The Schrödinger Representation and 3d Gauge Theories

V.P. Nair

Physics Department, City College New York, CUNY
New York, NY 10031

E-mail: vpnair@ccny.cuny.edu

Abstract

In these notes, we will consider the Hamiltonian analysis of Yang-Mills theory and some variants of it in three dimensions using the Schrödinger representation. This representation, although technically more involved than the usual covariant formulation, may be better suited for some nonperturbative issues. I will review string tension, Casimir energy, and comment on a few other questions.

(To be revised)

Contents

1 Introduction **3**

2 The gauge principle **8**

3 Confinement **18**

4 *Parametrization of gauge fields **22**

5 *The volume element for the gauge-orbit space **27**

6 The topology and geometry of $\mathcal{C}$ **33**

7 *The Hamiltonian **37**

8 *A propagator mass for the gluon **43**

9 *The Schrödinger equation: An expansion scheme **45**

9.1 The lowest order solution 47

9.2 The first order corrections to the vacuum wave functional 48

9.3 Another route to the vacuum wave functional 49

10 *Analytic results, comparison with numerics **50**

10.1 String tension 50

10.2 Comparison of string tension with numerical estimates 52

10.3 Comments regarding string tension 53

10.4 Casimir effect: Calculation 57

10.5 Casimir effect: Comparison with lattice data 59

10.6 Propagator mass: Alternate approaches 61

11 Screenable representations and string breaking **64**

11.1 A Schrödinger equation for the glue lump 65

11.2 Comparison with lattice simulations 69

12 Alternate candidates for the vacuum wave functional **71**

13 Extensions of the Yang-Mills theory and some comments **75**

13.1 An intuitive argument for the mass gap 75

13.2 Yang-Mills-Chern-Simons theory 77

13.3 *Supersymmetric theories 82

14 Entanglement in Yang-Mills (2+1) **87**

14.1 Entanglement in Maxwell theory 87

14.2 The case of YM(2+1) 91

15 Appendix A: Conventions and Notations **93**

16 Appendix B: Regularization **94**

16.1 The regularized form of the operators 94

16.2 A Lorentz-invariance argument 98

17 Appendix C: Corrections to string tension **99**

1 Introduction

Gauge theories have a foundational role in physics since they are the basic paradigm for the formulation of the Standard Model (SM) of fundamental particles and their interactions. The great success of the SM therefore makes it imperative that we understand the structure of gauge theories in different environments and kinematic regimes. Covariant perturbation theory for gauge theories is by now a well-developed and powerful technique and it is adequate for the analysis of the electroweak sector of the SM for most questions of interest. The situation for the strong nuclear forces, described by Quantum Chromodynamics (QCD), is very different. The high energy regime of QCD (energies $\gtrsim 10\,\mathrm{GeV}$) can be analyzed using perturbation theory by virtue of asymptotic freedom. But the low energy regime, where the interaction strength is large and where perturbation theory is no longer applicable, remains a real challenge. Decades of work have led to a fairly good qualitative understanding of the low energy regime of non-abelian gauge theories, but quantitative analysis of important questions such as how quarks bind together to form hadrons, what the nucleonic and nuclear matrix elements for the electroweak transitions of hadrons are, etc., is difficult. Lattice gauge theory, combined with large scale numerical simulations, has been the reliable workhorse for most questions of a nonpertrubative nature and, indeed, it has produced a number of useful results. However, it is important to correlate these results with an analytical approach to arrive at a complete or more comprehensive understanding of the physics of gauge theories.

In this review/lectures, we will try to develop an approach which is very different from covariant perturbation theory, namely, the Schrödinger representation in field theory where we use Hamiltonians and seek wave functionals which are solutions of the Schrödinger equation. Although this representation goes back to the early days of field theory, and has the conceptual simplicity of elementary quantum mechanics, it has rarely been used because of many perceived difficulties. Nevertheless, it may be more suitable for certain types of questions in field theory. To cite an elementary example, recall that a spacetime approach in terms of path integrals can be used to work out the bound state energy levels and transition matrix elements for the Hydrogen atom, but it is much simpler to use the Hamiltonian and the Schrödinger equation.

We will only be considering the application of this method to the three (or 2+1) dimensional Yang-Mills (YM) theory. However, it is useful to start with a few general

observations. Consider a simple scalar field theory with a classical action of the form

$$S = \int \left[\frac{1}{2} \partial_\mu \phi \, \partial^\mu \phi - \frac{m^2}{2} \phi^2 - \lambda \phi^4 \right] \tag{1.1}$$

In the canonical quantization of this theory, we start with the equal-time commutation rules, say at time $t = 0$,

$$\begin{aligned} [\phi(\vec{x}, 0), \phi(\vec{y}, 0)] &= 0 \\ [\phi(\vec{x}, 0), \pi(\vec{y}, 0)] &= i \, \delta(\vec{x} - \vec{y}) \\ [\pi(\vec{x}, 0), \pi(\vec{y}, 0)] &= 0 \end{aligned} \tag{1.2}$$

where $\pi(\vec{x}, 0) = \dot{\phi}(\vec{x}, 0)$. This suggests that we can define a set of ϕ-diagonal states $|\varphi\rangle$ obeying $\phi(\vec{x}, 0) \, |\varphi\rangle = \varphi(\vec{x}) \, |\varphi\rangle$, where $\varphi(\vec{x})$ is a c-number function. A Schrödinger wave function for a state $|\alpha\rangle$ will take the form

$$\Psi_\alpha[\varphi] = \langle \varphi | \alpha \rangle \tag{1.3}$$

It is a functional of φ. The commutation rules (1.2) can then be represented as

$$\begin{aligned} \langle \varphi | \, \phi(\vec{x}, 0) \, | \alpha \rangle &= \varphi(\vec{x}) \, \Psi_\alpha[\varphi] \\ \langle \varphi | \, \pi(\vec{x}, 0) \, | \alpha \rangle &= -i \frac{\delta}{\delta \varphi(\vec{x})} \, \Psi_\alpha[\varphi] \end{aligned} \tag{1.4}$$

This is the Schrödinger representation of the commutation rules.

The Hamiltonian corresponding to the action (1.1) has the form

$$\mathcal{H} = \int \left[\frac{1}{2} \pi^2 + \frac{1}{2} \phi (-\nabla^2 + m^2) \phi + \lambda \phi^4 \right] \tag{1.5}$$

The idea is that we can use this to write down and solve the Schrödinger equation. The vacuum state of the theory, represented by the wave functional $\Psi_0[\varphi]$, would thus satisfy

$$\begin{aligned} \mathcal{H} \, \Psi_0[\varphi] &= \int \left[-\frac{1}{2} \left(\frac{\delta^2}{\delta \varphi(\vec{x}) \delta \varphi(\vec{x})} \right) + \frac{1}{2} \varphi(x) \omega^2(x, y) \varphi(y) + \lambda \varphi^4(x) \right] \Psi_0[\varphi] \\ &= 0 \\ \int_{x,y} \varphi(x) \omega^2(x, y) \varphi(y) &\equiv \int_x \varphi(x) (-\nabla^2 + m^2) \varphi(x) \end{aligned} \tag{1.6}$$

where we have used the Schrödinger representation to write $\mathcal{H}$ as a functional differential operator which can act on $\Psi_0[\varphi]$.

A number of potential problems are evident at this stage. As with any field theory, we need regularization and renormalization. In covariant perturbation theory, the regularized action has the form

$$S = \int \left[Z_3 \left[\frac{1}{2} \partial_\mu \phi \, \partial^\mu \phi - \frac{m^2}{2} \phi^2 + \frac{\delta m^2}{2} \phi^2 \right] - Z_1 \lambda \phi^4 \right] \tag{1.7}$$

where Z_1, Z_3 and δm^2 will depend on the regularization parameter Λ (upper cutoff on momenta) and are chosen so as to render all correlation functions finite as $\Lambda \to \infty$. The situation in the Schrödinger representation is more complicated. We have functional derivatives at the same point $\vec{x}$ in the $\delta^2/\delta\varphi^2$-term, so it needs regularization and a Z-factor. A similar statement applies to the $\varphi(-\nabla^2\varphi)$-term. The mass term will need an additive renormalization as well, so we need a term $\frac{1}{2}\delta m^2 \varphi^2$. And finally we need regularization and a Z-factor for the interaction term. At this stage, we could envisage independent regularizations for the terms $\delta^2/\delta\varphi^2$ and $\varphi(-\nabla^2\varphi)$, since we have a separation of space and time and Lorentz invariance is not manifest. The requirement of Lorentz invariance will relate the Z-factors for these two terms. The regularization must be so chosen as to ensure this, Lorentz invariance is not automatic as in covariant perturbation theory. This is one of the complications of the Schrödinger representation for field theories.

There is one other issue associated with Poincaré invariance. One of the commutation rules for the Poincaré algebra is

$$[K_i, P_j] = i \, \delta_{ij} \mathcal{H} \tag{1.8}$$

where P_j is the total momentum and K_i is the generator of Lorentz boosts. Taking the expectation value of this with the vacuum state shows that if the vacuum is to be Lorentz invariant, we must have $\langle 0 | \, \mathcal{H} \, | 0 \rangle = 0$. So, for maintaining Poincaré invariance, $\mathcal{H}$ must be redefined by subtracting a certain c-number term to ensure this; this is the version of the familiar normal ordering in the present context.

In addition to the Hamiltonian, we must ensure that the wave functionals are well-defined. In general, this will require additional counterterms. One way to understand the genesis of such counterterms is to think of the wave function at time t_1 as defined by a path integral over the region $t < t_1$ as in

$$\Psi[\varphi, t_1] = \int [\mathcal{D}\tilde{\varphi}] \, e^{iS[\tilde{\varphi}, t_1, t_0]} \, \Psi[\varphi', t_0], \quad \tilde{\varphi}(\vec{x}, t_1) = \varphi(\vec{x}), \; \tilde{\varphi}(\vec{x}, t_0) = \varphi'(\vec{x}) \tag{1.9}$$

In the course of carrying out calculations using this form, we will be renormalizing an action defined on a spacetime region with boundaries (the time-slices at t_0 and t_1) and

it will require counterterms on the boundaries. These take the form

$$\Psi \to \exp\left[i \int Z_5 \varphi \partial_0 \varphi \ + \Delta \, \varphi^2 \right] \ \Psi \tag{1.10}$$

The Hamiltonian itself takes the form

$$\begin{aligned} \mathcal{H} \ &= \ \int \frac{1}{2} \left[-\frac{1}{Z_3 Z_0} \left(\frac{\delta^2}{\delta\varphi(\vec{x})\delta\varphi(\vec{x})} \right)_{\rm reg} + Z_3 Z_0 \Big((\nabla\varphi)^2 + (m^2 - \delta m^2)\varphi^2 \Big)_{\rm reg} \right] \\ &\quad + \int Z_1 \Big(\lambda \varphi(x)^4 \Big)_{\rm reg} - E_0 \end{aligned} \tag{1.11}$$

The factor Z_0 is related to Z_5 in (1.10). The boundary counterterms are another complication, in general, for the Schrödinger representation.

With the formalism as outlined above, and using a point-splitting regularization, Symanzik was able to prove the renormalizability of the ϕ^4-theory in the Schrödinger representation [1]. (While renormalizability of this theory in the covariant formalism was relatively straightforward, before Symanzik's work, there was even a general feeling that the theory was not renormalizable in the Schrödinger representation. There is some new physics which emerges in this formalism as well. Symanzik used the Schrödinger representation to analyze Casimir energies. Further, the additional Z-factor introduced by Symanzik can also be related to a new critical exponent, see [2].)

A useful observation worth mentioning at this stage is that the vacuum wave functional for the free theory (with $\lambda = 0$) is given by

$$\Psi_0[\varphi] = \left[\det \left(\frac{\sqrt{k^2 + m^2}}{\pi} \right)_{\rm reg} \right]^{\frac{1}{2}} \exp\left(-\frac{1}{2} \int_{x,y} \varphi(x) \big[\sqrt{k^2 + m^2} \big]_{x,y,{\rm reg}} \varphi(y) \right) \tag{1.12}$$

with $E_0 = \frac{1}{2} \big[\sqrt{k^2 + m^2} \big]_{x,x,{\rm reg}}$.

Given the additional complications with regard to regularization and renormalization, compared to covariant perturbation theory, one might wonder whether it is worth the trouble to pursue the Schrödinger representation in field theory. For certain questions of a nonperturbative nature, the answer seems to be a qualified yes. The kinetic operator in the Hamiltonian may be viewed as the Laplace operator on the infinite-dimensional space of field configurations and if we have some knowledge of the geometry and topology of this space, it can shed light on the spectrum of the Hamiltonian. A key inspirational paper in this context was by Feynman, who analyzed Yang-Mills theory in 2+1 dimensions [3]. These theories are rather optimal candidates for the

Schrödinger representation since there is no renormalization of the coupling constant, so some of the aforementioned problems can be avoided. Feynman tried to argue that the space of gauge-invariant configurations (gauge potentials modulo gauge transformations) is compact and hence can lead to a discrete spectrum for the Laplacian and ultimately a mass gap for the theory. This is not quite true, the configuration space is not compact, as shown by Singer, who however argued that the curvature of the space is positive [4].[1]

Feynman's arguments and Singer's analysis were carried out before we had an exact expression for the volume element for the configuration space. What we shall do here is to revisit this problem in the light of later developments. We will see that, modulo certain approximations and caveats as explained in detail below, there are a few key quantitative (and encouraging) results which emerge from our analysis:

1. There is an analytic formula for the string tension which compares very favorably with numerical estimates from lattice simulations.

2. One can calculate the Casimir energy for a parallel plate arrangement; this too compares very favorably with the lattice simulations.

