψ^μ are components of a section of $\Phi^{-1}(TM)$.

As always in topological field theory, symmetries are more important than the Lagrangian itself, and we will discuss them first. In addition to the topological symmetry we consider worldsheet diffeomorphisms a gauge symmetry. This additional symmetry distinguishes the model from a theory of topological worldsheet matter and makes it a string theory.

The presence of diffeomorphism invariance as an additional gauge symmetry requires new ghosts in the BRST multiplet, which now becomes

$$[Q, x^\mu] = \psi^\mu + c^a \partial_a x^\mu, \qquad \{Q, \psi^\mu\} = c^a \partial_a \psi^\mu, \qquad \{Q, c^a\} = c^b \partial_b c^a, \tag{4.3}$$

and causes a typical overcounting of gauge symmetries. As a consequence of this overcounting, the theory will enjoy a new, fermionic gauge symmetry given by

$$\delta_\epsilon x^\mu = 0, \qquad \delta_\epsilon \psi^\mu = \epsilon^a \partial_a x^\mu. \tag{4.4}$$

The standard strategy for taking care of this ghostly symmetry is to introduce a ghost for ghost field ϕ^a, and extend the BRST multiplet to

$$\begin{aligned} [Q, x^\mu] &= \psi^\mu + c^a \partial_a x^\mu, \qquad & \{Q, \psi^\mu\} &= c^a \partial_a \psi^\mu + \phi^a \partial_a x^\mu, \\ \{Q, c^a\} &= c^b \partial_b c^a - \phi^a, \qquad & [Q, \phi^a] &= c^b \partial_b \phi^a + \phi^b \partial_b c^a. \end{aligned} \tag{4.5}$$

The BRST multiplet is already becoming complicated, and we can simplify things by agreeing to work directly with diffeomorphism invariant configurations only. This restriction allows us to ignore the diffeomorphism ghosts c^a, and leads to the so-called equivariant BRST quantization. In fact, this equivariant approach will turn out to be very effective for the comparison of the topological rigid string to the large-N expansion of two-dimensional QCD, as the moduli spaces emerging in the latter are manifestly diffeomorphism invariant. In the equivariant theory, the BRST multiplet is reduced to

$$[Q, x^\mu] = \psi^\mu, \qquad \{Q, \psi^\mu\} = \phi^a \partial_a x^\mu, \qquad [Q, \phi^a] = 0, \tag{4.6}$$

and the BRST charge is only nilpotent on diffeomorphism invariant configurations.

We will actually go one step further, and throughout most of the paper will keep the fermionic gauge symmetry along with the ordinary diffeomorphism invariance as a manifest gauge symmetry of the theory, without explicitly fixing either of them. One of the benefits of this strategy is the simplification of the subsequent formulas, which would otherwise contain many terms depending on ϕ^a and the gauge-fixing multiplets associated with it.

To construct a Lagrangian, we need a gauge fixing condition; we choose the minimal-area condition,

$$\Delta x^\mu = 0. \tag{4.7}$$

The Laplacian in (4.7) is now the covariant Laplacian on x^μ, defined with respect to the induced metric on the worldsheet:*

$$\Delta x^\mu \equiv h^{ab}\nabla_a\partial_b x^\mu = h^{ab}\left(\delta^\mu_\nu - \partial_c x^\mu h^{cd}\partial_d x^\lambda g_{\lambda\nu}\right)\left(\partial_a\partial_b x^\nu + \Gamma^\nu_{\sigma\rho}\,\partial_a x^\sigma\partial_b x^\rho\right). \tag{4.8}$$

This of course means that the maps that satisfy (4.7) are harmonic in their own induced metric, i.e. they satisfy the harmonicity condition with respect to the induced connection

$$\Gamma^c_{ab} = h^{cd}\partial_d x^\mu\, g_{\mu\nu}\left(\partial_a\partial_b x^\nu + \Gamma^\nu_{\sigma\rho}\,\partial_a x^\sigma\partial_b x^\rho\right). \tag{4.9}$$

The theory is a non-minimal coupling of the harmonic topological sigma model to topological gravity. Note also that our gauge-fixing condition coincides with the equation of motion of the bosonic Nambu-Goto string. Once again, we are constructing a theory whose path integral will be localized to the moduli spaces of all classical solutions of an associated bosonic theory, in this case the Nambu-Goto string theory.

The gauge-fixing condition (4.7) has $D-2$ independent components, as it should, since two components of x^μ should stay unfixed by virtue of worldsheet diffeomorphism invariance. There are two constraints on Δx^μ, expressing the fact that Δx^μ (as the trace of the second fundamental form of Φ) is normal to $\Phi(\Sigma)$:

$$\partial_a x\cdot\Delta x = 0. \tag{4.10}$$

Our gauge fixing condition is a section of the normal bundle $\mathcal{N}$ to the worldsheet. Since we will be encountering sections of $\mathcal{N}$ very frequently, we introduce a special notation for the inner product induced in $\mathcal{N}$ by the target metric $g_{\mu\nu}$; from now on, we write

$$v * w \equiv v\cdot w - v\cdot\partial_a x\, h^{ab}\,\partial_b x\cdot w \tag{4.11}$$

for the inner product of the normal parts of any two vectors v^μ and w^μ from $\Phi^{-1}(TM)$.

The gauge fixing condition requires us to introduce antighosts and auxiliaries,

$$\{Q,\chi^\mu\} = B^\mu, \qquad [Q,B^\mu] = 0. \tag{4.12}$$

Since the gauge fixing function Δx^μ is a section of $\mathcal{N}$, so are χ^μ and B^μ. The fermionic gauge symmetry (4.4) extends to the antighost multiplet by

$$\delta_\epsilon\chi^\mu = 0, \qquad \delta_\epsilon B^\mu = \epsilon^a\partial_a\chi^\mu. \tag{4.13}$$

$\star$ Hoping not to create too much confusion, we keep the notation of the previous section. In this section, the worldsheet metric h_{ab} is always the induced metric, while in the previous section, it was always the fixed auxiliary metric.

4.1 Restoration of the Ghost-Antighost Symmetry

Unlike harmonic topological sigma models, topological rigid string theory does not exhibit manifest symmetry between its ghosts and antighosts. The ghost field ψ^μ represents infinitesimal deformations of the map Φ, and is a section of $\Phi^{-1}(TM)$. The antighost field, as we have just seen, is a section of the same bundle as the gauge fixing function Δx^μ. In harmonic topological sigma models, Δx^μ was also a section of $\Phi^{-1}(TM)$. In the topological rigid string, worldsheet diffeomorphism invariance makes Δx^μ a section of the normal bundle $\mathcal{N} \subset \Phi^{-1}(TM)$, hence spoiling the symmetry between ghosts and antighosts.

There are several motivations for restoration of the ghost-antighost symmetry:

1. The harmonic topological sigma model can be naturally interpreted as a deformation of the double-topological holomorphic sigma model, and as a twisted $\mathcal{N} = 4$ supersymmetric theory. Similar structure can be expected in the topological rigid string. In the double-topological theory, both the ghosts and the antighosts are members of the same BRST multiplet, and must be sections of the same bundle. Also, the theory is then described in terms of a single BRST multiplet, in analogy with a similar property of the Yang-Mills theory (cf. §1.2).

2. The normal bundle $\mathcal{N}$ is not always well-defined. In particular, generic maps to two-dimensional targets are not immersions, which makes $\mathcal{N}$ always ill-defined. In explicit calculations, it is more convenient to deal with sections of the regular bundle $\Phi^{-1}(TM)$ instead.

3. The existence of two BRST charges $Q, \bar{Q}$ in the harmonic topological sigma model have allowed us to write its Lagrangian as a very simple double commutator. A similar formula will hold for the Lagrangian of the topological rigid string, and will allow us to draw interesting analogies between the topological rigid string theory and the physics of polymers, disordered systems and stochastic quantization.

4. In the formulation with manifest symmetry between ghosts and antighosts, the overall ghost number anomaly is manifestly zero, which leads to a simple selection rule on physical correlation functions, and effectively decouples observables with positive ghost number from observables with ghost number zero (such as string winding modes).

The worldsheet gauge invariance that spoils the symmetry between ghosts and antighosts comes to the rescue, and allows us to restore this symmetry. Since the longitudinal part of ψ^μ is a pure gauge of the ghostly gauge symmetry, the gauge-invariant parts of ψ^μ and χ^μ are both sections of $\mathcal{N}$. Instead of gauge fixing the longitudinal part of ψ^μ, we can go in the opposite direction and restore the symmetry between ghosts and antighosts by enlarging the gauge symmetry. With this in mind, define

$$\begin{aligned} \delta_\varepsilon x^\mu &= 0, & \delta_\varepsilon \psi^\mu &= 0, \\ \delta_\varepsilon \chi^\mu &= \varepsilon^a \partial_a x^\mu, & \delta_\varepsilon B^\mu &= \varepsilon^a \partial_a \psi^\mu. \end{aligned} \tag{4.14}$$

This symmetry is imposed as yet another local fermionic symmetry of the theory. Instead of being set to zero from the outset, the longitudinal part of χ^μ is now a pure gauge of the new gauge symmetry.

Just like in harmonic topological sigma models, the restoration of the symmetry between ghosts and antighosts allows us to define a second fermionic nilpotent charge $\bar{Q}$:

$$\begin{aligned} [\bar{Q}, x^\mu] &= \chi^\mu, & \{\bar{Q}, \chi^\mu\} &= 0, \\ \{\bar{Q}, \psi^\mu\} &= -B^\mu, & [\bar{Q}, B^\mu] &= 0. \end{aligned} \tag{4.15}$$

Together with Q they form an extended supersymmetry algebra,

$$\{Q, \bar{Q}\} = 0, \qquad Q^2 = 0, \qquad \bar{Q}^2 = 0. \tag{4.16}$$

All fields of the model fall into an irreducible representation of (4.16).

4.2 Itoi-Kubota Symmetry

In the previous section we extended the gauge symmetry of the model, in order to restore the symmetry between ghosts and antighosts. In this process we have actually obtained more than we required. In addition to the new fermionic gauge symmetry, we have introduced a new bosonic gauge symmetry,

$$\delta_v B^\mu = v^a \partial_a x^\mu, \qquad \delta_v(\text{other fields}) = 0. \tag{4.17}$$

This new symmetry is produced from the two fermionic gauge symmetries (4.14) and (4.4), (4.13) as their anticommutator,

$$\delta_v \propto \{\delta_\epsilon, \delta_\varepsilon\}. \tag{4.18}$$

It is indeed a local symmetry, and allows us to consider B^μ a section of $\Phi^{-1}(TM)$, by making the longitudinal part of B^μ a gauge artifact.

It is interesting to note that the bosonic gauge symmetry (4.17) has actually been introduced in the bosonic rigid string theory quite some time ago, by Itoi and Kubota [28]. The original motivation of the authors of [28] for introducing this gauge symmetry was quite different from ours, however. Here we have seen how the Itoi-Kubota symmetry naturally emerges in the supersymmetry algebra of the topological rigid string, as a consequence of the symmetry between ghosts and antighosts.