There are also some additional insights obtained regarding supersymmetric theories, entanglement, etc., which we will comment on later.

Unlike the 1+1 dimensional cases, the 2+1 dimensional Yang-Mills theories have propagating degrees of freedom, so one might consider them to be closer to the 4d Yang-Mills theories, which is an added motivation for analyzing these theories. But they are also relevant for the high temperature (T) limit of 4d Yang-Mills theories. In this limit, the 4d (or 3+1 dimensional) theory reduces to a (Euclidean) 3d Yang-Mills theory with coupling constant $e^2 = g^2\mathsf{T}$, where g is the coupling constant of the 4d-theory. Electric fields and time-dependent processes become irrelevant. The mass gap of the 3d theory, from the point of view of the 4d theory, becomes the magnetic mass since it controls the screening of magnetic fields in the gluon plasma [6]. So the identification of the propagator mass in the 3d theory (either analytically or via the lattice simulation of the Casimir effect) can be useful for the 4d theory at high temperatures.

Three dimensional space is also famous as the home ground of the 3d Chern-Simons (CS) theory, with all its ramifications, knot theory, conformal field theory, etc. For the

[1] Feynman's analysis was modeled on his earlier very successful analysis of superfluid Helium. The comparison of the two cases and some of the nuances of the gauge theory are outlined in [5].

CS theory also, a beautiful analysis can be carried out in the Schrödinger representation, see [7].

In the following there are some sections or subsections marked with an asterisk; they capture the logical flow of the main arguments and the most interesting results. The remaining sections can be bypassed in a first reading. A return to this introduction may also be useful after reading through some of the following sections.

2 The gauge principle

The quintessential example of a gauge theory is quantum electrodynamics describing the interaction of electrons and positrons with the electromagnetic field. The starting point for this theory is the Lagrangian

$$L(\psi, \bar{\psi}, A) = \bar{\psi}\,[i\gamma^\mu(\partial_\mu - iA_\mu) - m]\,\psi - \frac{1}{4e^2}F_{\mu\nu}F^{\mu\nu} \tag{2.1}$$

where ψ is a 4-component spinor field in four dimensions representing the electron-positron field, $\bar{\psi} = \psi^\dagger\gamma^0$, and A_μ is the vector potential for the electromagnetic field. Also $F_{\mu\nu}$ is the field strength tensor defined as $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$. The components of $F_{\mu\nu}$ are related to the electric (E_i) and magnetic (B_i) fields as $F_{0i} = E_i$, $F_{ij} = \epsilon_{ijk}B_k$. The charge of the electron is e and its mass is m. Also, γ^μ are the Dirac γ-matrices obeying[2]

$$\gamma^\mu\,\gamma^\nu + \gamma^\nu\,\gamma^\mu = 2\,\eta^{\mu\nu}\,\mathbb{1} \tag{2.2}$$

The key property of the Lagrangian (2.1) for our analysis is gauge invariance. If we make a change of variables as

$$\begin{aligned} \psi \to \psi^g = g\,\psi, \qquad & \bar{\psi} \to \bar{\psi}^g = \bar{\psi}\,g \\ A_\mu \to A^g_\mu \;\; = \;\; & g\,A_\mu\,g^{-1} - i\partial_\mu g\,g^{-1} = A_\mu + \partial_\mu\theta, \end{aligned} \tag{2.3}$$

where $g = e^{i\theta}$, we find

$$L(\psi^g, \bar{\psi}^g, A^g) = L(\psi, \bar{\psi}, A) \tag{2.4}$$

Notice that $F_{\mu\nu}$ (i.e., E_i, B_i) is unchanged by the transformation (2.3). Classically the motion of a charged particle is governed by the Lorentz force law which involves only E_i, B_i. Hence, classically the entire dynamics is insensitive to the transformation (2.3).

[2] Our conventions and specific realizations are discussed in Appendix A.

Therefore, the gauge degree of freedom, namely $\theta(x)$, represents a redundancy in the dynamical variables used to describe the theory. Going to the quantum theory, notice that we can set A^g to zero along a line by defining $\theta(x)$ as

$$\theta(x) = \int_{x_0,C}^{x} dx^\mu A_\mu = \int_{x_0,C}^{x} A \tag{2.5}$$

where C denotes a path connecting the point x_0^μ to x^μ. In this case, ψ acquires a phase factor $e^{i\theta} = e^{i\int_C A}$. Thus, in the quantum theory (where phases are important and have observable consequences), we would need A_μ rather than just $F_{\mu\nu}$. But this statement needs qualification. If we consider deforming the path C to another path C' with the same end points, we get

$$\begin{aligned} e^{i\int_C A} &= e^{i\int_{C-C'} A}\, e^{i\int_{C'} A} \\ &= e^{i\int_\Sigma F}\, e^{i\int_{C'} A} \end{aligned} \tag{2.6}$$

where Σ is a surface with $C - C'$ as its boundary, i.e., $\partial\Sigma = C - C'$. (Here $-C'$ denotes C' traversed in the opposite direction from x^μ to x_0^μ.) Then by choosing C' as a fixed standard path, we can obtain the phase $e^{i\int_C A}$ for any path C in terms $F_{\mu\nu}$ again, and hence it is insensitive to the gauge transformation, when only phase differences are observable. (One might even use a C' in a region where $A = 0$.)

An exception to this can arise if there are noncontractible paths in the space under consideration, in other words, if the fundamental homotopy group Π_1 of the space is nonzero. This may arise due to an intrinsic nontrivial topology, or because certain regions are inaccessible to the particle/field ψ. In such cases, we might have different phases associated with different homotopy classes of paths, i.e., with different elements of Π_1. This will be in addition to the factor $e^{i\int_\Sigma F}$ which can relate phases for different paths within the same homotopy class. The conclusion is that, even in the quantum theory, apart from cases of nontrivial topology (specifically Π_1) for space, all dynamics can be obtained in terms of $F_{\mu\nu}$ (and other fields) which are insensitive to the gauge transformation. Thus we do have a true redundancy of description. The transformation (2.3) is often referred to as a gauge transformation and the invariance in (2.4) as gauge symmetry. But it is different from the usual Noetherian symmetries; it is actually a redundancy in the field variables used to describe the theory.

The function $g = e^{i\theta}$ is an element of the group $U(1)$, so the gauge symmetry in (2.3), (2.4) is a $U(1)$ gauge symmetry. The generalization of this to an arbitrary Lie group G is as follows. Consider a set of fields ψ^i, $i = 1, 2, \cdots, N$ which transform as an

N-dimensional representation R of the group G; i.e.,

$$(\psi^i)^g = g^{ij}\,\psi^j \tag{2.7}$$

We define a covariant derivative $D_\mu\psi$ as

$$(D_\mu\psi)^i = \partial_\mu\psi^i + (A_\mu)^{ij}\,\psi^j \tag{2.8}$$

where A_μ is an element of the Lie algebra of G, with (A_μ) as its matrix representative in the chosen representation R. Thus, if $\{T^a\}$ denote a basis for the Lie algebra of G, with $a = 1, 2, \cdots, \dim G$, realized as matrices in the representation R,

$$(A_\mu)^{ij} = -iA^a_\mu\,(T^a)^{ij} \tag{2.9}$$

We also define the gauge transform of A as

$$A^g_\mu = g\,A_\mu\,g^{-1} - \partial_\mu g\,g^{-1} \tag{2.10}$$

This is also in the matrix notation. The derivative $D_\mu\psi$ is covariant in the sense that

$$\begin{aligned}(D^g_\mu\psi^g)^i &= \left[\partial_\mu(g\psi) + (g\,A_\mu\,g^{-1} - \partial_\mu g\,g^{-1})(g\psi)\right]^i \\ &= g^{ij}(\partial_\mu\psi + A_\mu\psi)^j = g^{ij}(D_\mu\psi)^j\end{aligned} \tag{2.11}$$

As a particular case, if the fields transform according to the adjoint representation of the group, $(T^a)^{ij} = -if^{aij}$, where f^{aij} are the structure constants of the Lie algebra of G in the chosen basis. Thus they are given by $[T^a, T^b] = if^{abc}T^c$. In this case $(D_\mu\psi)^a = \partial_\mu\psi^a + f^{abc}A^b_\mu\psi^c$.

The commutator of covariant derivatives defines the field strength tensor as

$$\begin{aligned}[D_\mu, D_\nu] &= F_{\mu\nu} = (-iT^a)\,F^a_{\mu\nu} \\ F^a_{\mu\nu} &= \partial_\mu A^a_\nu - \partial_\nu A^a_\mu + f^{abc}A^b_\mu A^c_\nu\end{aligned} \tag{2.12}$$

By construction, $F_{\mu\nu}$ transforms homogeneously under gauge transformations as

$$F^g_{\mu\nu} = g\,F_{\mu\nu}\,g^{-1} \tag{2.13}$$

If we have a unitary representation of the group on the fields ψ, we have $\bar\psi^g = \bar\psi g^\dagger = \bar\psi g^{-1}$, so that a Lagrangian consistent with gauge invariance is

$$L(\psi, \bar\psi, A) = \bar\psi\,[i\gamma^\mu D_\mu - m]\,\psi - \frac{1}{4e^2}F^a_{\mu\nu}F^{a\mu\nu} \tag{2.14}$$

This is the kind of Lagrangian we use for coupling of quarks to the gluons (particles corresponding to A^a_μ) in quantum chromodynamics (QCD). The left and right chiral components of the fermion field couple to the gauge field in an identical fashion, so the coupling is vectorial in nature. The Standard Model also involves chiral or axial couplings of the quarks and leptons to various gauge fields. Most of of our analysis will be for the pure gauge theory, and when we discuss gauge fields in interaction with matter, we will mostly consider vectorial couplings. The action for the gauge field part of the Lagrangian (2.14) is the Yang-Mills action

$$\begin{aligned} S_{\text{Y-M}} &= -\frac{1}{4e^2}\int dt d\mu\, F^a_{\mu\nu}F^{a\mu\nu} \\ &= \frac{1}{2e^2}\int dt d\mu\, (E^a_i E^a_i - B^a_i B^a_i) \end{aligned} \tag{2.15}$$

where $d\mu$ is the volume element for the spatial manifold. For the special case of a $U(1)$ gauge theory where $\dim G = 1$, this action agrees with the action for the electric and magnetic fields in electrodynamics. The nonabelian analogs of these fields can be written out as

$$\begin{aligned} E^a_i &= F^a_{0i} = \frac{\partial A^a_i}{\partial t} - \partial_i A^a_0 + f^{abc}A^b_0 A^c_i \\ &= \frac{\partial A^a_i}{\partial t} - (D_i A_0)^a \\ B^a_i &= \frac{1}{2}\epsilon_{ijk}F^a_{jk} = \frac{1}{2}\epsilon_{ijk}(\partial_j A^a_k - \partial_k A^a_j + f^{abc}A^b_j A^c_k) \quad \text{(3+1 dim)} \\ B^a &= \frac{1}{2}\epsilon_{jk}F^a_{jk} = \frac{1}{2}\epsilon_{jk}(\partial_j A^a_k - \partial_k A^a_j + f^{abc}A^b_j A^c_k) \qquad \text{(2+1 dim)} \end{aligned} \tag{2.16}$$

The equations of motion for the Y-M theory are

$$\begin{aligned} (D_i E_i)^a &= 0 \\ \frac{\partial E^a_i}{\partial t} &= \begin{cases} -\epsilon_{ijk}(D_j B_k)^a & \text{3+1 dim} \\ -\epsilon_{ij}(D_j B)^a & \text{2+1 dim} \end{cases} \end{aligned} \tag{2.17}$$

The first of these is the Gauss law familiar from electrodynamics, now generalized to the nonabelian case. The second is an equation of motion, in the sense of defining time-evolution, for the field E^a_i.

Our aim is to consider the Hamiltonian formulation of the Y-M theory using the functional Schrödinger formulation. From now on, unless specifically indicated, we will consider $2+1$ dimensions. As a first step, by extending the action in (2.15) to a

general curved manifold with metric $g_{\mu\nu}$ as $F^a_{\mu\nu}F^{a\mu\nu} \to \sqrt{-g}\, g^{\mu\alpha}g^{\nu\beta}F^a_{\mu\nu}F^{a\alpha\beta}$ and taking the variation with respect $g^{\mu\nu}$, we find the energy-momentum tensor for the theory as

$$T_{\mu\nu} = \frac{1}{e^2}\left[-F^a_{\mu\alpha}\eta^{\alpha\beta}F^a_{\beta\nu} + \frac{1}{4}\eta^{\mu\nu}F^a_{\alpha\beta}F^{a\alpha\beta}\right] \tag{2.18}$$

This identifies the Hamiltonian as

$$\mathcal{H} = \int d\mu\, T_{00} = \frac{1}{2e^2}\int d\mu\, (E^2 + B^2) \tag{2.19}$$

To obtain the Poisson brackets, or the commutations rules for the fields in the quantum theory, we need the canonical structure for the fields. From the term involving time-derivatives of the fields in (2.15), we can identify this as

$$\Omega = \frac{1}{e^2}\int d\mu\, \delta E^a_i\, \delta A^a_i = \int d\mu\, \delta\Pi^a_i\, \delta A^a_i, \qquad \Pi^a_i = \frac{E^a_i}{e^2} \tag{2.20}$$

This is to be interpreted as a differential two-form in the space of field configurations (E^a_i, A^a_i); we use δ to denote the exterior derivative on the space of fields. On the spatial manifold at a fixed time, E^a_i is to be treated as an independent variable since it involves the time-derivative of A^a_i. It is proportional to the canonical momentum Π^a_i conjugate to A^a_i. The equal-time commutation rules defined by (2.20) are

$$\begin{aligned}
[A^a_i(x), A^b_j(y)] &= 0 \\
[E^a_i(x), E^b_j(y)] &= 0 \\
[E^a_i(x), A^b_j(y)] &= -i\, e^2\delta_{ij}\delta^{ab}\delta^{(2)}(x-y)
\end{aligned} \tag{2.21}$$

The commutation rules (2.21) show that Π^a_i is the variable canonically conjugate to A^a_i. There is no variable conjugate to A^a_0. Put another way, the canonical momentum for A^a_0 is zero. If we augment Ω by the addition of a term $\int \delta\Pi^a_0\, \delta A^a_0$, then we must carry out a reduction of the phase space by setting Π^a_0 to zero as a constraint, $\Pi^a_0 \approx 0$ (in the sense of Dirac's theory of constraints). As a conjugate constraint, we can use $A^a_0 \approx 0$. Thus the pair (Π^a_0, A^a_0) will be eliminated from the theory.

The Hamiltonian equations of motion which follow from the canonical brackets are obtained as

$$\begin{aligned}
\frac{\partial A^a_i}{\partial t} &= E^a_i \\
\frac{\partial E^a_i}{\partial t} &= -\epsilon_{ij}(D_j B)^a
\end{aligned} \tag{2.22}$$

Notice that the first of these equations requires $A_0^a = 0$ for consistency with the definition in (2.16). If we did not set A_0^a to zero, we would need to add a term to the Hamiltonian to obtain the result (2.16). The canonical Hamiltonian and the Hamiltonian defined by T_{00} would differ by terms proportional to the constraint. With $A_0^a = 0$, the first of the equations in (2.22) reproduces the definition of E_i^a. The second equation agrees with the second of the Lagrangian equations of motion in (2.17).

In terms of the canonical momentum, the first of the Lagrangian equations in (2.17) reads $(D_i\Pi_i)^a = 0$, so it does not involve time-derivatives. Therefore it cannot be reproduced as a Hamiltonian equation of motion. For equivalence of the Hamiltonian formulation to the Lagrangian given as (2.17), we have to impose $(D_i\Pi_i)^a = 0$ as an additional condition. It should be viewed as a constraint on the phase space variables or on the initial data.

We have restricted the field variables (by use of the freedom of gauge transformations) to some extent by setting $A_0^a = 0$. But the theory would still allow for gauge transformations g which do not depend on time, so that they preserve the condition $A_0^a = 0$. The constraint $D_iE_i = 0$ may be viewed as the statement of this residual gauge freedom. We can then choose a constraint conjugate to D_iE_i, say $\nabla \cdot A \approx 0$ for example, and carry out a further canonical reduction to obtain Ω on the reduced phase space (where $D_iE_i = 0$ and $\nabla_i A_i = 0$). We can then formulate Poisson brackets and commutators in terms of this reduced Ω. This is the approach of gauge-fixing, $\nabla_i A_i = 0$ being the gauge-fixing condition. Alternatively, in the quantum theory we can impose $D_iE_i = 0$ not as an operator condition but as a condition on states or wave functions. This is the approach we will be pursuing.

As is well-known, conditions imposed in terms of operators should be understood as valid with suitable smearing using test functions. The nature of the test functions is crucial to determining the physical consequences of the theory. We consider the smeared operator

$$G_0(\theta) = \int d\mu\, \theta^a (D_i\Pi_i)^a \tag{2.23}$$

If we impose the condition

$$G_0(\theta)\Psi = 0 \tag{2.24}$$

on the wave functions Ψ in the theory, for consistency, we will also need the commutator $[G_0(\theta), G_0(\theta')]$ to vanish on Ψ. From the canonical commutation rules (2.21) it is easy

to check that

$$\begin{aligned}[G_0(\theta), G_0(\theta')] &= iG_0(\theta \times \theta') + i \oint_{\partial V} (\theta \times \theta')^a \Pi_i^a \, dS_i \\ (\theta \times \theta')^a &\equiv f^{abc} \theta^b \theta'^c \end{aligned} \tag{2.25}$$

We see that we cannot consistently impose (2.24) unless Π_i^a vanishes fast enough as we approach ∂V or at spatial infinity. This would in turn amount to requiring all charges to vanish (this will be clearer soon), which is not something we can impose $a\ priori$ in the theory. The only other option is to require the test functions to vanish on ∂V. In this case, the surface term in (2.25) will vanish and we have a closed algebra for the $G_0(\theta)$'s and the condition (2.24) can be consistently imposed. In terms of its action on fields, we find

$$\begin{aligned} e^{iG_0(\theta)} \left[\int d\mu \, A_i^a v_i \right] e^{-iG_0(\theta)} &= \int d\mu \, A_i^a v_i + i[G_0(\theta), \int d\mu \, A_i^a v_i] + \cdots \\ &= \int d\mu \, A_i^a v_i + \int d\mu \, \theta^a (\nabla_i v_i) - f^{abc} \int d\mu \, A_i^b \theta^c v_i + \cdots \\ &= \int d\mu \, A_i^a v_i - \int d\mu \, (D_i \theta)^a v_i + \cdots + \oint_{\partial V} \theta^a v_i dS_i \\ &= \int d\mu \, (A_i - D_i \theta)^a v_i, \quad \text{if } \theta^a \to 0 \text{ on } \partial V \end{aligned} \tag{2.26}$$

For the electric field we find

$$e^{iG_0(\theta)} \left[\int d\mu \, E_i^a w_i \right] e^{-iG_0(\theta)} = \int d\mu \, E_i^a w_i + f^{abc} \int d\mu \, \theta^b E_i^c w_i + \cdots \tag{2.27}$$

(In (2.26) and (2.27), v_i and w_i are test functions for A_i^a and E_i^a.) The right hand sides of these equations are of the form of infinitesimal gauge transformations (2.10), (2.13) with $g = e^{-it^a \theta^a} \approx 1 - it^a \theta^a$. This means that the operator $G_0(\theta)$ will generate infinitesimal gauge transformations of (A_i^a, E_i^a) provided θ^a vanishes at spatial infinity (or on the boundary of the spatial volume under consideration). Since the Hamiltonian is invariant under gauge transformations, $[G_0(\theta), \mathcal{H}] = 0$, and hence the requirement $G_0(\theta)\Psi = 0$ will be preserved under time-evolution as well. The closed algebra (2.25) is a statement of the group property that a sequence of infinitesimal transformations of the form (2.26), (2.27) can be used to generate a finite transformation. We can now define an infinite dimensional group $\mathcal{G}_*$ as follows:

$$\mathcal{G}_* = \{\text{Set of } g(x) : \text{Space } \to G \text{ such that } g(x) \to 1 \text{ on } \partial V\} \tag{2.28}$$

If we consider all of $\mathbb{R}^2$, we may define $\mathcal{G}_*$ as

$$\mathcal{G}_* = \{\text{Set of } g(x) : \mathbb{R}^2 \to G \text{ such that } g(x) \to 1 \text{ as } |\vec{x}| \to \infty\} \tag{2.29}$$

The condition (2.24) is the statement that all wave functions in the theory are invariant under gauge transformations $g \in \mathcal{G}_*$. In this sense, $\mathcal{G}_*$ is the true gauge group of the theory. To distinguish wave functions or states which are more general and do not necessarily obey (2.24), we refer to states satisfying (2.24) as "physical states".

Given that states or wave functions obey (2.24), for the matrix element of an operator $\mathcal{O}$ we can write

$$\begin{aligned} \langle \Psi_1 | \, \mathcal{O} \, | \Psi_2 \rangle &= \langle \Psi_1 | \, e^{iG_0(\theta)} \mathcal{O} e^{-iG_0(\theta)} \, | \Psi_2 \rangle \\ &= \langle \Psi_1 | \, \mathcal{O} \, | \Psi_2 \rangle + i \, \langle \Psi_1 | \, [G_0(\theta), \mathcal{O}] \, | \Psi_2 \rangle + \cdots \end{aligned} \tag{2.30}$$

This will give an inconsistent result unless we have $[G_0(\theta), \mathcal{O}] = 0$. Therefore, we can say that an operator $\mathcal{O}$ is an observable and can have well-defined matrix elements only if it weakly commutes with $G_0(\theta)$, i.e., if $\langle \Psi_1 | \, [G_0(\theta), \mathcal{O}] \, | \Psi_2 \rangle = 0$, for all physical states Ψ_1, Ψ_2.

We now turn to another set of transformations of interest. Towards this, we first consider transformations of the type (2.10), (2.13) where $g \in G$ is a constant not necessarily equal to one on the spatial manifold; i.e., the transformations are

$$A_i \to g \, A_i g^{-1}, \quad E_i \to g \, E_i g^{-1} \tag{2.31}$$

The Hamiltonian (2.19) is clearly invariant under these. Further, this is not a gauge transformation and cannot be removed by the choice of a suitable element of $\mathcal{G}_*$ since elements of $\mathcal{G}_*$ must become the identity on ∂V or at spatial infinity. By choice of the action of $G_0(\theta)$, we can go from g to $g'(x)g$ where $g'(x) \to 1$ on the boundary. But the value of the combined transformation $g'(x)g$ still has the value g (which is not necessarily the identity) on the boundary. So these transformations (2.31) generate a Noether-type symmetry and the states of the system can be classified by representations of the group G. In fact the transformations (2.31) with constant g's correspond to charge rotations and the states in the various irreducible representations of G correspond to states with different possible charges.

It is also useful to consider another set of operators

$$G(\theta) = - \int d\mu \, (D_i \theta)^a \Pi_i^a \tag{2.32}$$

These coincide with $G_0(\theta)$ for those test functions θ^a which vanish on ∂V, but, in general, we can consider $G(\theta)$ even for those functions θ^a which do not vanish on ∂V. (Notationally, we distinguish the two by using the subscript on $G_0(\theta)$ to indicate that it

is for the case when $\theta^a \to 0$ on ∂V.) It is easy enough to check that

$$\begin{aligned} i[G(\theta), \int d\mu\, A_i^a v_i] &= \int d\mu\, (A_i - D_i\theta)^a v_i \\ [G(\theta), G(\theta')] &= i\, G(\theta \times \theta') \end{aligned} \tag{2.33}$$

So $G(\theta)$ does generate gauge transformations (as in (2.10), (2.13)) even for $\theta^a \neq 0$ on ∂V. (But recall that these are not true gauge transformations as they are not elements of $\mathcal{G}_*$.) We can use the freedom of gauge transformations by $G_0(\theta')$ to change the value of θ everywhere except on the boundary. Thus $G(\theta)$ is characterized by the boundary value θ (modulo the action of $G_0(\theta')$). The commutation rules also give

$$\begin{aligned} G_0(\theta')\, G(\theta)\, \Psi &= G(\theta)\, G_0(\theta')\, \Psi + iG_0(\theta' \times \theta)\, \Psi \\ &= 0 \end{aligned} \tag{2.34}$$

so that $G(\theta)\Psi$ are also states compatible with the requirement of (2.24). In other words, the action of $G(\theta)$ on Ψ's will generate physical states in the theory. Among the operators $G(\theta)$ there are the ones mentioned earlier where θ^a on ∂V or spatial infinity is a constant (that is, independent of angular directions), but not necessarily the identity. These generate charge rotations and hence they lead to the charged states of the theory. More generally, the operators $G(\theta)$ for those θ which may have nontrivial angular dependence or is a nonconstant function on ∂V generate observable dynamical degrees of freedom localized on the boundary. They are usually referred to as edge states. Notice that $[G_0(\theta), G(\theta)] = 0$.

The fact that the wave functions corresponding to physical states are gauge invariant means that their normalization has to be defined with a gauge-invariant volume element. Since A_i^a at different spatial points commute, we can consider A-diagonal wave functions $\Psi(A)$. (We can equally well consider E-diagonal ones, but for the moment we stay with $\Psi(A)$.) Thus $\Psi_1^* \Psi_2$ for physical states will be gauge-invariant and integration over all configurations A will clearly diverge. To define the proper volume element, we start by defining

$$\mathcal{A} = \{ \text{ Set of all gauge potentials such that } F_{ij} \to 0 \text{ as } |\vec{x}| \to 0\} \tag{2.35}$$

We impose a mild condition on the gauge potentials. Also, here, by gauge potential we mean a Lie-algebra valued one-form on the spatial manifold, $A = (-it^a) A_i^a dx^i$. This space is actually an affine space, i.e., any two points on $\mathcal{A}$ can be connected by a straight line as

$$A(\tau) = A^{(1)}\, (1 - \tau) + A^{(2)}\, \tau \tag{2.36}$$

where τ is a real parameter $0 \leq \tau \leq 1$. The straight line (2.36) connects $A^{(1)}$ at $\tau = 0$ to $A^{(2)}$ at $\tau = 1$. The key point here is that, for any value of τ, $A(\tau)$ transforms as a gauge potential,

$$A^g(\tau) = g\, A(\tau)\, g^{-1} - dg\, g^{-1} \tag{2.37}$$

Hence the entire straight line is in $\mathcal{A}$. Because of this property, the topology of the $\mathcal{A}$ is trivial, it is a flat contractible space. We can then consider the space $\mathcal{C} = \mathcal{A}/\mathcal{G}_*$ which is the space of all gauge potentials modulo gauge transformations. The configurations of the form $A^g = gAg^{-1} - dgg^{-1}$, for $g \in \mathcal{G}_*$, give the orbit of A under the action of $\mathcal{G}_*$. So $\mathcal{C}$ will also be referred to as the space of $\mathcal{G}_*$-orbits in $\mathcal{A}$, or the gauge-orbit space, for short. This is the space of physical configurations and the reduced phase space may be viewed as the cotangent space of $\mathcal{C}$. The wave functions are defined as functions on $\mathcal{C}$. Therefore the inner product for states should be defined with an integration measure (or the volume element) for the space $\mathcal{C}$. Expressed mathematically,

$$\langle 1|2\rangle = \int d\mu(\mathcal{C})\, \Psi_1^*\, \Psi_2 \tag{2.38}$$

A similar statement can be made if we choose to represent states by wave functions which are functions of E as well.