4.3 The Theory

The topological rigid string Lagrangian can be written as a sum of two parts. The first part is given by

$$\mathcal{L}_1 = \frac{1}{\alpha_0} \int_\Sigma \mathrm{d}^2\sigma \, \sqrt{h} \, \Big\{ h^{ab} \nabla_a B \cdot \partial_b x - h^{ab} \nabla_a \chi \cdot \nabla_b \psi - R_{\mu\sigma\rho\nu} \, h^{ab} \partial_a x^\sigma \partial_b x^\rho \, \chi^\mu \psi^\nu \\ + \Big(h^{ab} h^{cd} - h^{ac} h^{bd} - h^{ad} h^{bc} \Big) \nabla_a \psi \cdot \partial_b x \, \nabla_c \chi \cdot \partial_d x + \Delta x^\mu g_{\mu\nu} \Gamma^\nu_{\sigma\rho} \, \chi^\sigma \psi^\rho \Big\} . \tag{4.19}$$

$\mathcal{L}_1$ is linear in the auxiliary field B^μ and quadratic in the fermionic fields ψ^μ and χ^μ, and all these fields can in principle be integrated out. The integral over B^μ gives a delta function localized to the moduli space of minimal-area maps, while the integral over the fermions produces a volume element on the moduli space. $\mathcal{L}_1$ is of course constructed as an exact BRST commutator,

$$\mathcal{L}_1 = \{Q, \Psi_1\} = \{Q, \frac{1}{\alpha_0} \int_\Sigma \mathrm{d}^2\sigma \, \sqrt{h} \, h^{ab} \nabla_a \chi \cdot \partial_b x\}. \tag{4.20}$$

The second part of the topological rigid string Lagrangian smears out the delta fuction by introducing a term that is essentially $\propto B^2$, made covariant under all local symmetries. Its full expression in terms of all fields is quite complicated, and we only write it here implicitly as a BRST commutator,

$$\mathcal{L}_2 = \{Q, \Psi_2\}, \tag{4.21}$$

with Ψ_2 given by

$$\Psi_2 = \frac{1}{\alpha_0} \int_\Sigma \mathrm{d}^2\sigma \, \sqrt{h} \, \Big\{ \chi * \big(B - \Gamma_{\sigma\rho} \, \chi^\sigma \psi^\rho \big) \\ + h^{ab} \left(\psi * \chi \, \nabla_a \chi \cdot \partial_b x + \psi * \nabla_a \chi \, \chi \cdot \partial_b x - \chi * \nabla_a \chi \, \psi \cdot \partial_b x \right) \Big\} . \tag{4.22}$$

One can find the full form of the Lagrangian by performing the BRST commutator explicitly, if one wishes so.

The expression for $\mathcal{L}$ becomes manageable when we keep only the diffeomorphism symmetry, and fix all the other gauge symmetries in a special gauge. A particularly natural gauge choice for both of the fermionic gauge symmetries and the Itoi-Kubota symmetry is

$$\psi \cdot \partial_a x = 0, \qquad \chi \cdot \partial_a x = 0, \qquad B \cdot \partial_a x = 0. \tag{4.23}$$

In other words, we have used the gauge symmetries to set the longitudinal components of all worldsheet fields to zero. In this particular gauge, the Lagrangian simplifies and can be explicitely written as follows:

$$\begin{aligned}
\mathcal{L}' = \mathcal{L}_1 + a\mathcal{L}_2 = \frac{1}{\alpha_0} \int_\Sigma \mathrm{d}^2\sigma \sqrt{h} \, \Big\{ \Big(& - B \cdot \Delta x - h^{ab} \nabla_a \chi \cdot \nabla_b \psi - R_{\mu\sigma\rho\nu} h^{ab} \partial_a x^\sigma \partial_b x^\rho \, \chi^\mu \psi^\nu \\
& + (h^{ab} h^{cd} - h^{ac} h^{bd} - h^{ad} h^{bc}) \nabla_a \psi \cdot \partial_b x \, \nabla_c \chi \cdot \partial_d x + \Delta x^\mu g_{\mu\nu} \Gamma^\nu_{\sigma\rho} \, \chi^\sigma \psi^\rho \Big) \, (1 + a\psi \cdot \chi) \\
& + a \Big(- B^2 + 2B^\mu g_{\mu\nu} \Gamma^\nu_{\sigma\rho} \, \chi^\sigma \psi^\rho - R_{\mu\nu\sigma\rho} \chi^\mu \psi^\nu \chi^\sigma \psi^\rho - \chi^\sigma \psi^\rho \Gamma_{\sigma\rho} * \Gamma_{\mu\nu} \, \chi^\mu \psi^\nu \\
& \qquad + (\psi \cdot \nabla_a \chi) \, h^{ab} \, (\nabla_b \psi \cdot \chi) - (\psi \cdot \nabla_a \psi) \, h^{ab} \, (\nabla_b \chi \cdot \chi) \\
& + h^{ab} \nabla_a \psi \cdot \partial_b x \, (-\chi \cdot B + \chi^\mu g_{\mu\nu} \Gamma^\nu_{\sigma\rho}) + h^{ab} \nabla_a \chi \cdot \partial_b x \, (\psi \cdot B - \psi^\mu g_{\mu\nu} \Gamma^\nu_{\sigma\rho} \, \chi^\sigma \psi^\rho) \\
& \qquad + (\psi \cdot \nabla_a \chi - \chi \cdot \nabla_a \psi) \, h^{ab} \, \partial_b x^\mu g_{\mu\nu} \Gamma^\nu_{\sigma\rho} \, \psi^\sigma \chi^\rho \Big) \Big\} .
\end{aligned} \tag{4.24}$$

B^μ can of course be integrated out, its equation of motion being

$$\begin{aligned}
B^\mu = -\frac{1}{2a}(1 + a\psi\chi)\Delta x^\mu + \frac{1}{2}\chi^\mu(\nabla\psi \cdot \partial x) - \frac{1}{2}\psi^\mu(\nabla\chi \cdot \partial x) \\
+ \Big(\delta^\mu_\nu - \partial_a x^\mu h^{ab} \partial_b x^\lambda g_{\lambda\nu} \Big) \, \Gamma^\nu_{\sigma\rho} \, \chi^\sigma \psi^\rho.
\end{aligned} \tag{4.25}$$

The explicit expression (4.24) for the Lagrangian is not very illuminating, and is presented here only for completeness. For all practical puproses, the only important properties of the Lagrangian are:

1. The theory has been constructed as a topological string theory, according to the prescription of the Mathai-Quillen formalism. In this sense, the model is (formally) exactly integrable since its partition functions calculate an equivariant Euler number of the moduli spaces [29-32].
2. Even without invoking the topological character of the theory, we will be able to rewrite the Lagrangian in a surprisingly simple form, amenable to a simple physical interpretation (see eqns. (4.29) – (4.31) below).
3. The theory is a topological version of the rigid string theory.

Later on, it will prove useful to set $a = 1$ and add another BRST exact term to the Lagrangian. After that, the full Lagrangian is given by

$$
\begin{aligned}
\mathcal{L} = \{Q, \frac{1}{\alpha_0} \int_\Sigma \mathrm{d}^2\sigma \, \frac{\sqrt{h}}{(1 - \psi * \chi)^2} \Big\{ & h^{ab} \nabla_a \chi \cdot \partial_b x + \chi * \big(B - \Gamma_{\sigma\rho} \, \chi^\sigma \psi^\rho \big) \\
& + h^{ab} \left(\psi * \nabla_a \chi \, \chi \cdot \partial_b x - \chi * \nabla_a \chi \, \psi \cdot \partial_b x - \psi * \chi \, \nabla_a \chi \cdot \partial_b x \right) \Big\} \, .
\end{aligned} \tag{4.26}
$$

The new term that has been added to the original Lagrangian contains only higher-order terms in the fermi fields, and does not bring in any new branches of the moduli spaces from the infinity in the space of all field configurations. Hence, the term has been designed in such a way that its addition to the Lagrangian should not change the value of the path integral.

It is instructive to look at the bosonic sector of the resulting Lagrangian. Upon setting all fermi fields to zero, both $\mathcal{L}$ and $\mathcal{L}'$ simplify to

$$
\mathcal{L}_0 = \frac{1}{\alpha_0} \int_\Sigma \mathrm{d}^2\sigma \, \sqrt{h} \, \left(-B \cdot \Delta x - B * B \right) . \tag{4.27}
$$

Integrating out the auxiliary fields B^μ, one gets

$$
\mathcal{L}_0 = \frac{1}{4\alpha_0} \int_\Sigma \mathrm{d}^2\sigma \, \sqrt{h} \, \Delta x \cdot \Delta x , \tag{4.28}
$$

which is the Lagrangian of the bosonic rigid string theory (at zero string tension). Hence, our topological theory, derived from the requirement of localization to the moduli spaces of minimal-area maps, can indeed be considered a topological version of the rigid string theory. Remarkably, the bosonic rigid string theory (in four target dimensions) has been studied some time ago by Polyakov [33], Kleinert [34] and others [35], as a candidate for QCD string theory.

Although its explicit component form is complicated, the full Lagrangian of the topological string theory can be written in an extremely simple and form, using the second supersymmetry charge $\bar{Q}$ of (4.15). This supercharge allows us to write the Lagrangian as a double commutator. Thus, we can write $\mathcal{L}_1$ of (4.19) as

$$
\mathcal{L}_1 = \{Q, [\bar{Q}, \frac{1}{\alpha_0} \int_\Sigma \mathrm{d}^2\sigma \, \sqrt{h} \,]\} . \tag{4.29}
$$

Similarly, the second part of the Lagrangian, as given by (4.21) and (4.22), can be written as

$$\mathcal{L}_2 = \{Q, [\bar{Q}, \frac{1}{\alpha_0} \int\limits_{\Sigma} \mathrm{d}^2\sigma \, \sqrt{h} \, \chi * \psi]\}, \tag{4.30}$$

and the deformed Lagrangian (4.26) takes the form

$$\mathcal{L} = \{Q, [\bar{Q}, \frac{1}{\alpha_0} \int\limits_{\Sigma} \mathrm{d}^2\sigma \, \frac{\sqrt{h}}{1 - \psi * \chi}]\}. \tag{4.31}$$

Thus, the whole Lagrangian of the topological rigid string can be written as a simple double commutator, and its invariance under the U(1) symmetry that mixes ghosts and antighosts is now manifest. The simple form of (4.29) – (4.31) is the key to an analogy between the topological rigid string theory and some models studied in the physics of polymers and disordered systems [16], and leads to a quite non-trivial physics in higher target dimensions.

4.4 Partition Functions and Correlation Functions

The partition function of the topological rigid string theory for a given worldsheet Σ and target M calculates the equivariant Euler number of the (regular locus of the) moduli spaces of minimal-area maps from Σ to M. Similarly as in the harmonic topological sigma models, this claim can be confirmed by a direct semiclassical calculation at $\alpha_0 \ll 1$. A more elegant way to prove it is to notice that the theory has been constructed as an infinite-dimensional version of the Mathai-Quillen formalism, which is essentially a specific algorithm how to calculate Euler characters of vector bundles over manifolds (and their equivariant analogs, if there is a symmetry group acting on the vector bundle). The fact that topological field theories are infinite-dimensional versions of the Mathai-Quillen formalism has been first discussed by Atiyah and Jeffrey in [30]. Since this interpretation of topological field theories is well covered in the literature and is now considered standard, it will not be discussed here. (For an excellent short review aimed at physicists, see [32]; more details can be found in [29-31].)