The Hamiltonian (2.19) in terms of its action on Ψ can be written as

$$\mathcal{H}\,\Psi = \frac{1}{2}\int d\mu\, \left[-e^2 \frac{\delta^2}{\delta A_i^a \delta A_i^a} + \frac{B^a B^a}{e^2}\right]\, \Psi \tag{2.39}$$

where we have used the functional Schrödinger representation of E_i^a,

$$E_i^a = -ie^2 \frac{\delta}{\delta A_i^a} \tag{2.40}$$

The functional differential operator, or the kinetic energy term in (2.39) is the functional Laplace operator on the space $\mathcal{A}$. But since it acts on Ψ's which are gauge-invariant, it can be viewed as the Laplace operator on the space $\mathcal{C}$.

We are now in a position to assemble the ingredients needed for the Hamiltonian formulation of the theory. Firs of all, the Hamiltonian has the form

$$\mathcal{H} = \frac{1}{2}\int d\mu\, \left[-e^2 \Delta_{\mathcal{C}} + \frac{B^2}{e^2}\right]\, \Psi \tag{2.41}$$

where $\Delta_{\mathcal{C}}$ is the Laplace operator on the configuration space $\mathcal{C}$. The wave functions themselves are gauge-invariant, i.e., defined as functions on $\mathcal{C}$. Their inner product for states $|1\rangle$ and $|2\rangle$ is given by (2.38), where $d\mu(\mathcal{C})$ is the volume element on $\mathcal{C}$.

Thus the key ingredients we need to calculate are the Laplacian $\Delta_{\mathcal{C}}$ and the volume element $d\mu(\mathcal{C})$. Both of these have to be defined with suitable regularizations, as for any field theory. Further, as mentioned in section 1, since we are using the Hamiltonian approach, we do not have manifest Lorentz invariance. So we do have to verify that the regularizations are compatible with Lorentz invariance.

3 Confinement

One of the key features of a nonabelian gauge theory is the confinement of particles or fields in nontrivial representations of the gauge group. As indicated in the last section, *a priori* we should allow for charged states which are generated by $G(\theta)$ which was defined in (2.32). Confinement refers to the statement that, in the nonabelian Yang-Mills theory, the dynamics is such that there are no charged states in the physical spectrum. Put another way, such states have infinite energy and therefore cannot be dynamically excited. Although this is not a proven fact, there are strong indications to support the idea of confinement. However, a direct analysis of the spectrum of the Hamiltonian, with a view to elucidating confinement, has not yet been successful. A possible strategy would then be to look for observables which can serve as useful diagnostics of confinement and to try to calculate them in some way. The most important among these is the Wilson loop operator defined by

$$W_R(C) = \mathrm{Tr}\left[\mathcal{P}\exp\left(-\oint_C A_\mu dx^\mu\right)\right] \tag{3.1}$$

Here $A_\mu = -it_a A^a_\mu$ and t_a, which are the generators of the Lie algebra, are in the representation R. This is indicated by the subscript on $W(C)$. The integral is over a closed curve C. Since A_μ at different points along the curve or for different choices of μ do not commute in general, there has to be an ordering prescription in how the line integral is evaluated. This is taken to be path-ordering, by which we mean the following. Let us parametrize the curve as $x^\mu(\tau)$, $0 \le \tau \le 1$, and divide the interval of τ into a sequence of infinitesimal segments ϵ_1, ϵ_2, $\cdots$, ϵ_n, with $n \to \infty$ and $\epsilon_k \to 0$ in the end as usual. Then the path-ordered integral from $x^\mu = x^\mu(0)$ to $y^\mu = x^\mu(1)$ is given by

$$\begin{aligned} W(y,x,C) &= \mathcal{P}\exp\left(-\int_x^y A_\mu dx^\mu\right) \\ &= \exp\left(-\int_{\epsilon_{n-1}}^{\epsilon_n} A_\mu \frac{dx^\mu}{d\tau}d\tau\right)\exp\left(-\int_{\epsilon_{n-2}}^{\epsilon_{n-1}} A_\mu \frac{dx^\mu}{d\tau}d\tau\right)\cdots \\ &\quad \cdots\exp\left(-\int_{\epsilon_1}^{\epsilon_2} A_\mu \frac{dx^\mu}{d\tau}d\tau\right) \end{aligned} \tag{3.2}$$

For the closed curve, we have $y^\mu = x^\mu$ and we take the trace of the resulting expression. For an open interval, we have the gauge transformation property

$$\left[\mathcal{P}\exp\left(-\int_x^y A^g_\mu dx^\mu\right)\right]^{ij} = \left[g(y)\,\mathcal{P}\exp\left(-\int_x^y A_\mu dx^\mu\right)\,g^{-1}(x)\right]^{ij} \tag{3.3}$$

This shows that once we close the curve and take the trace, we get a gauge-invariant quantity. The Wilson loop operators are thus observables of the theory. In fact, by choosing all possible closed curves, we get an over-complete set of observables. All other observables can be constructed from $W_R(C)$.

The Wilson loop operator is important for another reason as well. The expectation value of $W_R(C)$ is related to the interaction energy of a heavy particle-antiparticle pair belonging to the representation R and its conjugate. Such a pair can be used as a probe into the dynamics of the gauge theory. They are taken to be heavy so that their own dynamics is trivial and does not complicate the interpretation of the result, since the focus is on the gauge theory.

In order to relate $W_R(C)$ to the energy of a particle-antiparticle pair, we will start by considering the process where we start with a heavy static particle-antiparticle pair separated by a spatial distance L at a certain time x^0. We will use ϕ and $\phi^\dagger$ as the annihilation and creation operators for the particle; χ and $\chi^\dagger$ will play a similar role for the antiparticle. Since these are taken to be heavy, the action for these fields is just the usual nonrelativistic action, but we can even omit the $(\nabla^2/2M)$-part. Thus

$$\mathcal{S}(\phi,\chi) = \int d^4x\ \left[i\phi^\dagger D_0\phi + i\chi^\dagger D_0\chi\right] \tag{3.4}$$

where $D_0\phi = \partial_t\phi + A_0\phi$ and $D_0\chi = \partial_t + A_0^*\chi$ are the covariant derivatives of ϕ and χ, respectively. This is in accordance with the fact that the fields transform under gauge transformations as $\phi \to g\phi$, $\chi \to g^*\chi$. We start with a gauge-invariant state corresponding to the particle-antiparticle pair separated by a spatial distance L. This state can be represented as

$$F^\dagger(x^0, x^1, x^1+L)\,|0\rangle = \phi^{\dagger i}(x)\,W^{ij}(x,y)\,\chi^{\dagger j}(y)\,|0\rangle \tag{3.5}$$

where $x = (x^0, x^1)$, $y = (x^0, x^1 + L)$ and $W^{ij}(x,y)$ is as in (3.2) over, say, a straight line segment. We have taken the separation of the pair to be along the x^1-direction, for simplicity.

Let $\mathcal{H}$ be the Hamiltonian for the Yang-Mills theory coupled to these matter fields ϕ, χ. As usual, we can set A_0 to zero; the A_0-dependent terms in (3.4) are then zero but

will contribute to $\mathcal{H}$ via the Gauss law, which now takes the form

$$\frac{(D_i E_i)^a}{e^2} + \phi^\dagger t^a \phi + \chi^\dagger(-\tilde{t}^a)\chi = 0 \tag{3.6}$$

Here $\tilde{t}^a$ is the transpose of t^a, corresponding to the conjugate representation. ($t^a \to -\tilde{t}^a$ is the conjugation operation in the Lie algebra.)

We now consider the time-evolution of the state (3.5) by an imaginary amount $-iT$ and then take its overlap with (3.5). The amplitude for this is given by

$$\langle 0|F\ e^{-HT} F^\dagger|0\rangle \approx \mathcal{N}\ e^{-E(L)T} \tag{3.7}$$

where $\mathcal{N}$ is some prefactor related to the normalization of F, and $E(L)$ is the energy of the pair. We are interested in taking T to be large, so that $E(L)$ will be the energy of the lowest energy state which can be created by $F^\dagger$. Since the particles are heavy and static, $E(L)$ is basically just the interaction energy of the pair due to the gauge field.

By the usual technique of the slicing of the time-interval, we can represent this amplitude as a Euclidean functional integral

$$\begin{aligned}\langle 0|F\ e^{-HT} F^\dagger|0\rangle &= \int [dA\, d\phi\, d\chi]\ \exp\left[-\mathcal{S}_E(A,\phi,\chi)\right] \\ &\qquad \chi^i(y') W^{*ji}(y',x')\phi^j(x')\ \phi^{\dagger r}(x) W^{rs}(x,y)\chi^{\dagger s}(y)\end{aligned} \tag{3.8}$$

where $x' = (x^0+T, x^1)$, $y' = (x^0+T, x^1+L)$. The (ϕ,χ)-part of the Euclidean action which appears in this functional integral is given by

$$\mathcal{S}_E(\phi,\chi) = \int d\mu\ \left[\phi^\dagger \frac{\partial \phi}{\partial \tau} + \chi^\dagger \frac{\partial \chi}{\partial \tau}\right] \tag{3.9}$$

This leads to the propagators

$$\begin{aligned}\langle \phi^i(x)\phi^{\dagger j}(x')\rangle &= \delta^{ij}\,\theta(\tau-\tau')\,\delta(\vec{x}-\vec{y}) \\ \langle \chi^i(x)\chi^{\dagger j}(x')\rangle &= \delta^{ij}\,\theta(\tau-\tau')\,\delta(\vec{x}-\vec{y})\end{aligned} \tag{3.10}$$

where $\theta(\tau-\tau')$ is the step function and τ denotes the Euclidean time-coordinate. The amplitude in (3.8) then reduces to

$$\begin{aligned}\langle 0|F\ e^{-HT} F^\dagger|0\rangle &= \int [dA]\ e^{-\mathcal{S}_{\rm YM}}\ W^{ij}(y',x')\ W^{ji}(x,y) \\ &= \int [dA]\ e^{-\mathcal{S}_{\rm YM}}\ W_R(C)\end{aligned}$$

$$= \langle W_R(C) \rangle \tag{3.11}$$

where C is the rectangle with vertices x, y, x', y'. Since $A_0 = 0$, we can put in the two time-like segments for free to complete the loop. Comparing this expression with (3.7), we see that

$$\langle W_R(C) \rangle \approx \mathcal{N} e^{-E(L)T} \tag{3.12}$$

This shows that the Euclidean expectation value of a large Wilson loop can be used to identify the interaction energy of a heavy static particle-antiparticle pair. Even though we used the $A_0 = 0$ gauge, $W_R(C)$ is gauge-invariant, and so are energies of gauge-invariant states. Thus the result holds true in general.

If the interaction energy $E(L)$ increases with the separation L, say, $E(L) \to \infty$ as $L \to \infty$, then it will cost arbitrarily large energy to remove a charged particle from its conjugate to an arbitrarily far away point, if the pair is created by any process. This is what we expect if there is confinement. In the case of nonabelian gauge theories, the expectation is that the interaction energy will grow linearly with L, i.e., $E(L) = \sigma L$. The coefficient σ is known as the string tension. In terms of the Wilson loop, this statement is expressed as

$$\begin{aligned} \langle W_R(C) \rangle &\approx \mathcal{N} \exp(-\sigma\, L\, T) \\ &\approx \mathcal{N} \exp(-\sigma\, A_C) \end{aligned} \tag{3.13}$$

where A_C is the area of the minimal surface whose boundary is C.

The use of the term "string tension" is related to the following qualitative picture of confinement. If we consider a heavy particle-antiparticle pair, the expectation is that the chromoelectric flux lines connecting the particle and the antiparticle are collimated to a thin tube of flux, which we refer to as the string, by the properties of the vacuum. Since the energy of a string would increase linearly with the length, the proportionality factor being the tension of the string, this picture would explain the linear rise of the potential.

Equation (3.13) shows that the area-law behavior of the expectation value $\langle W_R(C) \rangle$ can be used as a test of confinement. This works for all representations which cannot be screened. Since the average in $\langle W_R(C) \rangle$ is done with the Yang-Mills action, the theory allows for the dynamical generation of gluons, which belong to the adjoint representation of the group G. Thus when we impart energy to a particle-antiparticle

pair, separating the constituents, $E(L)$ can grow to a point where it becomes possible to create a number of gluons spontaneously. If the representation R is such that $R \otimes (\text{Adjoint})$ (or $R \otimes \text{Adjoint} \otimes \text{Adjoint} \cdots$) contains the trivial representation, (these are called screenable representations), then the pair-configuration can decompose into a particle-gluon(s) state (of zero charge) and an antiparticle-gluon(s) state (also of zero charge). The interaction energy between these composites is no longer $E(L)$, since each has zero charge, so they can be separated far from each other. Correspondingly, $\langle W_R(C) \rangle$ will not exhibit an area law. Thus, while confinement continues to be true (since the particle-gluon(s) state and the antiparticle-gluon(s) state each has zero charge), the expectation value of the Wilson loop is no longer a good diagnostic tool.

The picture in terms of the string of flux connecting the particle-antiparticle pair is that the string breaks by the spontaneous production of gluons, which leads to new composites of zero charge and hence there is no longer any string of flux connecting these states.

From the argument given above, we see that, strictly speaking, $\langle W_R(C) \rangle$ is useful only for nonscreenable representations, namely, those for which $R \otimes \text{Adjoint} \otimes \text{Adjoint} \cdots$ does not contain the trivial representation. Nevertheless, our argument with $E(L)$ shows that we should expect the area law to hold until $E(L)$ becomes large enough to create a pair of gluons. So for a limited range of L, the area law for $\langle W_R(C) \rangle$ can still be obtained and can still be useful.

4 *Parametrization of gauge fields

We will now consider a special parametrization for the gauge fields which will facilitate working out the Hamiltonian and the volume element $d\mu(\mathcal{C})$ in terms of manifestly gauge-invariant variables [8], see also [9]. We are primarily interested in Yang-Mills theories on flat $(2+1)$-dimensional space, so the spatial manifold is $\mathbb{R}^2$. The two spatial coordinates x^1, x^2 can be combined into the complex combinations $z = x^1 - ix^2$, $\bar{z} = x^1 + ix^2$, with the corresponding derivatives

$$\partial \equiv \partial_z = \frac{1}{2}(\partial_1 + i\partial_2), \quad \bar{\partial} \equiv \partial_{\bar{z}} = \frac{1}{2}(\partial_1 - i\partial_2) \tag{4.1}$$

As explained before, we can take $A_0 = 0$. For the Abelian gauge theory, for the spatial components of A, we can use the Hodge decomposition

$$A_i = -i(\partial_i \varphi + \epsilon_{ij} \partial_j \chi) \tag{4.2}$$

for real functions φ and χ on $\mathbb{R}^2$. We use antihermitian A_i so that the covariant derivative is $(\partial_i + A_i)$, similar in form to what is usually used for the nonabelian case. For the complex components, we can write

$$A \equiv A_z = \frac{1}{2}(A_1 + iA_2) = -\partial_z \Theta, \quad \Theta = \chi + i\varphi \tag{4.3}$$

with $\bar{A} \equiv A_{\bar{z}} = -(A_z)^\dagger$.

The gauge potentials for the nonabelian case are of the form $A_i = (-it_a)A_i^a$. We will consider the gauge group $SU(N)$ for simplicity, so that t_a may be taken as $N \times N$ hermitian traceless matrices. For a small neighborhood around $A = 0$, the fields may be considered as Abelian and we expect a result similar to (4.3). We may thus write

$$A \equiv A_z = -\partial_z \Theta + \mathcal{O}(\Theta^2) \tag{4.4}$$

where Θ is also an $N \times N$ traceless matrix. Because it is complex, we may regard it as the group parameter of an element of $SL(N, \mathbb{C})$ (represented as an $N \times N$ matrix). The expression (4.4) is then of the form of a pure gauge near the identity in $SL(N, \mathbb{C})$, i.e., for an element $M = e^\Theta \approx 1 + \Theta$. We can then "integrate" (4.4) (i.e., compose it with a series of infinitesimal group translations in $SL(N, \mathbb{C})$) and write it in the form

$$A = -\partial M \, M^{-1}, \quad M \in SL(N, \mathbb{C}) \tag{4.5}$$

With $A_{\bar{z}} = -(A_z)^\dagger$, the full parametrization is thus

$$\boxed{A = -\partial M \, M^{-1}, \quad \bar{A} = M^{\dagger -1} \bar{\partial} M^\dagger} \tag{4.6}$$

While we have obtained this result for the group $SU(N)$, it is easy to see how it will generalize. For a Lie group G, Θ is combination of the generators of the group with complex coefficients, so the parametrization (4.6) will hold in general with M as an element of the complexification $G^{\mathbb{C}}$ of the group G.

In (4.2), the term $\partial_i \varphi$ denotes the gauge transformation for the group $U(1)$. More generally, for the nonabelian case, gauge transformations take the form[3]

$$M \to g \, M, \quad g \in SU(N) \text{ (or more generally } \in G) \tag{4.7}$$

The gauge invariant degrees of freedom are thus given by

$$H = M^\dagger M \tag{4.8}$$

[3]There are other ways to parametrize A's. One could even use $A_z = -\partial_z \Theta$, without any further terms of order Θ^2. In this case, Θ will transform in a rather complicated way under gauge transformations. The simple transformation law (4.7) is the real advantage of using the $SL(N, \mathbb{C})$ version.

The factors of g and $g^\dagger$ in the transformation of M, $M^\dagger$ cancel out and H is invariant. Since M modulo the $SU(N)$ transformations g define $SL(N,\mathbb{C})/SU(N)$, the gauge-invariant degrees of freedom can be taken as the set of mappings from $\mathbb{R}^2$ to this coset space $SL(N,\mathbb{C})/SU(N)$ (or more generally to $G^{\mathbb{C}}/G$). The hermitian matrix H parametrizes the coset $SL(N,\mathbb{C})/SU(N)$.

The advantage of the parametrization (4.6) is precisely that the gauge transformations take the homogeneous form in (4.7), as left translations by G on the matrix M, so that we can easily identify all gauge-invariant degrees of freedom.

There is another way to argue for the parametrization (4.6). We can obtain a similar parametrization on S^2 viewed as the complex projective space $\mathbb{CP}^1$, and then take a large radius limit to get the result (4.6) for $\mathbb{R}^2$. (The parametrization of gauge fields for this case has been worked out in [10].) The space $\mathbb{CP}^1 \sim S^2$ is equivalent to the coset space $SU(2)/U(1)$. We can thus use an element u of $SU(2)$ as coordinates for $\mathbb{CP}^1$, with the identification $u \sim u\,h$, $h \in U(1) \subset SU(2)$. Local coordinates z, $\bar{z}$ can be related to this using the parametrization

$$u = \frac{1}{\sqrt{1+\bar{z}z}} \begin{pmatrix} 1 & z \\ -\bar{z} & 1 \end{pmatrix} \begin{pmatrix} e^{i\alpha/2} & 0 \\ 0 & e^{-i\alpha/2} \end{pmatrix} \tag{4.9}$$

The $U(1)$ angle α can be eliminated via the identification $u \sim uh$. We can define three coordinates x^a by $u\sigma^3 u^{-1} = -\sigma^a\, x^a$; For the parametrization (4.9),

$$x^1 = \frac{z+\bar{z}}{1+\bar{z}z}, \quad x^2 = i\frac{z-\bar{z}}{1+\bar{z}z}, \quad x^3 = \frac{\bar{z}z-1}{1+\bar{z}z} \tag{4.10}$$

These correspond to the embedding of S^2 in $\mathbb{R}^3$, with a stereographic projection onto the complex plane, with the south pole mapped to $z=0$ and the north pole mapped to $|z| \to \infty$. The coordinates cover S^2 except for a small region around the north pole. (A second coordinate patch can be used around the north pole, by choosing $e^{i\alpha/2} = \sqrt{z/\bar{z}}$ (away from the south pole, so $z \neq 0$). Effectively this amounts to an inversion of z. The two coordinate patches will give full coverage of the sphere.) The metric on the coset space $SU(2)/U(1)$ is the Fubini-Study metric for $\mathbb{CP}^1$ given by

$$ds^2 = \frac{dz\, d\bar{z}}{(1+\bar{z}z)^2} \tag{4.11}$$

We now consider unitary irreducible representations (UIR) of $SU(2)$. A basis for the Lie algebra of $SU(2)$ in the defining 2×2 matrix representation is given by $\sigma_a/2$, so that we may write u as

$$u = e^{i\sigma_a\theta^a/2} \tag{4.12}$$

where the parameters θ^a can be taken as functions of z, $\bar{z}$, α or vice versa. Let T_a denote the generators of the group $\sigma_a/2$ in an arbitrary representation. Then a general UIR is specified by the spin value s, defined by $T_a T_a = s(s+1)$. The matrix corresponding to u is given by

$$\mathcal{D}^{(s)}_{m,m'}(u) = \langle s,m|\,\hat{u}\,|s,m\rangle\,, \qquad \hat{u} = e^{iT_a\theta^a} \tag{4.13}$$

The states within the representation are labeled by m, m' which are the eigenvalues of T^3 and take the values $m, m' = -s, -s+1, \cdots, s$.

The matrix-valued functions $\mathcal{D}^{(s)}_{m,m'}(u)$ form a complete set for $SU(2)$, so that any function on $SU(2)$ can be expanded as

$$f(u) = \sum_{s,m,m'} C^{(s)}_{mm'}\ \langle s,m|\,\hat{u}\,|s,m'\rangle \tag{4.14}$$

The action of the $U(1)$ transformation $u \to uh$, $h = e^{i\theta\sigma^3/2}$ is represented as

$$f(uh) = \sum_{s,m,m'} C^{(s)}_{m,m'}\ \langle s,m|\,\hat{u}e^{iT_3\theta^3}\,|s,m'\rangle = \sum_{s,m,m'} C^{(s)}_{m,m'}\ \langle s,m|\,\hat{u}\,|s,m'\rangle\, e^{im'\theta^3} \tag{4.15}$$

Functions on the coset $SU(2)/U(1)$ must be invariant under these transformations. Therefore they have a similar mode expansion with the state on the right side $|s,m'\rangle$ having $m' = 0$. Thus, a function on $\mathbb{CP}^1$ has the expansion,

$$f(u) = \sum_{s,m} C^{(s)}_{m}\ \langle s,m|\,\hat{u}\,|s,0\rangle \tag{4.16}$$

The coefficients $C^{(s)}_m$ define the function.

To define derivative operators, we define the right translation operators R_a by

$$R_a\, u = u\,\frac{\sigma_a}{2} \tag{4.17}$$

This can be lifted to any representation by using $\hat{u}$ and T_a in this equation. Further, the left-invariant one-forms E^a on $SU(2)$ are given by

$$\begin{aligned} u^{-1}du &= -i\frac{\sigma_a}{2}E^a_k d\theta^k \\ E^1 &= i\frac{dz-d\bar{z}}{1+\bar{z}z}, \qquad E^2 = -\frac{dz+d\bar{z}}{1+\bar{z}z}, \qquad E^3 = i\frac{zd\bar{z}-\bar{z}dz}{1+\bar{z}z} \end{aligned} \tag{4.18}$$

E^1, E^2 are the frame fields for the coset space $\mathbb{CP}^1$. From this equation, we see that we can realize R_a as the differential operators

$$R_a = i(E^{-1})^k_a\frac{\partial}{\partial\theta^k} \tag{4.19}$$

In particular, we find

$$R_+ = (R_1 + iR_2) = (1 + \bar{z}z)\partial, \quad R_- = (R_1 - iR_2) = -(1 + \bar{z}z)\bar{\partial} \tag{4.20}$$

From (4.19) we see that R_3 generates the $U(1)$ transformation on the right of u. It corresponds to the isotropy group and is thus the analog of the Lorentz group for Minkowski space. In particular, while functions are invariant under R_3, vectors should transform nontrivially, with the same transformation properties as $R_\pm$. Since $[R_3, R_\pm] = \pm R_\pm$, a vector corresponding to holomorphic components will have the mode expansion

$$A_+ = \sum_{s,m} a_m^{(s)} \ \langle s, m| \, \hat{u} \, |s, 1\rangle \tag{4.21}$$

Since the state $|s, 1\rangle$ can be obtained from $|s, 0\rangle$ as $|s, 1\rangle \sim R_+ \, |s, 0\rangle$, we can write (4.21) as

$$A_+ = R_+ \sum_{s,m} a_m^{(s)} \ \langle s, m| \, \hat{u} \, |s, 0\rangle = -R_+ \Theta \tag{4.22}$$

where Θ is the function $-\sum_{s,m} a_m^{(s)} \ \langle s, m| \, \hat{u} \, |s, 0\rangle$. This A_+ is written using a tangent frame. Using (4.20) and going to the coordinate frame, (4.22) becomes

$$A = -\partial \, \Theta \tag{4.23}$$

This is adequate for an Abelian gauge potential, with $\bar{A} = -(A)^\dagger$. The generalization to the nonabelian case follows the arguments given after (4.3) and we arrive at

$$A = \partial M M^{-1}, \quad \bar{A} = M^{\dagger -1} \bar{\partial} M^\dagger \tag{4.24}$$

These are still on the space $\mathbb{CP}^1$ in terms of components in the coordinate frame. (For the components in the tangent frame, these will be multiplied by $(1 + \bar{z}z)$.) If we now scale $z \to z/r$ and take the large r limit, $\mathbb{CP}^1$ approximates to the flat space $\mathbb{R}^2$ and we recover the parametrization (4.6) for the flat case as well.

We close this section with a comment on what we shall refer to as the holomorphic ambiguity or holomorphic invariance. From the definition in (4.6) it is clear that, for a given A, M is not unique. It is easy to see that M and $M\bar{V}$, where $\bar{V} = \bar{V}(\bar{z})$ is an $SL(N, \mathbb{C})$-matrix whose matrix elements are antiholomorphic functions, lead to the same potential. Similarly, $M^\dagger$ and $V(z)M^\dagger$ lead to the same $\bar{A}$, where V is holomorphic in its dependence on the coordinates. For the two-sphere or for the Riemann sphere, the only (nonsingular and globally defined) antiholomorphic/holomorphic function is a constant by Liouville's theorem, Thus $\bar{V}$ has to be constant. We can eliminate the

ambiguity by requiring a condition like $M \to 1$ at spatial infinity. In general, this is not adequate. The $(M, M^\dagger)$ or $H = M^\dagger M$ corresponding to given potentials $(A, \bar{A})$ can have singularities. To avoid these and obtain a nonsingular description, one has to resort to a patchwise definition of $(M, M^\dagger)$ with transition functions on the intersections of coordinate patches. Notice that $(A, \bar{A})$ are themselves defined only patchwise in general, with gauge transformations acting as the transitions on intersections. By using H which is gauge-invariant we avoid this issue, but we may still need to modify $(M, M^\dagger)$ or H as we move from one coordinate patch to another. The values on coordinate patches U_1 and U_2 will be related on the intersection by $M_1 = M_2 \bar{V}_{12}$, etc., or $H_1 = V_{12} H_2 \bar{V}_{12}$. Since this is an ambiguity of choice of field variables, all observable results must be invariant under this. In particular, we will choose regularizations in such a way as to preserve this invariance. This holomorphic ambiguity in the choice of H and the need for antiholomorphic/holomorphic transition functions also play a role in connection with the Gribov problem, we discuss this briefly in section 6.