Consider a minimal-area map Φ, given in coordinates by x^μ. Any deformation $x^{\mu\prime} = x^\mu + \delta x^\mu$ that is still a minimal area map to lowest order in δx^μ must satisfy the linearized minimal-area equation

$$\mathcal{J}^\mu_\nu \, \delta x^\nu = 0, \tag{4.32}$$

where

$$\mathcal{J}^\mu_\nu \equiv \left\{ \left(\delta^\mu_\nu - \partial_c x^\mu h^{cd} \partial_d x^\lambda g_{\lambda\nu} \right) \Delta - \left(g^{\mu\lambda} - \partial_c x^\mu h^{cd} \partial_d x^\lambda \right) R_{\lambda\sigma\rho\nu} \, h^{ab} \partial_a x^\sigma \partial_b x^\rho \right\}. \tag{4.33}$$

(Here we have chosen the gauge $\delta x \cdot \partial_a x = 0$, in order to eliminate the apparent deformations of Φ that correspond to reparametrizations of Σ.) This equation happens to be identical to the equation that defines the fermionic zero modes, i.e. it coincides with the linearized equation of motion for both the ghost and antighost fields (in the gauge (4.23)). Even in the theory of minimal-area maps, this equation is called the Jacobi equation, and its solutions are tangent to the moduli spaces of minimal-area maps.

Because of the symmetry between ghosts and antighosts, there is always an equal number of ghost and antighost zero modes. Hence, as in the sigma model case, there is no ghost number anomaly, and observables with positive ghost numbers effectively decouple from observables with ghost number zero.

5. Topological Rigid Strings in Two Dimensions

In this section, we continue our discussion of the topological rigid string, in the restricted class of two dimensional targets. In the next section, the partition functions of the topological rigid string theory will be compared to the results of the large N expansion in 2D QCD.

5.1 Suppression of Folds and Worldsheet Self-Avoidance

The functional integral of the topological rigid string theory is localized to the infinitesimal vicinity of the moduli spaces of all solutions to the gauge-fixing constraint, which in our case is

$$\Delta x^{\mu} = 0. \tag{5.1}$$

In D target dimensions, this condition has just $D-2$ independent components, as a result of worldsheet diffeomorphism invariance. In two target dimensions, we are apparently left with $D-2=0$ conditions! Yet, the condition (5.1) is non-trivial even in two dimensions. The naive counting of independent components of (5.1) was based on the transversality condition

$$\partial_a x \cdot \Delta x \equiv 0, \tag{5.2}$$

which represents two constrains on Δx whenever $\partial_a x^{\mu}$ is non-degenerate as a two-by-two matrix (i.e. in those points where Φ is an immersion). In two dimensions, generic maps are not immersions, and their induced metric is always degenerate somewhere. This fact makes the condition (5.1) non-trivial even in two dimensions. (5.1) is the equation of motion of the Nambu-Goto string and represents the minimal-area (or more precisely, critical-area) condition on Φ. Maps with folds of non-zero length are not critical-area maps, hence they violate (5.1) and do not contribute to the path integral. Thus, in two dimensions, the sole purpose of the minimal-area condition (5.1) is to suppress maps with folds of non-zero length.⋆

The suppression of folds has been identified as one of the crucial properties of the large-N expansion of 2D QCD, and is on general grounds (such as the strong coupling expansion on the lattice) expected from higher dimensional QCD string theory as well. The absence of folds in string theory means that the string worldsheet is self-avoiding. This self-avoidance is local on the worldsheet; in particular, it is different from spacetime self-avoidance of the strings, since two distinct worldsheet points are still allowed to occupy the same spacetime point. As it turns out, the topological rigid string can be derived from the bosonic Nambu-Goto string by imposing a simple condition of worldsheet self-avoidance on the latter [16].

5.2 Moduli Spaces of Minimal-Area Maps

We have treated diffeomorphism invariance of the theory as an equivariant symmetry, therefore we must keep it a manifest symmetry of our theory. In particular, the moduli spaces must be parametrized in a diffeomorphism invariant way. Fortunately, this requirement is a virtue rather than a constraint, and will allow us to understand the geometry of the moduli spaces in detail. (Here we follow the insight of Gross and Taylor, who parametrized the moduli spaces emerging from the large N expansion of 2D QCD in a closely related way.) With a simple parametrization of the moduli spaces at hand, we will be able to calculate their Euler numbers.

Since maps with no folds of non-zero length are coverings of M_G almost everywhere, we obtain a very simple classification of maps that solve (5.1). In a given homotopy class $[\Sigma_g, M_G]$, minimal-area maps are coverings of M_G outside a fininte number of points $P_1, \ldots, P_k \in M_G$. At each of these points, a generic map exhibits one of the following moduli:

1. A simple branchpoint of degree one. (We define the degree of a branchpoint as the number of covering sheets above a generic target point minus the number of sheets above the branchpoint.)
2. A collapsed handle. (Whenever a handle of Σ_g is mapped to M_G in a homotopically trivial way, the minimal-area condition (in the sense of stratified surfaces [36]) requires the handle to be mapped to a point in M_G.)

⋆ In general, the minimal-area equation (5.1) requires interpretation, and a mathematically precise specification of the class of maps that are considered solutions of (5.1) is a subtle issue [36]. It leads to a necessary extension of the naive definition of a surface to more general objects, such as stratified surfaces and multivarifolds [36]. In this paper we will try to avoid introducing these complicated mathematical objects, and only note here that our definition of minimal-area maps is that of stratified surfaces.

3. A collapsed disk. (A map with a homotopically trivial domain of Σ_g mapped into one point; although these maps also satisfy the minimal-area condition, we will see below that their contribution to the partition function is in fact zero.)

In addition to maps with various combinations of these moduli, other solutions of $\Delta x^\mu = 0$ exist, and are actually quite important in the path integral. These additional solutions represent maps of "critical area" rather than minimal area, and are ustable solutions of the associated bosonic theory (although they are of course absolute minima of the topological rigid string action). The fact that unstable solutions of the bosonic theory will contribute to the partition function is yet another similarity to the Yang-Mills theory in spacetime.

The critical-area maps add one more modulus type to the ones listed above:

4. A collapsed neck between two sheets of opposite orientation.

Henceforth we call the moduli of these four types the "simple moduli," to distinguish them from the moduli that emerge when two or more simple moduli coalesce (i.e. at the compactification locus of the moduli spaces). Notice also that maps from Σ to M with any combination of the simple moduli are all smooth maps.*

The moduli spaces of critical-area maps can then be described as follows. Consider a map that covers M_G by n ($\tilde{n}$) sheets of the same (opposite) orientation everywhere except in k distinct points $(P_1, \ldots, P_k)$, and pick one simple modulus for each P_i. To fix the homotopy class uniquely, in addition to the simple moduli at P_i we must specify how the sheets of the cover are permuted when we go around any non-contractible loop in M_G. Hence, we choose $4G$ elements a_j, b_j, $j = 1, \ldots, 2G$ of the group of permutations of the cover, $S_n \otimes S_{\tilde{n}}$. Altogether, these choices are constrained by one homotopy condition,

$$\prod_{i=1}^{k} p_i(P_i) \prod_{i=1}^{G} a_j b_j a_j^{-1} b_j^{-1} = 1 \quad \in \quad S_n \otimes S_{\tilde{n}}, \tag{5.3}$$

where $p_i(P_i)$ is the element of $S_n \otimes S_{\tilde{n}}$ that represents the permutation of the covering sheets when we go around P_i. When the simple modulus at P_i is a simple branchpoint, $p_i(P_i)$ is a permutation of two sheets of identical orientation; permutations associated with the other three types of simple moduli are trivial.

Each of the moduli is point-like in the target, and its target location P_i serves as a natural coordinate in the bulk of the moduli space. For a given choice of data that satisfy (5.3), define

$$\tilde{\mathcal{M}}^0_{G,k} = \{(P_1, \ldots, P_k), \quad P_i \in M_G, \quad P_i \neq P_j \text{ for } i \neq j\}\,. \tag{5.4}$$

As we will see later, the removed points provide a natural compactification of $\tilde{\mathcal{M}}^0_{G,k}$.

$\star$ While the differential topology of the maps is smooth, the induced metric is of course singular.

The actual bulk moduli spaces $\mathcal{M}^0_{G,k}$ are factors of $\tilde{\mathcal{M}}^0_{G,k}$ by the permutation group S_k that permutes P_i's,

$$\mathcal{M}^0_{G,k} = \tilde{\mathcal{M}}^0_{G,k}/S_k. \tag{5.5}$$

From the point of view of the worldsheet path integral, this factorization has two different sources:

1. When two simple moduli of the same type are on the same connected component of the worldsheet, they are indistinguishable (i.e. they are related by a global worldsheet diffeomorphism).
2. Two different simple moduli on the same connected worldsheet components, as well as two simple moduli of the same type but on different non-isomorphic worldsheet components, are distinguishable. Their contribution is however overcounted when we sum independently over all possible simple moduli at all P_i's. In order to avoid this overcounting, we must factorize $\tilde{\mathcal{M}}^0_{G,k}$ by S_k.

The non-compact bulk moduli spaces $\mathcal{M}^0_{G,k}$ (as well as their finite coverings $\tilde{\mathcal{M}}^0_{G,k}$) have a natural compactification dictated by the theory itself. Whenever two or more simple moduli coalesce in one point P in M_G, a composite modulus is formed. Among the composite moduli are maps with several branchpoints of degree one at a given location P, maps with a branchpoint of higher degree at P, maps with collapsed manifold of higher genus at P, maps with twisted connecting necks between sheets of different orientation, etc.; the structure of all possible composite moduli is uniquely determined by the simple moduli. The locations of composite moduli serve as coordinates on the compactification locus of the moduli spaces $\mathcal{M}^0_{G,k}$. Thus, given a component of the bulk moduli space $\mathcal{M}^0_{G,k}$, all possible composite moduli compactify the cover $\tilde{\mathcal{M}}^0_{G,k}$ to $\tilde{\mathcal{M}}_{G,k}$,

$$\tilde{\mathcal{M}}_{G,k} \equiv (M_G)^k = \{(P_1, \ldots, P_k), \quad P_i \in M_G\}. \tag{5.6}$$

The factorization of $\tilde{\mathcal{M}}_{G,k}$ by the symmetry group that permutes the copies of M_G turns this component of the compactified moduli space into an orbifold:

$$\mathcal{M}_{G,k} = \tilde{\mathcal{M}}_{G,k}/S_k. \tag{5.7}$$

So far we have compactified a single bulk component of the moduli space, with the homotopy class of Φ fixed uniquely by our fixed choice of a_j, b_j and the simple moduli at P_i. Since in general a given composite modulus can be created in several ways when different groups of simple moduli coalesce, two or more different bulk components of the total moduli space can have a common compactification locus. Consequently, the total moduli space is strictly speaking not even an orbifold, but it can always be decomposed into orbifolds, and its Euler number can always be uniquely defined using this decomposition.