5 *The volume element for the gauge-orbit space

The next logical step for us should be to make the change of variables from A, $\bar{A}$ to M and $M^\dagger$ and obtain the volume element of the configuration space $\mathcal{C}$. Our strategy will be to start with the space of gauge potentials $\mathcal{A}$ and divide out the volume of gauge transformations. (The calculation we present is from [11, 9, 12]. See also [13] for more details regarding regularization.) As mentioned earlier, $\mathcal{A}$ is an affine space and we would expect the metric on this space to be the standard flat Euclidean one. We can confirm that this is indeed the relevant metric for the dynamics by considering the Yang-Mills action. With $A_0 = 0$, we have

$$S_{\text{Y-M}} = \int dt d^2x \, \left[\frac{1}{2} \frac{\partial A_i^a}{\partial t} \frac{\partial A_i^a}{\partial t} - \frac{1}{2} B^2 \right] \tag{5.1}$$

A field theory can be thought of as describing the dynamics of a point-particle moving in an infinite dimensional ambient space of fields. Thus comparing (5.1) to the action for a point-particle, namely,

$$S = \int dt \left[\frac{1}{2} g_{\mu\nu} \frac{dx^\mu}{dt} \frac{dx^\nu}{dt} - V \right], \tag{5.2}$$

we see that (5.1) does indeed correspond to the case where the ambient space has the Euclidean metric

$$ds^2 = \int d^2x \, (\delta A_i^a \, \delta A_i^a) = -8 \int d^2x \, \mathrm{Tr}(\delta A \, \delta \bar{A}) \tag{5.3}$$

This is our starting point. Now we can use the parametrization (4.6) to write

$$\begin{aligned}\delta A &= -\left(\partial(\delta M M^{-1}) + [-\partial M M^{-1}, \delta M M^{-1}]\right) \\ &= -D(\delta M M^{-1}) \\ \delta \bar{A} &= \bar{D}(M^{\dagger -1}\delta M^{\dagger})\end{aligned} \tag{5.4}$$

where D, $\bar{D}$ denote covariant derivatives $D\phi = \partial\phi + [A, \phi]$, $\bar{D}\phi = \bar{\partial}\phi + [\bar{A}, \phi]$. Using these expressions we find

$$\begin{aligned}ds^2 &= 8\int d^2x \,\mathrm{Tr}\left[D(\delta M M^{-1})\bar{D}(M^{\dagger -1}\delta M^{\dagger})\right] \\ &= 8\int d^2x \,\mathrm{Tr}\left[(\delta M M^{-1})(-D\bar{D})(M^{\dagger -1}\delta M^{\dagger})\right]\end{aligned} \tag{5.5}$$

As shown in section 4, M and $M^{\dagger}$ can be thought of as elements of $SL(N, \mathbb{C})$. The Cartan-Killing metric for $SL(N, \mathbb{C})$ viewed as the complexification of $SU(N)$ is of the form $\mathrm{Tr}(\delta M M^{-1}\, M^{\dagger -1}\delta M^{\dagger})$. Extending this to $SL(N, \mathbb{C})$-valued functions on $\mathbb{R}^2$, the metric is given as

$$ds^2_{SL(N,\mathbb{C})} = 2\int d^2x \,\mathrm{Tr}(\delta M M^{-1}\, M^{\dagger -1}\delta M^{\dagger}) \tag{5.6}$$

Given the structure of (5.5), the volume element for $\mathcal{A}$ can be written as

$$d\mu(\mathcal{A}) = \det(-D\bar{D})\, d\mu(M, M^{\dagger}) \tag{5.7}$$

where $d\mu(M, M^{\dagger})$ is the volume element associated with the metric (5.6) for M, $M^{\dagger}$. (There are some constant multiplicative factors which are irrelevant for us, since we will be using this to normalize the wave functions. Any such factor will cancel out in matrix elements.)

There are two further simplifications to be done. We must write $d\mu(M, M^{\dagger})$ in terms of $H = M^{\dagger}M$ and a unitary part which corresponds to the $SU(N)$ gauge degrees of freedom. Secondly, we have to calculate the Jacobian determinant $\det(-D\bar{D})$ arising from the change of variables from A, $\bar{A}$ to M, $M^{\dagger}$.

The volume element for $SL(N, \mathbb{C})$ is given by the top-rank differential form constructed from dMM^{-1} and $M^{\dagger -1}dM^{\dagger}$. It is given by

$$\begin{aligned}dV(M, M^{\dagger}) \;\propto\; & \epsilon_{a_1 \ldots a_n}(dMM^{-1})_{a_1} \wedge \cdots \wedge (dMM^{-1})_{a_n} \\ & \times \epsilon_{b_1 \ldots b_n}(M^{\dagger -1}dM^{\dagger})_{b_1} \wedge \cdots \wedge (M^{\dagger -1}dM^{\dagger})_{b_n}\end{aligned} \tag{5.8}$$

where $n = \dim G = N^2 - 1$. (Again we use a proportionality relationship, some constant numerical factors, which are irrelevant for us, are ignored.) The components indicated are of the form $(dMM^{-1})_a = 2\mathrm{Tr}(t_a dMM^{-1})$, $(M^{\dagger -1}dM^\dagger)_b = 2\mathrm{Tr}(t_b M^{\dagger -1}dM^\dagger)$.

We now use a polar decomposition for the matrices M, $M^\dagger$, given as $M = U\rho$, $M^\dagger = \rho U^\dagger$, where ρ is hermitian and U is unitary. Since gauge transformations act on M as $M^g = gM$, we see that U corresponds to the gauge degree of freedom in M. By direct substitution of $M = U\rho$, (5.8) becomes

$$\begin{aligned} dV(M, M^\dagger) &\propto \epsilon_{a_1 \ldots a_n} (d\rho\rho^{-1} + \rho^{-1}d\rho)_{a_1} \wedge \ldots \wedge (d\rho\rho^{-1} + \rho^{-1}d\rho)_{a_n} \\ &\quad \times \epsilon_{b_1 \ldots b_n} (U^{-1}dU)_{b_1} \wedge \ldots \wedge (U^{-1}dU)_{b_n} \\ &\propto \epsilon_{a_1 \ldots a_n} (H^{-1}dH)_{a_1} \wedge \ldots \wedge (H^{-1}dH)_{a_n} dV_U \end{aligned} \tag{5.9}$$

Here dV_U is the Haar measure for $SU(N)$. If we parametrize H as $H = e^{t_k \varphi^k}$ in terms of the real functions φ^k we can also write the H-dependent terms in (5.9) as

$$\epsilon_{a_1 \ldots a_n} (H^{-1}dH)_{a_1} \ldots (H^{-1}dH)_{a_n} = (\det r)\; d\varphi^1 d\varphi^2 \cdots d\varphi^n \tag{5.10}$$

where $H^{-1}dH = d\varphi^a r_{ak}(\varphi)\, t_k$. Thus is the volume element for $SL(N, \mathbb{C})/SU(N)$ obtained by reduction from the Cartan-Killing metric for $SL(N, \mathbb{C})$.

An important feature of (5.9) is that the volume of $SU(N)$, namely, dV_U factors out from the terms involving H. There is no obstruction to this, because $SL(N, \mathbb{C})/SU(N)$ is a contractible space.

Upon taking the product of dV_U and the expression in (5.10) over all points of space to convert to a functional integration measure for $SL(N, \mathbb{C})$-valued fields, we can write

$$\begin{aligned} d\mu(M, M^\dagger) &= \prod_x dV(M, M^\dagger) = \left[(\det r)\, d\varphi^1 d\varphi^2 \cdots d\varphi^n\right] \prod_x dV_U \\ &= d\mu(H)\, d\mu(U) \end{aligned} \tag{5.11}$$

$d\mu(H) = \prod_x (\det r) d\varphi^1 d\varphi^2 \cdots d\varphi^n$ is the Haar measure for hermitian matrix-valued fields. We also note that $d\mu(U) = \prod_x dV_U$ gives the volume of $\mathcal{G}_*$. The volume element in (5.7) can now be written as

$$d\mu(\mathcal{A}) = \det(-D\bar{D})\, d\mu(H)\, d\mu(U) \tag{5.12}$$

It is now straightforward to factor out the volume of gauge transformations ($d\mu(U)$) and write the volume element for $\mathcal{C} = \mathcal{A}/\mathcal{G}_*$ as

$$d\mu(\mathcal{C}) = \left[d\mu(\mathcal{A})/d\mu(U)\right] = \det(-D\bar{D})\, d\mu(H) \tag{5.13}$$

The real advantage of our parametrization (4.6) is in this expression where we can factor out the volume of gauge transformations exactly. What remains is to calculate the determinant of the operator $(-D\bar{D})$. Towards this, we start with

$$\Gamma = \log\det \bar{D} = \mathrm{Tr}\log \bar{D} \tag{5.14}$$

Taking a variation of $\bar{A}$ we find

$$\delta\Gamma = \mathrm{Tr}(\bar{D}^{-1}\delta\bar{A}) = \int d^2x\, \mathrm{Tr}\left[(\bar{D}^{-1})_{x,y}\delta\bar{A}(y)\right]_{y\to x} \tag{5.15}$$

(Here Tr on the right hand side denotes the trace over the Lie algebra while Tr on the left hand side of (5.15) denotes the full functional trace.) We see from this equation that the result for $\delta\Gamma$ will depend on the coincident point limit of the Green's function $\bar{D}^{-1}$. It is easy to verify that

$$\begin{aligned}
\bar{G}(x,y) &= (\bar{\partial})^{-1}_{x,y} = \frac{1}{\pi(x-y)}, \quad x = x_1 - ix_2,\ y = y_1 - iy_2 \\
G(x,y) &= (\partial)^{-1}_{x,y} = \frac{1}{\pi(\bar{x}-\bar{y})}
\end{aligned} \tag{5.16}$$

For the gauge-covariant Green's functions we then find

$$\begin{aligned}
(\bar{D}^{-1})_{x,y} &= M^{\dagger -1}(x)\left[\frac{1}{\pi(x-y)}\right]M^{\dagger}(y) \\
(D^{-1})_{x,y} &= M(x)\left[\frac{1}{\pi(\bar{x}-\bar{y})}\right]M^{-1}(y)
\end{aligned} \tag{5.17}$$

The coincident point limit of $\bar{D}^{-1}$ is singular and so we need regularized expressions in place of (5.17). We will take up this issue in more detail later, but for now, notice that for small infinitesimal but nonzero separations

$$(\bar{D}^{-1})_{x,y} \approx \frac{1}{\pi(x-y)} + \frac{1}{\pi}\partial M^{\dagger -1}\, M^{\dagger}(y) + \frac{\bar{x}-\bar{y}}{\pi(x-y)}\bar{\partial}M^{\dagger -1}M^{\dagger}(y) + \cdots \tag{5.18}$$

Since $\mathrm{Tr}\,\delta\bar{A} = 0$, the use of this expression in (5.15) gives

$$\delta\Gamma = \frac{1}{\pi}\int d^2x\, \mathrm{Tr}\left[(\partial M^{\dagger -1}M^{\dagger})\delta\bar{A}(y)\right] + \frac{\bar{x}-\bar{y}}{\pi(x-y)}\int d^2x\, \mathrm{Tr}\left[\bar{\partial}M^{\dagger -1}M^{\dagger}\delta\bar{A}(y)\right] + \cdots \tag{5.19}$$

If we now take the limit $y \to x$ in a rotationally symmetric fashion, (so that $(\bar{x}-\bar{y})/(x-y) \to 0$), we find

$$\delta\Gamma = -\frac{1}{\pi}\int d^2x\, \mathrm{Tr}\left[M^{\dagger -1}\partial M^{\dagger})\delta\bar{A}\right]$$

$$= \frac{1}{\pi}\int d^2x\,\mathrm{Tr}\left[\bar{D}(M^{\dagger-1}\partial M^\dagger)\,M^{\dagger-1}\delta M^\dagger\right] \tag{5.20}$$

We define the Wess-Zumino-Witten action for a matrix-valued field M as

$$S_{\rm wzw}(M) = \frac{1}{2\pi}\int \mathrm{Tr}\left(\partial M\,\bar{\partial}M^{-1}\right) + \frac{i}{12\pi}\int \mathrm{Tr}\left(M^{-1}dM\right)^3 \tag{5.21}$$

The first term on the right hand side involves the integral over the 2-manifold while the last term is the integral of the 3-form over a 3-manifold whose boundary is the 2-manifold of interest. By direct calculation

$$S_{\rm wzw}(NM) = S_{\rm wzw}(N) + S_{\rm wzw}(M) - \frac{1}{\pi}\int \mathrm{Tr}\left(N^{-1}\bar{\partial}N\,\partial M M^{-1}\right) \tag{5.22}$$

This result is known as the Polyakov-Wiegmann identity [14]. The key point about it is the chiral splitting in the last term; N has only the antiholomorphic derivative, M has only the holomorphic derivative. By taking $NM \to M^\dagger(1+\theta)$, we find

$$\begin{aligned} S_{\rm wzw}(M^\dagger(1+\theta)) - S_{\rm wzw}(M^\dagger) &= -\frac{1}{\pi}\int \mathrm{Tr}\left(M^{\dagger-1}\bar{\partial}M^\dagger\,\partial\theta\right) \\ &= \frac{1}{\pi}\int \mathrm{Tr}\left(\partial(M^{\dagger-1}\bar{\partial}M^\dagger)\,\theta\right) \\ &= \frac{1}{\pi}\int \mathrm{Tr}\left(\bar{D}(M^{\dagger-1}\partial M^\dagger)\,M^{\dagger-1}\delta M^\dagger\right) \end{aligned} \tag{5.23}$$

where we have used the identity

$$\partial(M^{\dagger-1}\bar{\partial}M^\dagger) - \bar{D}(M^{\dagger-1}\partial M^\dagger) = 0 \tag{5.24}$$

and the fact that $\theta = M^{\dagger-1}\delta M^\dagger$. Comparing with (5.20), we see that we can identify

$$\delta\Gamma = 2\,c_A\,\delta S_{\rm wzw}(M^\dagger) \tag{5.25}$$

where c_A is the value of the quadratic Casimir operator in the adjoint representation. The trace in (5.20) is over the adjoint representation, while we wrote the WZW action using traces in the fundamental representation. The identity $\mathrm{Tr}(t_a t_b)_A = 2\,c_A\,\mathrm{Tr}(t_a t_b)_F$ leads to the factor $2c_A$ in (5.25). The integrated version of (5.25) gives the result

$$\Gamma = \mathrm{Tr}\log\bar{D} = 2\,c_A\,S_{\rm wzw}(M^\dagger), \tag{5.26}$$

up to an additive constant. Although we used a simple expansion of $\bar{D}^{-1}$, what we have is really an anomaly calculation, namely, the change of $\det\bar{D}$ under an $SL(N,\mathbb{C})$ transformation. So, as with anomaly calculations, the answer is robust and is obtained by other regularizations as well. In a similar way we get

$$\mathrm{Tr}\log D = 2\,c_A\,S_{\rm wzw}(M) \tag{5.27}$$

If we write $\mathrm{Tr}\log(-D\bar{D}) = \mathrm{Tr}\log D + \mathrm{Tr}\log\bar{D}$ with (5.26), (5.27), the result is not gauge-invariant. Basically, the regularization we used is not gauge-invariant. However, as with the calculation of effective actions from quantum corrections, changing regularizations is equivalent to adding local counterterms. In the present case we can add the local counterterm

$$S_{\mathrm{counter}} = \frac{2\,c_A}{\pi}\int \mathrm{Tr}(\bar{A}\,A) = -\frac{2\,c_A}{\pi}\int \mathrm{Tr}\left(M^{\dagger -1}\bar{\partial}M^{\dagger}\,\partial M M^{-1}\right) \tag{5.28}$$

With this counterterm, or with the corresponding choice of regularization,

$$\begin{aligned}
\log\det(-D\bar{D}) &= \mathrm{Tr}\log D + \mathrm{Tr}\log\bar{D} + \frac{2\,c_A}{\pi}\int \mathrm{Tr}(\bar{A}\,A) \\
&= 2\,c_A\left[S_{\mathrm{wzw}}(M) + S_{\mathrm{wzw}}(M^{\dagger}) - \frac{1}{\pi}\mathrm{Tr}\left(M^{\dagger -1}\bar{\partial}M^{\dagger}\,\partial M M^{-1}\right)\right] \\
&= 2\,c_A\,S_{\mathrm{wzw}}(M^{\dagger}M) \\
&= 2\,c_A S_{\mathrm{wzw}}(H)
\end{aligned} \tag{5.29}$$

where we have used the Polyakov-Wiegmann identity again to combine terms. Since H is gauge-invariant, we have a gauge-invariant result for the determinant. Since we used the variation of the determinant, this calculation does not fix an overall multiplicative constant for the determinant. The constant however, corresponds to the case of when $M = M^{\dagger} = \mathbb{1}$, i.e., to $\det(-\partial\bar{\partial})$. Combining all results, we can then write

$$\det(-D\bar{D}) = \left[\frac{\det'(-\partial\bar{\partial})}{\int d^2x}\right]^{\dim G} e^{2\,c_A S_{\mathrm{wzw}}(H)} \tag{5.30}$$

The prime on $\det'(-\partial\bar{\partial})$ indicates that the constant modes, which are zero modes of the Laplacian are not to be included in the determinant. The division by $\int d^2x$ is to take account of the normalization of the same zero modes. Using this back in (5.13), we get the volume for the gauge-orbit space as

$$\boxed{d\mu(\mathcal{C}) = \mathcal{N}\,d\mu(H)\,e^{2\,c_A S_{\mathrm{wzw}}(H)}, \qquad \mathcal{N} = \left[\frac{\det'(-\partial\bar{\partial})}{\int d^2x}\right]^{\dim G}} \tag{5.31}$$

A worthwhile remark regarding this result is that $S_{\mathrm{wzw}}(VH\bar{V}) = S_{\mathrm{wzw}}(H)$. This follows from the Polyakov-Wiegmann identity (5.22). We also have $d\mu(VH\bar{V}) = d\mu(H)$, since V, $\bar{V}$ are matrices of unit determinant. Thus the volume element (5.31) has the required holomorphic invariance.

6 The topology and geometry of $\mathcal{C}$

The space of gauge-invariant field configurations $\mathcal{C}$ can be identified as $\mathcal{A}/\mathcal{G}_*$, where $\mathcal{A}$ is the space of gauge potentials (which are Lie-algebra valued 1-forms) and $\mathcal{G}_*$ is the set of gauge transformations, i.e., group elements $g : \mathbb{R}^2 \to G$, wit the condition that $g(x) \to \mathbb{1}$ as $|\vec{x}| \to \infty$. One can think of $\mathcal{A}$ as a fiber bundle with $\mathcal{G}_*$ as the structure group and $\mathcal{C} = \mathcal{A}/\mathcal{G}_*$ as the base manifold,

$$\begin{array}{ccc} \mathcal{G}_* & \longrightarrow & \mathcal{A} \\ & & \downarrow \\ & & \mathcal{C} \;=\; \mathcal{A}/\mathcal{G}_* \end{array} \tag{6.1}$$

These are all infinite dimensional spaces. The topology and geometry of these spaces are clearly important for the study of gauge theories. The bundle structure (6.1) shows that, locally on a patch U of $\mathcal{C}$, we have the product structure $\mathcal{A}_U \sim \mathcal{C}_U \times \mathcal{G}_*$. On the patch U we have a set of gauge potentials (corresponding to points in $\mathcal{C}$) with a fiber corresponding to the orbit of each such configuration by gauge transformations. One can specify the gauge-invariant degrees of freedom by choosing a representative configuration for each orbit; this is the process of gauge-fixing and is equivalent to specifying a section of the bundle. While this can be done on a local patch on $\mathcal{C}$, $\mathcal{A}$ as a $\mathcal{G}_*$-bundle is nontrivial and does not admit a global section. Thus there is no gauge fixing which is valid for all gauge potentials. This is the Gribov problem [15]; for a more general discussion, see also [4].

The nontriviality of the bundle can be seen by a slight variant of the *reductio ad absurdum* argument due to Singer [4]. Assume that we can write $\mathcal{A} = \mathcal{C} \times \mathcal{G}_*$ globally, i.e., for all gauge potentials. As mentioned in section 2, the space $\mathcal{A}$ is an affine space and all homotopy groups of $\mathcal{A}$ are trivial. If the condition $\mathcal{A} = \mathcal{C} \times \mathcal{G}_*$ is correct, then we must have trivial homotopy groups for $\mathcal{C}$ and for $\mathcal{G}_*$. Consider now $\Pi_1(G_*)$. A typical element of this would be a sequence of group elements $g(x_1, x_2, \sigma)$ where σ is a parameter (with values in $[0, 1]$) along the loop of $\mathcal{G}_*$ elements. Specifically, we consider a loop starting and ending at the identity element, which implies that $g(x_1, x_2, \sigma) \to 1$ at $\sigma = 0, 1$. From the definition of $\mathcal{G}_*$ we also have $g(x_1, x_2, \sigma) \to 1$ as $|\vec{x}| \to \infty$. Thus g is a map from a cylinder (coordinatized by x_1, x_2, σ) to G, with $g = 1$ on the boundary. Topologically, this equivalent to maps from a sphere to G,

$$g(x_1, x_2, \sigma) : S^3 \to G \tag{6.2}$$

The homotopy classes of such maps are classified by $\Pi_3(G)$, implying $\Pi_1(\mathcal{G}_*) = \Pi_3(G)$.

This is nontrivial for all nonabelian Lie groups; for simple groups, we have

$$\Pi_3(G) = \begin{cases} \mathbb{Z} & \text{Any simple } G, \text{ except } SO(4) \\ \mathbb{Z}\times\mathbb{Z} & SO(4) \end{cases} \tag{6.3}$$

The nontriviality of $\Pi_1(\mathcal{G}_*)$ shows that the initial assumption that $\mathcal{A} = \mathcal{C}\times\mathcal{G}_*$ cannot be valid. This establishes the nontriviality of the bundle (6.1). There is a Gribov problem for any nonabelian group.

Consider now a two-parameter family of gauge potentials of the form

$$A(x,1,x_2,\sigma,\tau) = \tau\, A(x_1,x_2) + (1-\tau)A^{g_1}(x_1,x_2,\sigma) \tag{6.4}$$

The σ-dependence of the potentials is due to the σ-dependence of g_1 which we take to be a nontrivial element of $\mathcal{G}_*$. Taking σ, τ as coordinates in $\mathcal{A}$, this defines potentials over a disc in $\mathcal{A}$. The potentials on the boundary of the disc are A at $\sigma = 0,1$ and at $\tau = 1$, and A^{g_1} at $\tau = 0$. Since these boundary values are all gauge-equivalent, they correspond to a single point in $\mathcal{C}$, so that the disc is a closed 2-surface in $\mathcal{C}$. If this surface is contractible to a point in $\mathcal{C}$, the pre-image of that point is a disc in $\mathcal{A}$ where all potentials inside are also gauge-equivalent to A, of the form A^g with $g(x_1,x_2,\sigma,\tau)$ such that $g(x_1,x_2,\sigma,0) = g_1(x_1,x_2,\sigma)$ and $g(x_1,x_2,\sigma,1) = 1$. Thus $g(x_1,x_2,\sigma,\tau)$ is a homotopy between $g_1(x_1,x_2,\sigma)$ and the identity, which is impossible since g_1 is a nontrivial element of $\Pi_3(G)$. The conclusion is that the closed 2-surface in $\mathcal{C}$ is not contractible. Rather than this long argument, we could also have used exact homotopy sequence

$$\begin{array}{ccccccc} \Pi_2(\mathcal{A}) & \to & \Pi_2(\mathcal{C}) & \to & \Pi_1(\mathcal{G}_*) & \to & \Pi_1(\mathcal{A}) \\ 0 & \to & \Pi_2(\mathcal{C}) & = & \Pi_1(\mathcal{G}_*) & \to & 0 \end{array} \tag{6.5}$$

to arrive at the same conclusion. The implication of the nontrivial nature of the bundle at the level of using $\mathcal{A}$ is the Gribov problem and the impossibility of a global section. At the level of directly using $\mathcal{C}$, it is manifest in the nontrivial topology of $\mathcal{C}$, the lowest dimensional such feature being $\Pi_2(\mathcal{C}) \neq 0$.

Our aim now is to construct an example of the set of configurations which form a noncontractible two-surface, i.e., a nontrivial element of $\Pi_2(\mathcal{C})$. (This discussion follows [12].) The winding number, which we may take as characterizing the element of $\Pi_2(\mathcal{C})$ can be related to the instanton number of a four-dimensional gauge theory. This can be seen as follows. In addition to the homotopy group $\Pi_2(\mathcal{C})$ being nontrivial, the second cohomology group of $\mathcal{C}$ is nontrivial as well. Thus there is a closed but not exact two-form on $\mathcal{C}$. In terms of the potentials, the generating element of this cohomology can

be written as

$$\Omega = \frac{1}{4\pi} \int \mathrm{tr}(\delta A \, \delta A) \tag{6.6}$$

Here A is a one-form on the spatial manifold, δ denotes the exterior derivative on $\mathcal{A}$. If we use w, $\bar{w}$ to denote the coordinates along the two-surface in $\mathcal{C}$, δ is given by $\delta = dw\partial_w + d\bar{w}\partial_{\bar{w}}$. The integration in (6.6) is over the spatial manifold, making Ω a two-form on $\mathcal{C}$. The integral of Ω over the closed noncontractible two-surface in $\mathcal{C}$ will give a winding number ν by $\int \Omega = 2\pi\nu$.