Several other facts are worth noticing:

1. The parametrization of the moduli spaces by the target location of the allowed singularities is manifestly invariant under worldsheet diffeomorphisms.
2. Since this parametrization of the moduli spaces does not allow us to keep track of the connectivity of the worldsheet, our partition function sums over all worldsheet topologies, not only the connected ones. This and the connected partition function are of course related to each other exponentially.

5.3 Euler Numbers of the Moduli Spaces

The partition functions of the topological rigid string theory calculate the Euler number of the moduli spaces. As we have seen, the structure of the moduli spaces is in fact quite simple, and we can calculate their Euler numbers directly, for example by cell decomposition. Let us first recall the definition of the Euler number of an orbifold [37].

Just as manifolds are locally modelled by regions $\mathcal{U}_i$ in $\mathbf{R}^n$, an orbifold $\mathcal{O}$ is locally modelled by regions $\mathcal{U}_i$ of $\mathbf{R}^n/G_i$, where i goes over the set of all coordinate systems on $\mathcal{O}$, and G_i is a finite group that acts on $\mathbf{R}^n$. Coordinate changes are required to respect the group action by G_i in a natural manner, which allows us to define for each point x in $\mathcal{O}$ a group G_x, called the "isotropy group" of x, as the smallest G_i associated with a domain containing x. With this notation, the Euler number of $\mathcal{O}$ is defined as follows. Pick a cell decomposition of $\mathcal{O}$ which respects the isotropy groups on $\mathcal{O}$, i.e. all points in a given cell $\mathcal{C}$ have the same isotropy group, which we denote by $G_{\mathcal{C}}$. The Euler number is then given by a sum over all cells,

$$\chi(\mathcal{O}) = \sum_{\mathcal{C}} (-1)^{\dim \mathcal{C}} \frac{1}{|G_{\mathcal{C}}|}. \tag{5.8}$$

This definition of the Euler number is natural with respect to products and disjoint unions of orbifolds, a fact that will be used below.

Using (5.8), the orbifold Euler number of $(M_G)^k/S_k$ can then be calculated as follows:

$$\chi((M_G)^k/S_k) = \frac{1}{k!}\left\{\chi(\tilde{\mathcal{M}}^0_{G,k})\right\} + \ldots = \frac{1}{k!}(2-2G)^k. \tag{5.9}$$

The expression in the parentheses is the Euler number of the locus in $(M_G)^k$ on which S_k acts freely ("free locus" from now on), while the dots represent contributions from the subset of $(M_G)^k$ where at least two P_i's coincide, i.e. from the points with non-zero isotropy group.

The Euler number of $\mathcal{M}^0_{G,k}$ can be easily calculated by induction in k. For $k = 2$, the Euler number is easily computed directly. First we use the multiplicativity property of the Euler number to get $\chi((M_G)^2) = [\chi(M_G)]^2 = (2-2G)^2$, and then, using the additivity of

χ, we subtract from this result the Euler number of the diagonal part of $(M_G)^2$, which is equal to $2-2G$. That gives the Euler number of the free locus in $(M_G)^2$. Since S_2 acts on the free locus freely, the Euler number of $\mathcal{M}^0_{G,2}$ (which is the factor of the free locus by S_2) is $1/(2!)$ times the Euler number of $\tilde{\mathcal{M}}^0_{G,k}$:

$$\chi(\mathcal{M}^0_{G,2}) = \frac{1}{2}\left\{(2-2G)^2-(2-2G)\right\} = \frac{(2G-2)(2G-1)}{2}. \tag{5.10}$$

For $k=3$, the direct calculation is still simple, and gives

$$\begin{aligned}\chi(\mathcal{M}^0_{G,3}) &= \frac{1}{3!}\left\{(2-2G)^3-3\left[(2-2G)^2-(2-2G)\right]-(2-2G)\right\}\\ &= -\frac{(2G-2)(2G-1)2G}{6}.\end{aligned} \tag{5.11}$$

In this expression, $(2-2G)^3$ is the Euler number of $(M_G)^3$, the term in the brackets subtracts the Euler number of the three subsets in $(M_G)^3$ where exactly two P_i's coincide as elements of M_G, while the last term subtracts the Euler number of the diagonal M_G, which is the set of points where all three coordinates P_1,P_2,P_3 coincide as elements of M_G.

In order to derive the general formula, assume first that we have calculated the Euler number of $\mathcal{M}^0_{G,k}$; the Euler number of $\mathcal{M}^0_{G,k+1}$ is then calculated as follows. We can represent $\tilde{\mathcal{M}}^0_{G,k+1}$ as $M_G\times\tilde{\mathcal{M}}^0_{G,k}$, minus the set of diagonal points. There are exactly k possibilities how the added point can coincide with another point as an element of M_G, and each possibility leads to a subspace of $M_G\times\tilde{\mathcal{M}}^0_{G,k}$ isomorphic to $\tilde{\mathcal{M}}^0_{G,k}$ itself. Since these k copies of $\tilde{\mathcal{M}}^0_{G,k}$ are non-intersecting in $M_G\times\tilde{\mathcal{M}}^0_{G,k}$, we obtain the following recursion relation,

$$\chi(\tilde{\mathcal{M}}^0_{G,k+1}) = \chi(M_G)\,\chi(\tilde{\mathcal{M}}^0_{G,k}) - k\,\chi(\tilde{\mathcal{M}}^0_{G,k}). \tag{5.12}$$

Since $\mathcal{M}^0_{G,k}$ is a factor of $\tilde{\mathcal{M}}^0_{G,k}$ by the free action of the permutation group S_k, the recursion relation (5.12) can be rewritten as

$$(k+1)\chi(\mathcal{M}^0_{G,k+1}) = (2-2G-k)\chi(\mathcal{M}^0_{G,k}). \tag{5.13}$$

This relation can be easily solved, and the general formula for the Euler numbers at arbitrary values of G and k finally is

$$\chi(\mathcal{M}^0_{G,k}) = (-1)^k\begin{pmatrix}2G+k-3\\ k\end{pmatrix}. \tag{5.14}$$

We can summarize these Euler numbers in a generating formula, by introducing an auxiliary variable x and defining $\chi(x)\equiv\sum\chi(\mathcal{M}^0_{G,k})\,x^k$. Using the explicit expressions (5.14) for the

Euler numbers, we can write the generating function $\chi(x)$ in a surprisingly simple form:

$$\chi(x) \equiv \sum_{k=0}^{\infty} \chi(\mathcal{M}^0_{G,k})\, x^k = \sum_{k=0}^{\infty} (-1)^k \begin{pmatrix} 2G+k-3 \\ k \end{pmatrix} x^k = \frac{1}{(1+x)^{2G-2}}. \qquad (5.15)$$

This formula will prove very valuable in §6, where we compare the partition functions of the topological rigid string theory with the results of the large-N expansion in 2D QCD.

5.4 Partition Functions of the Topological Rigid String

The partition function of the topological rigid string theory on a fixed target M_G of genus G contains a contribution from various components of the moduli spaces of minimal area maps as analyzed in the previous subsections. We can impose the homotopy constraint (5.3) in the form of a delta function, which allows us to write the partition function as an unrestricted sum over all possible moduli as well as values of the homotopies a_j, b_j,

$$\mathcal{Z} = \sum_{n,\tilde{n}} g_{\rm string}^{(n+\tilde{n})(2G-2)} \frac{1}{n!\tilde{n}!} \sum_{k=0}^{\infty} \sum_{a_j,b_j} \zeta_{G,k}\, \delta\left(\sigma_{n,\tilde{n}}(P_1)\ldots\sigma_{n,\tilde{n}}(P_k) \prod_{j=1}^{G} a_j b_j a_j^{-1} b_j^{-1}\right). \qquad (5.16)$$

Here $\sigma_{n,\tilde{n}}(P_i)$ is a sum over all possible moduli in P_i, each modulus being represented by its element of $S_n \otimes S_{\tilde{n}}$ and weighted by its contribution to the overall power of $g_{\rm string}$, and $\zeta_{G,k}$ are numbers that implicitly depend on all the data that are being summed over. The intergration over the moduli spaces and over the integrable zero modes of the ghosts and antighosts in the topological rigid string theory gives the Euler number of the moduli spaces, and we expect

$$\zeta_{G,k} \propto \chi(\mathcal{M}^0_{G,k}). \qquad (5.17)$$

Before we write an explicit expression for $\sigma_{n,\tilde{n}}(P_i)$, let us analyze this expectation in detail. Although essentially true, the naive statement (5.17) receives corrections from the integration over the remaining modes in the path integral.

First we show that collapsed disks do not contribute to the partition function, since the corresponding moduli spaces have Euler number zero. Consider an arbitrary fixed configuration of moduli other than collapsed disks, in points $(P_1,\ldots,P_k)$ in the target. Adding s collapsed disks in additional points $(P'_{k+1},\ldots,P'_{k+s})$ does not change the genus of the worldsheet. When two collapsed disks at P'_{k+i}, P'_{k+j} coalesce, they again form a collapsed disk. In this sense, the moduli space of maps with s collapsed disks is a compactification locus of the moduli space of maps with $s+1$ collapsed disks. Hence, for fixed moduli at $(P_1,\ldots,P_k)$, there are two disconnected components of the moduli spaces: one corresponds to maps with no collapsed disks, and consists of just one point for each set of moduli at fixed values of $(P_1, \ldots, P_k)$; the other corresponds to maps with an arbitrary number of collapsed disks. This second component of the moduli space is nominally infinitely dimensional, and all moduli spaces with finite s are nested in it. As $s \to \infty$, the Euler number of this moduli space goes to zero, and the only contribution to the partition function thus comes from maps with no collapsed disks.

One reason why the Euler numbers of the moduli spaces are not the whole story comes from the existence of minimal-area maps with additional zero modes of the ghost-for-ghost ϕ^a. Path integrals in these sectors are notoriously hard to calculate since the zero modes of ϕ^a make the moduli space of all zero modes non-compact, but general arguments exist that these contributions typically vanish. Even in the topological rigid string, this is a subtle issue, and its full clarification would go well beyond the scope of this paper. To see which classes of maps lead to additional zero modes of ϕ^a, we will use a BRST fixed point theorem.

BRST fixed-point theorems (see e.g. [38]) use the BRST invariance of the theory to argue that the only non-zero contribution to the path integral comes from infinitesimal vicinity of the set of configurations annihilated by the BRST charge (i.e. are "fixed points" of the BRST supersymmetry transformation). Recall first the action of the equivariant BRST charge on the fields of the topological rigid string,

$$[Q, x^\mu] = \psi^\mu, \qquad \{Q, \psi^\mu\} = \phi^a \partial_a x^\mu. \tag{5.18}$$

Whenever a two-dimensional domain $\mathcal{D}$ in Σ is mapped to one point in the target (e.g. when the map has a collapsed handle or a collapsed disk), the two-by-two matrix $\partial_a x^\mu$ is degenerate everywhere in $\mathcal{D}$. The condition that the BRST transformation of all fields be zero entails

$$\psi_0^\mu = 0, \qquad \phi_0^a \partial_a x_0^\mu = 0, \tag{5.19}$$

and a degenerate matrix $\partial_a x_0^\mu$ leads to many non-trivial solutions for ϕ_0^a.

Additional zero modes of ϕ^a also exist for some maps with composite moduli, i.e. in some components of the compactification locus of the moduli spaces. The simplest modulus leading to additional zero modes is generated when two identical simple branchpoints coalesce in P and form a connecting tube between two sheets of the same orientation. More complicated composite moduli of this type (i.e. twisted connecting tubes) are created when multiple branchpoints coalesce. Maps with these composite moduli map a homotopically non-trivial loop on Σ to the target point P. Along this loop, the two-by-two matrix $\partial_a x_0^\mu$ degenerates and non-trivial solutions ϕ_0^a of (5.19) exist.

Even on a regular component of the moduli spaces, the partition function is not necessarily equal to the Euler number, and can in fact differ from $\chi(\mathcal{M})$ by a sign. Indeed, while for minimal-area maps the fermionic and bosonic one-loop determinants cancel each other exactly, an extra minus sign can appear for unstable critical-area map. A generic unstable critical-area map contains at least one neck that connects two covering sheets of opposite orientations; such a map can be deformed to a map with lower area by opening the neck into a connecting tube of non-zero radius, with a fold of non-zero length. This deformation corresponds to a negative eigenvalue of the fermionic operator (4.33), which modifies the contribution of these moduli spaces to

$$(-1)^v \chi(\mathcal{M}^0_{G,k}), \tag{5.20}$$

where v is the number of simple connecting necks among the moduli at $P_1, \ldots, P_k$.

The contribution of the regular moduli spaces to the partition function can thus be summarized in the following expression for $\sigma_{n,\tilde{n}}(P)$,

$$\begin{aligned}\sigma_{n,\tilde{n}}(P) = &\sum_{\sigma\otimes\tau\in S_n\otimes S_{\tilde{n}}} \sigma\otimes\tau\, g_{\text{string}}^{n+\tilde{n}-K_\sigma-K_\tau} \\ &\times \prod_{\ell=1}^{\min(n,\tilde{n})} \left(\sum_{v_\ell=0}^{\min(\sigma_{(\ell)},\tau_{(\ell)})} (-1)^{v_\ell} \ell^{v_\ell} v_\ell! \begin{pmatrix} \sigma_{(\ell)} \\ v_\ell \end{pmatrix} \begin{pmatrix} \tau_{(\ell)} \\ v_\ell \end{pmatrix} g_{\text{string}}^{2v_\ell} \right).\end{aligned} \tag{5.21}$$

Here $\sigma_{(\ell)}$ and $\tau_{(\ell)}$ is the number of cycles of length ℓ in σ and τ, K_σ is the number of all cycles in σ, and the sum is restricted to σ and τ that are not simultaneously trivial.

The combinatorial factors are present in (5.21) to ensure that each component of the moduli spaces that contributes to (5.21) does so exactly once. More explicitly, the geometry of the factors is as follows. ℓ denotes the number of sheets of the orientation-preserving and orientation-reversing cover that are connected by an orientation-reversing collapsed neck; this neck can be visualized as the simple orientation-reversing connecting neck, twisted ℓ times. In the sum over all values of ℓ, v_ℓ is the number of orientation-reversing collapsed necks that connect a cycle with ℓ sheets of a given orientation with a cycle of ℓ sheets with the opposite orientation, the binomial coefficients count the number of combinations in which the orientation-preserving and orientation-reversing cycles of length ℓ can be combined to form v_ℓ connecting necks, $v_\ell!$ represents all possible permutations of the combinations, and the additional power ℓ^{v_ℓ} comes from the fact that a given cycle of length ℓ can be combined with a given cycle of the same length and opposite orientation in ℓ different ways. The power of g_{string} in the sum over all ℓ just weighs the contribution of the necks to the overall Euler number of the worldsheet.

Since in (5.21) we have already absorbed into $\sigma_{n,\tilde{n}}(P)$ the additional minus signs $(-1)^v$ that come from the contribution of the negative modes to the one-loop determinant, each component of the moduli spaces that contributes in (5.21) contributes exactly $\chi(\mathcal{M}^0_{G,k})$.

Hence, we can write our final formula for the partition functions of the topological rigid string as follows,

$$\mathcal{Z} = \sum_{n,\tilde{n}} \frac{1}{n!\tilde{n}!} g_{\text{string}}^{(n+\tilde{n})(2G-2)} \sum_{k=0}^{\infty} \sum_{a_1,b_1\dots a_G,b_G} \chi(\mathcal{M}^0_{G,k})\, \delta\left(\prod_{i=1}^{k} \sigma_{n,\tilde{n}}(P_i) \prod_{j=1}^{G} a_j b_j a_j^{-1} b_j^{-1}\right). \quad (5.22)$$

This formula can be rewritten in terms of an integral over the total moduli space,

$$\mathcal{Z} = \int_{\mathcal{M}} \mathrm{e}^{-\mathcal{R}_0(\mathcal{M})}, \quad (5.23)$$

where $\mathcal{R}_0(\mathcal{M})$ is a suitably defined two-form on the total moduli space $\mathcal{M}$. This two-form is the sum of the induced curvature two-forms along the directions that contribute to the partition functions, multiplied by an additional minus sign when the modulus is a connecting neck.

6. Large-N QCD Strings in Two Dimensions

We have seen in the previous sections how the topological rigid string theory is defined, and that its path integral gives the Euler numbers of moduli spaces of minimal-area maps. Here we show how this reproduces the results of the large-N expansion in 2D QCD.

The large-N expansion in the two-dimensional Yang-Mills theory with gauge group SU(N) can be written as

$$\begin{aligned}\mathcal{Z}(G,\lambda A,N) = \sum_{n,\tilde{n}} \frac{1}{n!\tilde{n}!}\, \mathrm{e}^{-(n+\tilde{n})\lambda A/2} \sum_{s,k} (-1)^s \frac{(\lambda A)^{s+k}}{s!k!} N^{(n+\tilde{n})(2-2G)-s-2k} \frac{(n-\tilde{n})^{2k}}{2^k} \\ \times \sum_{p_1,\dots p_s\in T_2} \sum_{a_1,b_1,\dots a_G,b_G} \delta(p_1\dots p_s\, \Omega^{2-2G}_{n,\tilde{n}} \prod_{j=1}^{G} a_j b_j a_j^{-1} b_j^{-1})\end{aligned} \quad (6.1)$$

Here the "Ω-points" correspond to (in the notation of [13])

$$\Omega_{n,\tilde{n}} = 1 + \tilde{\Omega}_{n,\tilde{n}} = \sum_{\sigma,\tau} \sigma\otimes\tau \sum_{v,v'} (-1)^{K_v} C_v N^{K_{\sigma\backslash v}+K_{\tau\backslash v}-n-\tilde{n}}. \quad (6.2)$$

6.1 Zero Target Area/Zero String Tension

At $\lambda = 0$, the partition function can be written as a sum over the number of orientation-preserving and orientation reversing sheets n and $\tilde{n}$, each contribution being equal to

$$\mathcal{Z}_{n,\tilde{n}}(G, \lambda A = 0, N) = \frac{1}{n!\tilde{n}!} N^{(n+\tilde{n})(2-2G)} \sum_{a_1,b_1,\ldots a_G,b_G} \delta(\Omega_{n,\tilde{n}}^{2-2G} \prod_{j=1}^{G} a_j b_j a_j^{-1} b_j^{-1}) \tag{6.3}$$

Since $\Omega_{n,\tilde{n}} = 1 + \tilde{\Omega}_{n,\tilde{n}}$, we can expand $\Omega_{n,\tilde{n}}$ in the powers of $\tilde{\Omega}_{n,\tilde{n}}$, and write (6.3) as

$$\begin{aligned} \mathcal{Z}_{n,\tilde{n}}(G, \lambda A = 0, N) &= \frac{1}{n!\tilde{n}!} N^{(n+\tilde{n})(2-2G)} \\ &\quad \times \sum_{k=0}^{\infty} (-1)^k \binom{2G+k-3}{k} \sum_{a_1,b_1,\ldots a_G,b_G} \delta(\tilde{\Omega}_{n,\tilde{n}}^{k} \prod_{j=1}^{G} a_j b_j a_j^{-1} b_j^{-1}) \end{aligned} \tag{6.4}$$

Recall the generating function of the Euler numbers of the moduli spaces of minimal area maps, (5.15). Hence, the binomial factor in (6.4) is essentially the Euler number of the moduli space of maps that are covers of M_G except in k points, where $\tilde{\Omega}_{n,\tilde{n}}$ is inserted. The expression (6.2) for $\tilde{\Omega}_{n,\tilde{n}}$ is not very transparent, but luckily, it can be rewritten as [13]

$$\Omega_{n,\tilde{n}} = \sum_{\sigma,\tau} \sigma \otimes \tau \, N^{-n-\tilde{n}+K_\sigma+K_\tau} \prod_{\ell=1}^{\min(n,\tilde{n})} \left(\sum_{v_\ell=0}^{\min(\sigma_{(\ell)},\tau_{(\ell)})} (-1)^{v_\ell} \ell^{v_\ell} v_\ell! \binom{\sigma_{(\ell)}}{v_\ell} \binom{\tau_{(\ell)}}{v_\ell} \frac{1}{N^{2v_\ell}} \right). \tag{6.5}$$

In this expression we recognize our expression for $\sigma_{n,\tilde{n}}(P)$ that summarizes the contributions of different components of the moduli spaces of minimal-area maps to the partition function of the topological rigid string. Because of the remarkable generating formula for the Euler numbers of the moduli spaces, (5.15), we can also identify the combinatorial factors in front of the sum over all homotopies a_j, b_j in (6.4) as the Euler numbers of the moduli spaces. Hence, the partition function of the large-N QCD in two dimensions at $\lambda = 0$, as given by (6.4), is equal to the partition function of the topological rigid string as summarized by eqn. (5.22), assuming we set $g_{\text{string}} = 1/N$. This is one of the central results of this paper.

6.2 Non-Zero Target Area/Non-Zero String Tension

At non-zero area/non-zero couping constant, the full results of [13] can also be interpreted in simple geometrical terms. For SU(N) Yang-Mills theory, the large-N expansion gives

$$\mathcal{Z}(G,\lambda A,N) = \sum_{n,\tilde{n}} \frac{1}{n!\tilde{n}!} \mathrm{e}^{-(n+\tilde{n})\lambda A/2} \sum_{s,t} (-1)^s \frac{(\lambda A)^{s+t}}{s!t!} N^{(n+\tilde{n})(2-2G)-s-2t} \frac{(n-\tilde{n})^{2t}}{2^t}$$
$$\times \sum_{k=0}^{\infty} (-1)^k \begin{pmatrix} 2G+k-3 \\ k \end{pmatrix} \sum_{p_1,\ldots p_s \in T_2} \sum_{a_1,b_1,\ldots a_G,b_G} \delta(p_1 \ldots p_s \tilde{\Omega}^k_{n,\tilde{n}} \prod_{j=1}^{G} a_j b_j a_j^{-1} b_j^{-1}). \tag{6.6}$$

Following [13], we decompose $(n-\tilde{n})^2$ into $-2n\tilde{n}+[n(n-1)+\tilde{n}(\tilde{n}-1)]+[n+\tilde{n}]$, and interpret the terms as coming from simple orientation-reversing connecting necks, simple orientation-preserving connecting tubes, and collapsed handles respectively. s in (6.6) counts the number of simple branchpoints that contribute to the area dependence of the partition function.

We have already interpreted the $\lambda=0$ part of the partition function as a calculation of the Euler number of certain moduli spaces of minimal-area maps, and can be consequently written as an integral over the moduli space of the exponential of a specific two-form, cf. (5.23). It can be straightforwardly shown that the whole partition function at non-zero λ can also be written as an integral of a specific form over the same moduli space. In order to get a better insight into the situation, consider first the following integral

$$\int_{\mathcal{M}_{G,k}} \mathrm{e}^{-\mathcal{R}-\lambda\mathcal{W}} \tag{6.7}$$

on the moduli spaces $\mathcal{M}_{G,k} \equiv \tilde{\mathcal{M}}_{G,k}/S_k$, with $\tilde{\mathcal{M}}_{G,k} \equiv (M_G)^k$. Here we have defined two-forms $\mathcal{R}$ and $\mathcal{W}$ on the moduli spaces by

$$\mathcal{R} = \sum_{i=1}^{k} \mathcal{R}_i, \qquad \mathcal{W} = \sum_{i=1}^{k} \mathcal{W}_i, \tag{6.8}$$

where $\mathcal{R}_i$ and $\mathcal{W}_i$ are the curvature two-form and the volume two-form on the i-th copy of M_G. On $\tilde{\mathcal{M}}_{G,k}$, the integral gives

$$\int_{\tilde{\mathcal{M}}_{G,k}} \mathrm{e}^{-\mathcal{R}-\lambda\mathcal{W}} = \frac{(-1)^k}{k!} \int_{\tilde{\mathcal{M}}_{G,k}} (\mathcal{R}+\lambda\mathcal{W})^k = \frac{(-1)^k}{k!} \sum_{s=0}^{k} \begin{pmatrix} k \\ s \end{pmatrix} \lambda^s \int_{\tilde{\mathcal{M}}_{G,k}} \mathcal{W}^s \wedge \mathcal{R}^{k-s}. \tag{6.9}$$

Since $\tilde{\mathcal{M}}_{G,k}$ is a direct product of k copies of M_G, this expression can be further reduced to

$$\begin{aligned}
\int\limits_{\tilde{\mathcal{M}}_{G,k}} \mathrm{e}^{-\mathcal{R}-\lambda\mathcal{W}} &= \frac{(-1)^k}{k!}\sum_{s=0}^{k}\begin{pmatrix} k \\ s \end{pmatrix}\lambda^s \int\limits_{\tilde{\mathcal{M}}_{G,k}} \sum_{i_1\ldots i_s}\mathcal{W}_{i_1}\wedge\ldots\mathcal{W}_{i_s}\wedge\sum_{j_1\ldots j_{k-s}}\mathcal{R}_{j_1}\wedge\ldots\mathcal{R}_{j_{k-s}}\\
&= \frac{(-1)^k}{k!}\sum_{s=0}^{k}\begin{pmatrix} k \\ s \end{pmatrix}\lambda^s\, k! \int\limits_{\tilde{\mathcal{M}}_{G,k}} \mathcal{W}_1\wedge\ldots\wedge\mathcal{W}_s\wedge\mathcal{R}_{s+1}\wedge\ldots\wedge\mathcal{R}_k\\
&= (-1)^k\sum_{s=0}^{k}\begin{pmatrix} k \\ s \end{pmatrix}\lambda^s \int\limits_{\tilde{\mathcal{M}}_{G,s}} \mathcal{W}_1\wedge\ldots\wedge\mathcal{W}_s \int\limits_{\tilde{\mathcal{M}}_{G,k-s}} \mathcal{R}_1\wedge\ldots\wedge\mathcal{R}_{k-s}\\
&= (-1)^k\sum_{s=0}^{k}\begin{pmatrix} k \\ s \end{pmatrix}\lambda^s\,\mathrm{Vol}\,(\tilde{\mathcal{M}}_{G,s})\;\chi(\tilde{\mathcal{M}}_{G,k-s}).
\end{aligned} \tag{6.10}$$

Our moduli spaces $\mathcal{M}_{G,k}$ are factors of $\tilde{\mathcal{M}}_{G,k}$ by S_k, and an analogous evaluation of the integral leads to

$$\int\limits_{\mathcal{M}_{G,k}} \mathrm{e}^{-\mathcal{R}-\lambda\mathcal{W}} = (-1)^k\sum_{s=0}^{k}\lambda^s\,\mathrm{Vol}\,(\mathcal{M}_{G,s})\;\chi(\mathcal{M}_{G,k-s}). \tag{6.11}$$

This is exactly the area dependence encountered in (6.6), since $\mathrm{Vol}\,(\mathcal{M}_{G,s}) = A^s/s!$.

To facilitate our further discussion, it is useful to place the theory with non-zero λ into a wider context. It is indeed well known that in two dimensions, one can deform the Yang-Mills Lagrangian by an infinite number of new terms, and write

$$\mathcal{L}' = \int\limits_M \phi F + \int\limits_M \mathrm{d}^2x\sqrt{g}\; f(\phi), \tag{6.12}$$

where $f(\phi)$ is an arbitrary class function on the Lie algebra of the Yang-Mills gauge group $\mathcal{G}$, and can be written as a sum over the infinite number of Casimir operators of $\mathcal{G}$. In the context of the topological interpretation of the Yang-Mills theory as summarized in §1.2, these new terms correspond to the higher BRST cohomology classes expressed in terms of the ghost-for-ghost field ϕ. We can still interpret the partition function of the deformed Yang-Mills theory (6.12) as a correlation function of the BRST cohomology classes in the underlying topological Yang-Mills theory (cf. (1.10)):

$$\langle 1\rangle_{f(\phi)} = \left\langle \exp\left\{-\int\limits_M(\phi F - \psi\wedge\psi) - \int\limits_M\sqrt{g}\; f(\phi)\right\}\right\rangle_{\text{topo. YM}} \tag{6.13}$$

At large N, all the new terms in the Yang-Mills Lagrangian have a corresponding string interpretation, discussed in [39].

In the topological rigid string theory, these deformed partition functions can be reconstructed as follows. The moduli spaces of minimal-area maps, as discussed in the analysis of the topological rigid string theory in §5, carry a natural cohomological structure. In particular, they carry an infinite number of natural cohomology classes, which are the analogy of the "stable" or "universal" cohomology classes that can be naturally defined on moduli spaces of other moduli problems studied in the literature [40]. Recall that we have parametrized our moduli spaces by the target locations $P_1, \ldots, P_s$ of the moduli of minimal-area maps. The fixed metric on the target manifold M induces a natural induced-area two-form $\mathcal{W}(P_i)$ on $\mathcal{M}$ for each P_i, the only non-zero components of $\mathcal{W}(P_i)$ being along the direction of P. The natural cohomology classes are then generated by specific two-forms on $\mathcal{M}$, which are in one-to-one correspondence with the elements of the set of all conjugacy classes of all possible moduli ("modulus types" from now on). Given a fixed modulus type α, for example a branchpoint of degree p, define a two-form $\mathcal{O}_\alpha$ on $\mathcal{M}$ as the sum of the induced-area two-forms $\mathcal{W}(P_i)$ where i runs over all moduli in the conjugacy class of α.

In addition to these natural cohomology classes $\mathcal{O}_\alpha$ on $\mathcal{M}$ which are all two-forms, another cohomology class is needed to establish a relation with the partition functions of the generalized Yang-Mills theory at large N. This class is a zero-form on $\mathcal{M}$, whose value in each point of the moduli space is equal to the induced area of the worldsheet of the corresponding minimal-area map. We will denote this cohomology class by $\mathcal{O}_0$.

By analogy with the Yang-Mills formula (6.13), we now claim that the large-N expansion of the partiton functions in the generalized Yang-Mills theory can be written as integrals of the exponential of a linear combination of the natural cohomology classes, combined with the density on the moduli spaces that is already present in the theory at $f(\phi) \equiv 0$,

$$\mathcal{Z}(G, f(\phi), N) = \int_{\mathcal{M}} \mathrm{e}^{-\mathcal{R}_0(\mathcal{M}) - c_0\mathcal{O}_0 - \sum c_\alpha \mathcal{O}_\alpha} \tag{6.14}$$

More precisely, a natural map Υ exists that associates with each choice of $f(\phi)$ in the generalized Yang-Mills theory a linear combination of the natural cohomology classes $\mathcal{O}_0, \mathcal{O}_\alpha$ of the moduli spaces, such that formula (6.14) is valid. The results of [39] can be considered a direct verification of this statement. For a given $f(\phi)$, the specific coefficients c_0, c_α of $\Upsilon(f(\phi)) \equiv c_0\mathcal{O}_0 + c_\alpha\mathcal{O}_\alpha$ can be directly inferred from [39].

Hence, in the string representation, any coupling constant of the generalized Yang-Mills theory multiplies a linear combination of the natural cohomology classes of the moduli spaces, and the large-N expansion of the partition function of the generalized Yang-Mills theory can be written as a correlation function of the corresponding cohomology classes in the topological rigid string. In particular, the λ term of the standard Yang-Mills Lagrangian is interpreted in the topological rigid string theory as a specific linear combination of the natural cohomology classes of the moduli spaces,

$$\Upsilon(\lambda\phi^2) = \lambda\left(\mathcal{O}_0 + 2\mathcal{O}_b + 2\mathcal{O}_r - 2\mathcal{O}_p - \mathcal{O}_h\right). \tag{6.15}$$

Here $\mathcal{O}_0$ is the zero-form that has been defined above, while the remaining contributions come from the natural two-forms $\mathcal{O}_\alpha$: $\mathcal{O}_b$ is the natural two-form that corresponds to the conjugacy class of a simple branchpoint, $\mathcal{O}_r$ corresponds to the simple orientation-reversing collapsed neck, $\mathcal{O}_p$ to the orientation-preserving collapsed tube, and $\mathcal{O}_h$ to the collapsed handle. The last three contributions would be missing if we change the gauge group from $\mathrm{SU}(N)$ to $\mathrm{U}(N)$.

Several remarks are in order:

1. Even though in the $\mathrm{U}(N)$ theory the moduli that correspond to collapsed handles do not contribute to either the partition function at zero λ or to the theory where λ is the only non-zero coupling in $f(\phi)$, collapsed handles do emerge when we consider higher Casimirs in $f(\phi)$ (see [39]). Hence, collapsed handles are not specifics of the $\mathrm{SU}(N)$ theory, and we cannot get rid of them by restriction to $\mathrm{U}(N)$. It is an advantage of our formulation of the QCD string theory over possible alternative formulations in terms of holomorphic maps that collapsed handles emerge as simple moduli from the outset.
2. In the topological rigid string theory, the difference between $\mathrm{SU}(N)$ and $\mathrm{U}(N)$ with only λ non-zero seems to be a matter of choice of the specific combination of the cohomology classes in (6.14), and neither choice seems to be particularly singled out.

An explicit field representation of the universal cohomology classes in the topological rigid string would require a detailed analysis of the equivariant BRST cohomology of the topological rigid string and is not an easy task, but we can at least discuss the simplest observables that emerge in the theory at non-zero λ.

Notice first that if we integrate out the auxiliary field B^μ in the topological rigid string theory defined by the Lagrangian (4.24), the on-shell BRST algebra is given (in the specific gauge (4.23)) by

$$\begin{aligned}
&[Q, x^\mu] = \psi^\mu, \qquad \{Q, \psi^\mu\} = 0,\\
&\{Q, \chi^\mu\} = -\frac{1}{2a}(1 + a\psi\chi)\Delta x^\mu + \frac{1}{2}\chi^\mu(\nabla\psi\cdot\partial x) - \frac{1}{2}\psi^\mu(\nabla\chi\cdot\partial x)\\
&\qquad\qquad + \left(\delta^\mu_\nu - \partial_a x^\mu h^{ab}\partial_b x^\lambda g_{\lambda\nu}\right)\Gamma^\nu_{\sigma\rho}\,\chi^\sigma\psi^\rho.
\end{aligned} \tag{6.16}$$

Define now

$$\delta\mathcal{L} = \lambda\int_\Sigma \mathrm{d}^2\sigma\,\sqrt{h}\,\left\{1 + 2a\psi * \chi\right\}. \tag{6.17}$$

In two dimensions, one can demonstrate by a direct calculation that $\delta\mathcal{L}$ is on-shell BRST invariant up to gauge transformation of the worldsheet gauge symmetries, i.e.

$$[Q, \delta\mathcal{L}] = \text{pure gauge}. \tag{6.18}$$

This is all we need to be able to evaluate the deformed partition function,

$$\mathcal{Z}(\lambda) = \int \mathrm{e}^{-\mathcal{L}-\delta\mathcal{L}}, \tag{6.19}$$

perturbatively in λ.

The first term in $\delta\mathcal{L}$ is just the bosonic Nambu-Goto Lagrangian (i.e. the induced area term), and λ is the string tension. The fermionic term in $\delta\mathcal{L}$ is an improvement that makes $\delta\mathcal{L}$ an admissible deformation from the point of view of the topological BRST symmetry.

We claim that $\delta\mathcal{L}$ is a field-theoretical representation of the cohomology class $\mathcal{O}_0$ defined above. To show this, we evaluate the deformed partition function (6.19) perturbatively in λ. Because of the nice BRST properties of the Lagrangian, we can evaluate the deformed path integral semiclassically. The part with no fermions is the bosonic Nambu-Goto action and measures the total induced area of the worldsheet, which on the moduli spaces of minimal area maps reduces to

$$\lambda \int_\Sigma \mathrm{d}^2\sigma\, \sqrt{h} \ = \lambda(n+\tilde{n})A, \tag{6.20}$$

with n and $\tilde{n}$ being the number of sheets in the orientation preserving and orientation reversing sectors respectively. λ is indeed the string tension. The term linear in quantum fluctuations δx^μ is proportional to the first variation of the Nambu-Goto action, and vanishes on the moduli spaces of minimal-area maps. The terms quadratic in quantum fluctuations could only affect the one-loop determinants in target dimensions higher than two. Hence, the one-loop determinants still cancel each other (up to a possible sign), and $\delta\mathcal{L}$ indeed reduces to the evaluation of the induced area of the worldsheet in the minimal-area map, which is the definition of the cohomology class $\mathcal{O}_0$.

In the topological rigid string theory defined by the deformed Lagrangian (4.26), it is natural to consider

$$\delta\mathcal{L}' = \lambda \int_\Sigma \mathrm{d}^2\sigma\, \sqrt{h}\ \left\{1 + \psi * \chi + (\psi * \chi)^2\right\}. \tag{6.21}$$

Notice that $\delta\mathcal{L}'$ is related to the Lagrangian $\mathcal{L}$ of the topological rigid string as defined in

(4.26) by a double-commutator formula,

$$\mathcal{L} = \{Q, [\bar{Q}, \int_\Sigma \mathrm{d}^2\sigma \, \frac{\sqrt{h}}{1-\psi * \chi}]\} = \{Q, [\bar{Q}, \frac{1}{\lambda}\delta\mathcal{L}']\}. \tag{6.22}$$

This relation is not accidental, but its more thorough explanation would lead us beyond topological theory [16]. Thus, we will only show that the deforming term $\delta\mathcal{L}'$ is weakly BRST invariant, which also allows us to use it as a deformation of the Lagrangian. On the moduli space of minimal-area maps, where $\psi^\mu = \chi^\mu = 0$ and $\Delta x^\mu = 0$, the Lagrangian $\mathcal{L}$ vanishes, and the double commutator formula (6.22) reduces to

$$\{\bar{Q}, [Q, \delta\mathcal{L}']\} = 0. \tag{6.23}$$

Hence, on the moduli space, the new term in the Lagrangian is BRST invariant up to a $\bar{Q}$-closed term,

$$[Q, \delta\mathcal{L}'] = \Gamma, \qquad \{\bar{Q}, \Gamma\} = 0, \tag{6.24}$$

which again allows us to add $\delta\mathcal{L}'$ to the Lagrangian of the topological rigid string and evaluate the deformed path integral semiclassically. The only difference from the previous calculation comes from the zero-mode integration, where we are left with the four-fermi term $(\psi * \chi)^2$ that accompanies the curvature-dependent four-fermi term of the topological rigid string Lagrangian in the saturation of the zero mode integral in the fermi sector. Effectively, this changes the measure on the moduli spaces, from the Euler measure to a measure that depends on the volume element of the components of the moduli spaces, according to formula (6.14). Consequently, $\delta\mathcal{L}'$ is a field-theoretical representation of a certain linear combination of the cohomology classes $\mathcal{O}_0$ and $\mathcal{O}_\alpha$.

The whole dependence of the partition functions on all coupling constants of the generalized Yang-Mills theory would require a detailed information about the field-theoretical realization of the natural cohomology classes of the moduli spaces of minimal-area maps introduced above, as well as a more detailed technical understanding of some puzzling aspects of the path integral of the theory (in particular, in the degenerate homotopy classes). In general, we should not expect the full structure of observables and correlation functions of the topological rigid string theory to be much simpler that that of conventional topological string theory [40]. Already at this stage, however, we can claim that the simplest deforming term corresponds to the Nambu-Goto induced-area term, improved by fermionic terms in order to become a BRST invariant observable. For generic coupling constants, the topological rigid string interpretation of the large-N partition functions follows from the cohomological formula (6.14).

7. Conclusions

In this paper, we have presented a topological rigid string theory, and discussed it in two dimensions as a theory of QCD strings. Since an extension of the theory to higher dimensions will be discussed elsewhere [16], we limit our conclusions to several remarks on those aspects of the two-dimensional theory that were not discussed in the body of this paper.

(1) Although we focused out attention on partition functions of the topological rigid string and their relation to the large-N expansion of the QCD partition functions, our results can be easily extended to correlation functions of Wilson loops in arbitrary representations of SU(N). It is a straightforward exercise to show that the Wilson loop correlation functions at $\lambda = 0$ calculate the Euler numbers of moduli spaces of minimal-area maps from worldsheets with boundaries to the targets with Wilson loops, while the area dependence emerges from the volume of the moduli spaces, exactly as in the case of the partition functions. The only difference from the calculation of the partition functions is in the slightly more complicated geometry of the moduli spaces of such minimal-area maps.

(2) Instead of SU(N) Yang-Mills theory, one can study the large-N expansion for alternative series of gauge groups, SO(N) and Sp(N). The results of [13] have been extended to these alternative cases in [41], leading to theories of unoriented closed strings (as expected). We will argue now that the corresponding string theories can also be described by a topological rigid string theory, as follows. We have seen in §2.5 that for harmonic topological sigma models, a canonical orbifold theory exists; when extended to the topological rigid string, this orbifold construction makes closed strings unoriented, and introduces open strings (with the standard, Neumann boundary conditions on both ends) as twisted states. We claim that this orbifold theory describes the large-N string theory of QCD with the alternative gauge groups. (The difference between SO(N) and Sp(N) corresponds to the sign choice in the definition of the $\mathbf{Z}_2$ orbifold action on the closed string sector). This conjecture raises an obvious question, as we certainly do not expect open strings in the large-N expansion of QCD Yang-Mills theory without matter, for either gauge group. This apparent paradox has a surprising resolution: The orbifold version of the topological rigid string theory does indeed contain an open string sector, with Neumann boundary conditions of both ends of open strings; all open-string states are however unphysical, since they are all homotopically trivial.

(3) As a next logical step, one can try to couple the string to dynamical quarks (perhaps by a choice of worldsheet boundary conditions that break the worldsheet $\mathcal{N} = 2$ supersymmetry of the topological rigid string down to $\mathcal{N} = 1$ along the boundary), and compare the results to the known properties of the 't Hooft model [7]. The existence of a natural coupling between the strings and quarks would serve as an independent check on the validity of our string theory as a two dimensional QCD string theory.

(4) Worldsheet and spacetime duality. 2D QCD string theory contains essentially two coupling constants: the string tension λ (or more precisely, its dimensionless version $\hat{\lambda} \equiv \lambda A$), and the string coupling constant $g_{\rm string} \equiv 1/N$. Results of [42] suggest that the theory exhibits interesting duality properties under $\hat{\lambda} \to \hat{\lambda}^{-1}$ (at least on the torus). In the spacetime Yang-Mills theory, such a duality would represent a strong-weak coupling duality (S-duality), while in the string theory it would correspond to a T-duality (since λ is a worldsheet coupling constant). We can also speculate about dualities that would mix $\hat{\lambda}$ and the string coupling constant $1/N$. Indeed, dualities of this type (i.e. dualities that interchange the rank of the gauge group with the gauge coupling constant) are not unknown in low-dimensional gauge theories [43]. In the string representation, these dualities would mix a worldsheet coupling constant with the string coupling constant, and would be an example of what has come to be called U-duality [44]. In QCD string theory, a U-duality interchanging $\hat{\lambda}$ and N would map the large-N string theory to a Wilson-like strong-coupling string theory [3], thus leading to two alternative string descriptions of a given Yang-Mills gauge theory.

(5) With a Lagrangian formulation of the QCD string theory, one can write down its corresponding string field theory, following the standard lore [45]. Such a QCD string field theory should be equivalent to the Yang-Mills theory, although not manifestly so. Since the spacetime Yang-Mills theory is completely solved, 2D QCD might represent an excellent opportunity to test concepts of string field theory. Thus, one can study the relatively well-understood string non-perturbative effects in 2D QCD (such as the Douglas-Kazakov phase transition and the finite-N effects [46]) from the point of view of string theory, and even look for their possible worldsheet interpretation (cf. the recent ideas of Green and Polchinski, [47]). It would also be very instructive to see how spacetime gauge invariance and the Yang-Mills field emerge in string field theory. Any progress in that direction would be helpful in the search for a microscopic derivation of the QCD string theory.

Acknowledgements

It is a pleasure to thank M. Douglas, D. Gross, Y. Nambu, A. Polyakov, I. Singer, L. Susskind, W. Taylor and E. Witten for discussions at various stages of this work. I am grateful to the Aspen Center for Physics, where a part of this work was done, for its hospitality. This work has been supported in part by NSF Grant No. PHY90-21984.

REFERENCES

1. For a review, see e.g.: "Dual Theory," ed.: M. Jacob, *Phys. Rep. Reprint Series* **1** (North-Holland, 1974)
   P.H. Frampton, "Dual Resonance Models and Superstrings" (World Scientific, 1986)
2. Y. Nambu, "Duality and Hadrodynamics," Chicago notes (1970), reprinted in: "Strings, Lattice Gauge Theory, High Energy Phenomenology," eds: V. Singh and S.R. Wadia (World Scientific, 1987)
3. K.G. Wilson, "Confinement of Quarks," *Phys. Rev.* **D10** (1974) 2445
4. G. 't Hooft, "A Planar Diagram Theory for Strong Interactions," *Nucl. Phys.* **B72** (1974) 461
5. J. Polchinski, "Strings and QCD?," Austin preprint (June 1992)
   D.J. Gross, "Some New/Old Approaches to QCD," in: "Integrable Quantum Field Theories," eds: L. Bonora et al. (Plenum Press, 1993)
6. N. Seiberg and E. Witten, "Electric-Magnetic Duality, Monopole Condensation, and Confinement in $N = 2$ Supersymmetric Yang-Mills Theory," *Nucl. Phys.* **B426** (1994) 19 (Erratum: *Nucl. Phys.* **B430** (1994) 485); "Monopoles, Duality and Chiral Symmetry Breaking in $N = 2$ Supersymmetric QCD," *Nucl. Phys.* **B431** (1994) 484
   N. Seiberg, "Electric-Magnetic Duality in Supersymmetric Non-Abelian Gauge Theories," hep-th/9411149
7. G. 't Hooft, "A Two-Dimensional Model for Mesons," *Nucl. Phys.* **B75** (1974) 461
   C.G. Callan, N. Coote and D.J. Gross, "Two-Dimensional Yang-Mills Theory: A Model for Quark Confinement," *Phys. Rev.* **D13** (1976) 1649
   M.B. Einhorn, "Confinement, Form Factors, and Deep-Elastic Scattering in Two-Dimensional Quantum Chromodynamics," *Phys. Rev.* **D14** (1976) 3451
8. A.A. Migdal, "Recursion Equations in Gauge Field Theories," *Zh. Eksp. Theor. Fiz.* **69** (1975) 810 [*Sov. Phys. JETP* **42** (1975) 413]
9. E. Witten, "On Quantum Gauge Theories in Two Dimensions," *Commun. Math. Phys.* **141** (1991) 153
   M. Blau and G. Thompson, "Quantum Yang-Mills Theory on Arbitrary Surfaces," *Int. J. Mod. Phys.* **A7** (1992) 3781
10. B.Ye. Rusakov, "Loop Averages and Partition Functions in U(N) Gauge Theory on Two-Dimensional Manifolds," *Mod. Phys. Lett.* **A5** (1990) 693
11. E. Witten,"Two Dimensional Yang-Mills Theory Revisited," *J. Geom. Phys.* **9** (1992) 303
12. D.J. Gross, "Two Dimensional QCD as a String Theory," *Nucl. Phys.* **B400** (1993) 161

13. D.J. Gross and W. Taylor, "Two Dimensional QCD is a String Theory," *Nucl. Phys.* **B400** (1993) 181; "Twists and Wilson Loops in the String Theory of Two Dimensional QCD," *Nucl. Phys.* **B403** (1993) 395

14. V.A. Kazakov and I.K. Kostov, "Non-Linear Strings in Two Dimensional U(∞) Gauge Theory," *Nucl. Phys.* **B176** (1980) 199
V.A. Kazakov, "Wilson Loop Average for an Arbitrary Contour in Two-Dimensional U(N) Gauge Theory," *Nucl. Phys.* **B179** (1981) 283

15. P. Hořava, "Topological Strings and QCD in Two Dimensions," hep-th/9311156, in: "Quantum Field Theory and String Theory," eds: L. Baulieu et al., Proceedings of the Cargèse Workshop on "New Developments in String Theory, Conformal Models, and Topological Field Theory," May 1993 (Plenum Press, 1995)

16. P. Hořava, in preparation

17. S. Cordes, G. Moore and S. Ramgoolam, "Large-N 2D Yang-Mills Theory and Topological String Theory," Yale preprint (February 1994), hep-th/9402107

18. E. Witten, "Topological Sigma Models," *Commun. Math. Phys.* **118** (1988) 601
L. Baulieu and I.M. Singer, "The Topological Sigma Model," *Commun. Math. Phys.* **125** (1989) 227

19. J. Eells and L. Lemaire, "A Report on Harmonic Maps," *Bull. London Math. Soc.* **10** (1978) 1; "Another Report on Harmonic Maps," *Bull. London Math. Soc.* **20** (1988) 385

20. J. Eells and L. Lemaire, "On the Construction of Harmonic and Holomorphic Maps Between Surfaces," *Math. Ann.* **252** (1980) 27

21. F.E. Burstall and J.H. Rawnsley, "Twistor Theory for Riemannian Symmetric Spaces with Applications to Harmonic Maps of Riemann Surfaces," *L. N. Math.* **1424** (Springer, 1990)

22. M. Blau and G. Thompson, "$N = 2$ Topological Gauge Theory, the Euler Characteristic of Moduli Spaces, and the Casson Invariant," *Commun. Math. Phys.* **152** (1993) 41

23. C. Vafa and E. Witten, "A Strong Coupling Test of S-Duality," *Nucl. Phys.* **B431** (1994) 3

24. E. Witten, "Supersymmetry and Morse Theory," *J. Diff. Geom.* **17** (1982) 661
D. Friedan and P. Windey, "Supersymmetric Derivation of the Atiyah-Singer Index and the Chiral Anomaly," *Nucl. Phys.* **B235[FS11]** (1984) 395
L. Alvarez-Gaumé, "Supersymmetry and the Atiyah-Singer Index Theorem," *Commun. Math. Phys.* **90** (1983) 161

25. P. Hořava, "Two Dimensional String Theory and the Topological Torus," *Nucl. Phys.* **B386** (1992) 383; "Spacetime Diffeomorphisms and Topological w_∞ Symmetry in Two Dimensional Topological String Theory," *Nucl. Phys.* **B414** (1994) 485

26. P. Hořava, "Equivariant Topological Sigma Models," *Nucl. Phys.* **B418** (1994) 571

27. E. Witten, "Chern-Simons Gauge Theory as a String Theory," IAS preprint

28. C. Itoi and H. Kubota, "Gauge Invariance Based on the Extrinsic Geometry in the Rigid String," *Z. Phys.* **C44** (1989) 337; "BRST Quantization of the String Model with Extrinsic Curvature," *Phys. Lett.* **B202** (1988) 381

29. V. Mathai and D. Quillen, "Superconnections, Thom Classes, and Equivariant Differential Forms," *Topology* **25** (1986) 85

30. M.F. Atiyah and L. Jeffrey, "Topological Lagrangians and Cohomology,"*J. Geom. Phys.* **7** (1990) 119

31. N. Berline, E. Getzler and M. Vergne, "Heat Kernels and Dirac Operators" (Springer, 1992)

32. M. Blau, "The Mathai-Quillen Formalism and Topological Field Theory," *J. Geom. Phys.* **11** (1993) 95

33. A.M. Polyakov, "Fine Structure of Strings," *Nucl. Phys.* **B268** (1986) 406; "Gauge Fields and Strings" (Harwood Academic Publishers, 1987)

34. H. Kleinert, "The Membrane Properties of Condensing Strings," *Phys. Lett.* **B174** (1986) 335

35. R.D. Pisarski, "Perturbative Stability of Smooth Strings," *Phys. Rev. Lett.* **58** (1987) 1300; "Smooth Strings at Large Dimension," *Phys. Rev.* **D38** (1988) 578; "Heavy and Smooth Strings in QCD," in: "String Theory – Quantum Cosmology and Quantum Gravity; Integrable and Conformal Invariant Theories," eds: H.J. de Vega and N. Sánchez (World Scientific, 1987)
T.L. Curtright, G.I. Ghandour, C.B. Thorn and C.K. Zachos, "Trajectories of Strings with Rigidity," *Phys. Rev. Lett.* **57** (1986) 799
T.L. Curtright, G.I. Ghandour and C.K. Zachos, "Classical Dynamics of Strings with Rigidity," *Phys. Rev.* **D34** (1986) 3811
E. Braaten and C.K. Zachos, "Instability of the Static Solution to the Closed String with Rigidity," *Phys. Rev.* **D35** (1987) 1512
E. Braaten, R.D. Pisarski and S.-M. Tse, "Static Potential for Smooth Strings," *Phys. Rev. Lett.* **58** (1987) 93
P. Olesen and S.-K. Yang, "Static Potential in a String Model with Extrinsic Curvatures," *Nucl. Phys.* **B283** (1987) 73
F. Alonso and D. Espriu, "On the Fine Structure of Strings," *Nucl. Phys.* **B283** (1987) 393

36. T.T. Dao and A.T. Fomenko, "Minimal Surfaces, Stratified Multivarifolds, and the Plateau Problem," AMS Translations Vol. **84** (Providence, 1991)

37. W. Thurston, "The Geometry and Topology of Three-Manifolds ," Ch. 13, Princeton notes, 1977 (unpublished)

38. E. Witten, "Mirror Manifolds and Topological Field Theory," in: "Essays on Mirror Symmetry," ed.: S.-T. Yau (International Press, 1992)

39. O. Ganor, J. Sonnenschein and S. Yankielowicz, "The String Theory Approach to Generalized 2D Yang-Mills Theory," Tel Aviv preprint TAUP-2182-94 (July 1994)

40. E. Witten, "On the Structure of the Topological Phase of Two-Dimensional Gravity," *Nucl. Phys.* **B340** (1990) 281; "Two Dimensional Gravity and Intersection Theory on Moduli Space," *Surv. Diff. Geom.* **1** (1991) 243

41. S.C. Naculich, H.A. Riggs and H.J. Schnitzer, "Two-Dimensional Yang-Mills Theories are String Theories," *Mod. Phys. Lett.* **A8** (1993) 2223
S. Ramgoolam, "Comment on Two Dimensional O(N) and Sp(N) Yang-Mills Theories as String Theories," Yale preprint (1993)

42. M. Douglas, "Conformal Field Theory Techniques for Large N Group Theory," hep-th/9303159
R. Rudd, The String Partition Function for QCD on the Torus," hep-th/9407176

43. S.C. Naculich, H.A. Riggs and H.J. Schnitzer, "Group-Level Duality in WZW Models and Chern-Simons Theory," *Phys. Lett.* **B246** (1990) 417

44. C.N. Hull and P. Townsend, "Unity of String Dualities," QMW preprint (October 1994), hep-th/9410167
E. Witten, "String Theory Dynamics in Various Dimensions," IAS preprint IASSNS-HEP-95-18 (March 1995), hep-th/9503124

45. B. Zwiebach, "Closed String Field Theory: Quantum Action and the B-V Master Equation, *Nucl. Phys.* **B390** (1993) 33

46. J. Baez and W. Taylor, "Strings and Two-Dimensional QCD for Finite N," MIT preprint (January 1994), hep-th/9401041

47. M.B. Green and J. Polchinski, "Summing over Worldsheet Boundaries," Santa Barbara preprint (June 1994), hep-th/9406012
J. Polchinski, "Combinatorics of Boundaries in String Theory," Santa barbara preprint (July 1994), hep-th/9407031
M.B. Green, "A Gas of D-Instantons," CERN preprint CERN-TH/95-78 (March 1995)