The two-surface in $\mathcal{C}$ (with the coordinates w, $\bar{w}$) and the two-dimensional spatial manifold can be considered together as a four-dimensional space. The instanton number on this 4d-space is given by

$$\nu = \frac{1}{8\pi^2} \int \mathrm{Tr}(\tilde{F} \, \tilde{F}) \tag{6.7}$$

where $\tilde{F} = (d+\delta)\tilde{A} + \tilde{A}\tilde{A}$. The operator $(d+\delta)$ denotes the full exterior derivative on the four-dimensional space and $\tilde{A}$ is the four-dimensional gauge potential. The 4d-potential can be constructed from the two-dimensional potential A as $\tilde{A} = A + A'$, where we take A' to be given in terms of M, $M^\dagger$ by

$$A' = -\partial_w M \; M^{-1} \, dw + (M^\dagger)^{-1} \partial_{\bar{w}} M^\dagger d\bar{w} \tag{6.8}$$

While A transforms as a connection under gauge transformations $g(x)$, A' is gauge-covariant since $g(x)$ does not depend on w, $\bar{w}$. In other words, $\delta g = 0$ for gauge transformations. The field strength can be written out as

$$\tilde{F} = F + F' + \delta A + DA', \qquad DA' = dA" + AA' + A'A \tag{6.9}$$

It is then easy to see that

$$\mathrm{tr}(\tilde{F}\tilde{F}) = \mathrm{tr}(\delta A \, \delta A) + d\left[\mathrm{tr}\left(A' \, DA' + 2 \, \delta A \, A'\right)\right] + \delta\left[\mathrm{tr}(F \, A')\right] \tag{6.10}$$

In integrating this expression over the spatial manifold and the internal closed two-surface, the terms which are total derivatives give zero. (Notice that the integrands are gauge-invariant, so there is not problem of the potentials being patchwise defined with transitional gauge transformations on the overlap regions. Therefore the total derivatives indeed integrate to zero.) From the integral of (6.10), we see that

$$\nu = \int \frac{\Omega}{2\pi} \tag{6.11}$$

where Ω is as given in (6.6). Using the expression for A' from (6.8) and the parametrization (4.6) for the spatial components, Ω takes the form

$$\Omega = \frac{1}{2\pi}\int \mathrm{Tr}\left[\partial(H^{-1}\bar{\partial}H)\delta(H^{-1}\bar{\delta}H) \; + \partial(H^{-1}\bar{\delta}H)\delta(H^{-1}\bar{\partial}H)\right] \tag{6.12}$$

We can exploit this connection between the two-form Ω on $\mathcal{C}$ and the instanton number to construct an example of the noncontractible two-surface of configurations. Towards this, we write the standard instanton in $\mathbb{R}^4$ using complex coordinates and interpreting one pair of complex coordinates as internal coordinates parametrizing the two-surface in $\mathcal{C}$. Explicitly this gives the expression

$$H = \exp(2fJ^3) \; = \cosh 2f \; + J^3 \sinh 2f \tag{6.13}$$

Here $J^3 = \sigma \cdot n$ with σ^a, $a = 1, 2, 3$, being the Pauli matrices and the unit vector n^a is given by

$$n^a = \frac{1}{(\bar{z}z + \bar{w}w)}\,(\bar{z}w + \bar{w}z,\; i(\bar{w}z - \bar{z}w),\; \bar{z}z - \bar{w}w) \tag{6.14}$$

The function f is given by

$$f = \frac{1}{2}\log\left(\frac{\bar{z}z + \bar{w}w + \mu^2}{\bar{z}z + \bar{w}w}\right) \tag{6.15}$$

μ is a scale parameter and $(w, \bar{w})$ parametrize the two-surface in $\mathcal{C}$. Using the formula (6.12), it is easy to verify that

$$\nu = \int \frac{\Omega}{2\pi} = 3 \tag{6.16}$$

for this set of configurations (6.13). Therefore (6.13) does correspond to a noncontractible two-surface in $\mathcal{C}$.

We have specified the configurations (6.13) directly in terms of the gauge-invariant variable H, so there is no Gribov problem *per se*. However, the existence of nontrivial elements in $\Pi_2(\mathcal{C})$ means that we have to choose coordinate patches (in $\mathcal{C}$) to specify the whole set of configurations in a nonsingular way. This will be related to the freedom of the holomorphic transformations mentioned earlier. We will illustrate this in our explicit example now.

Notice that, as $\bar{z}z \to \infty$, $H \to 1$. Further, for almost all $w, \bar{w}$, H is nonsingular; however, the particular configuration at $w = 0$ has a singularity at the spatial point $z = 0$. We can change the position of this singularity by transformations of the type

$H \to VH\bar{V}$, where V is holomorphic in z. Consider the configuration for which $w = \bar{w} = 0$; it is given by

$$\begin{aligned} f &= \frac{1}{2}\log(\bar{z}z + \mu^2/\bar{z}z) \\ H &= \exp(2f\sigma^3) = \exp\left(\sigma^3\left[\log(\bar{z}z+\mu^2) - \log z - \log\bar{z}\right]\right) \\ &= \exp\left(-\sigma^3\log z\right)\,\exp\left(\sigma^3\log(\bar{z}z+\mu^2)\right)\,\exp\left(-\sigma^3\log\bar{z}\right) \end{aligned} \tag{6.17}$$

Using $V = e^{\sigma^3\log(z/z-a)}$, we find

$$V\,H\,\bar{V} = \exp\left(\sigma^3[\log(\bar{z}z+\mu^2) - \log(z-a) - \log(\bar{z}-\bar{a})]\right) \tag{6.18}$$

We see that the singularity has been shifted from $z = 0$ to $z = a$.

This tells us that, at least for configurations of the type given here, we can specify field configurations by nonsingular formulae for H in different coordinate patches with transition relations given by transformations of the form $H \to V(z)\,H\,\bar{V}(\bar{z})$. This shows the importance of the holomorphic invariance.

Since the singularity in our example is at a point, namely at $w = 0$, for this specific case, even if we simply use the formulae (6.13-6.15) with the coordinate singularity, the effect on the quantum wave functions is minimal. This is something that can be verified in terms of the wave functions given later. Also, as remarked earlier, the WZW-action $S_{\rm wzw}(H)$ is invariant under transformations of the type $H \to VH\bar{V}$ and therefore we do not expect any pathology for the wave function. Explicitly, for the set of configurations (6.13), the WZW-action is given by

$$\begin{aligned} \mathcal{S}(H) &= \frac{5\mu^2+4w\bar{w}}{w\bar{w}+\mu^2} - \frac{3\mu^2+4w\bar{w}}{\mu^2}\log\left[\frac{\mu^2+w\bar{w}}{w\bar{w}}\right] \\ &= 5 + 3\log(w\bar{w}) + \mathcal{O}(w\bar{w}) \end{aligned} \tag{6.19}$$

When $w \to 0$, $\exp(2c_A\mathcal{S})$ vanishes as $(w\bar{w})^{6c_A}$. The coordinate singularity does not lead to difficulties, at least for this case.

7 *The Hamiltonian

In section 5, we obtained the volume element of the gauge orbit space. As discussed in section 2, the wave functions for the physical states must obey the invariance condition $G_0(\theta)\Psi = 0$. The inner product is then given by integration with $d\mu(\mathcal{C})$, see (2.38), and

it can be written out as

$$\langle 1|2\rangle = \int d\mu(H)\, e^{2\,c_A S_{\rm wzw}(H)}\, \Psi_1^*\, \Psi_2 \tag{7.1}$$

The Hamiltonian has the form given in (2.39) or (2.41). Since it involves products of operators at the same point, a regularized version has to be defined, consistent with all the symmetries which have to be maintained. (The construction of the Hamiltonian, including regularization issues, is discussed in detail in [13].) Towards this, we first define translation operators on the $SL(N,\mathbb{C})$ group elements M and $M^\dagger$ by

$$\begin{aligned} [p_a(\vec{x}), M(\vec{y})] &= M(\vec{y})(-it_a)\,\delta^{(2)}(\vec{x}-\vec{y}) \\ [\bar{p}_a(\vec{x}), M^\dagger(\vec{y})] &= (-it_a)M^\dagger(\vec{y})\,\delta^{(2)}(\vec{x}-\vec{y}) \end{aligned} \tag{7.2}$$

Here M and $M^\dagger$ are taken to be $N\times N$ matrices, corresponding to the fundamental representation of $SL(N,\mathbb{C})$. Correspondingly, t_a are $N\times N$ matrices which form a basis for the Lie algebra of $SU(N)$. We take them to be normalized as $\mathrm{Tr}(t_a t_b) = \frac{1}{2}\delta_{ab}$. Parametrizing $M, M^\dagger$ in terms of $\Theta^a(\vec{x}),\ \bar{\Theta}^a(\vec{x})$ respectively, we can write

$$M^{-1}\delta M \ = \delta\Theta^a R_{ab}(\Theta)\, t_b, \qquad \delta M^\dagger M^{\dagger -1} \ = \delta\bar{\Theta}^a R^*_{ab}(\bar{\Theta}) t_b \tag{7.3}$$

These equations define $R_{ab}(\Theta)$ and $R^*_{ab}(\bar{\Theta})$.[4] From the parametrization of the gauge potentials, we also have

$$\delta A = -D(\delta M M^{-1}), \qquad \delta\bar{A} = \bar{D}(M^{\dagger -1}\delta M^\dagger) \tag{7.4}$$

Using these relations, we can solve for Θ^a and $\bar{\Theta}^a$ in terms of δA^a and $\delta\bar{A}^a$ and identify the functional derivatives (which are the electric fields up to a factor of e^2) as

$$\begin{aligned} -\frac{i}{2}\frac{\delta}{\delta\bar{A}_k(\vec{x})} &= \frac{i}{2}M^\dagger_{ak}(\vec{x})\int_y \bar{G}(\vec{x},\vec{y})\,\bar{p}_a(\vec{y}) \\ -\frac{i}{2}\frac{\delta}{\delta A_k(\vec{x})} &= -\frac{i}{2}M_{ka}(\vec{x})\int_y G(\vec{x},\vec{y})\,p_a(\vec{y}) \end{aligned} \tag{7.5}$$

where $M_{ab} = 2\,\mathrm{Tr}(t^a M t^b M^{-1})$ is the adjoint representation of M. The kinetic energy operator in (2.41) can now be written down as

$$\begin{aligned} T &= -\frac{e^2}{2}\Delta_{\mathcal{C}} = -\frac{e^2}{2}\int_x \frac{\delta^2}{\delta A_k(\vec{x})\delta\bar{A}_k(\vec{x})} \\ &= \frac{e^2}{2}\int_x K_{ab}(\vec{x})(\bar{G}\bar{p}_a)(\vec{x})(Gp_b)(\vec{x}) \end{aligned} \tag{7.6}$$

[4] They are basically the frame fields on the group $SL(N,\mathbb{C})$.

where $K_{ab} = M^\dagger_{ak} M_{kb} = 2\,\mathrm{Tr}(t^a H t^b H^{-1})$ and $Gp_b(\vec{x}) \equiv \int_y G(\vec{x},\vec{y}) p_b(\vec{y})$, etc.

Another way to write T, which shows explicitly that it is a symmetric operator, is

$$\begin{aligned}
\langle 1|T|2\rangle &= \frac{e^2}{4}\int d\mu(H) e^{2c_A S_{\rm wzw}(H)} \left[\overline{Gp_a\Psi_1}\, K_{ab}\,(Gp_b\Psi_2) + \overline{\bar{G}\bar{p}_a\Psi_1}\, K_{ba}\,(\bar{G}\bar{p}_b\Psi_2)\right] \\
&= \langle T1|2\rangle \\
&= \frac{e^2}{4}\int d\mu(H) e^{2c_A S_{\rm wzw}(H)} \Psi_1^* \Bigg[\int_x e^{-2c_A S_{\rm wzw}(H)} \Big[\bar{G}\bar{p}_a(\vec{x}) K_{ab}(\vec{x})\, e^{2c_A S_{\rm wzw}(H)} Gp_b(\vec{x}) \\
&\qquad\qquad + Gp_a(\vec{x}) K_{ba}(\vec{x})\, e^{2c_A S_{\rm wzw}(H)} \bar{G}\bar{p}_b(\vec{x})\Big]\Bigg]\Psi_2 \qquad (7.7)
\end{aligned}$$

In this expression, if we try to move $\bar{G}\bar{p}_a$ through $K_{ab}\, e^{2c_A\mathcal{S}}$ to act on $Gp_b(\vec{x})\Psi_2$, we will encounter the singular commutator $[\bar{G}\bar{p}_a(\vec{x}), K_{ab}(\vec{x})]$. The regularized version of (7.7) should be such that it agrees with (7.6).

The regularization of a field theory in the Schrödinger formulation in terms of the Hamiltonian and wave functions is more involved (and less well-known) than the case of covariant perturbation theory. We will discuss this and related issues in some detail separately in Appendix B. But for now, we make an observation about observables and the wave function. Since we are considering the gauge theory without matter fields, the Wilson loop operators $W(C)$, over all closed curves C, constitute a complete (in fact, overcomplete) set of observables. These are given by

$$W(C) = \mathrm{Tr}\,\mathcal{P} e^{-\oint_C A dz + \bar{A} d\bar{z}} = \mathrm{Tr}\,\mathcal{P} e^{\oint_C \partial H H^{-1} dz} = \mathrm{Tr}\,\mathcal{P} e^{(\pi/c_A)\oint_C J dz} \qquad (7.8)$$

Here $\mathcal{P}$ signifies path-ordering of the matrices in the exponent and J is the current given by

$$J = \frac{c_A}{\pi}\partial H\, H^{-1} \qquad (7.9)$$

This is the current associated with the WZW action $S_{\rm wzw}(H)$ which is part of the volume of the gauge orbit space. We are starting with wave functions which are functions of A, $\bar{A}$, or equivalently, M and $M^\dagger$. Since they must be gauge-invariant by the Gauss law condition $G_0(\theta)\Psi = 0$, we can take them to be functions of H. But since all observables can be given in terms of J, we can further assume Ψ's to be functions of J. Thus it is advantageous to express the Hamiltonian entirely in terms of J. The kinetic energy operator then takes the form

$$\boxed{\begin{aligned}
T\,\Psi(J) &= m\left[\int J_a(\vec{z})\frac{\delta}{\delta J_a(\vec{z})} + \int_{z,w}\Omega_{ab}(\vec{z},\vec{w})\frac{\delta}{\delta J_a(\vec{w})}\frac{\delta}{\delta J_b(\vec{z})}\right]\Psi(J) \\
\Omega_{ab}(\vec{z},\vec{x}) &= \left(\frac{c_A}{\pi^2}\frac{1}{(z-w)^2} + i f_{abc}\frac{J^c(w)}{\pi(z-w)}\right) + \mathcal{O}(\epsilon)
\end{aligned}} \qquad (7.10)$$

where $m = e^2 c_A/(2\pi)$. We have done the regularization using ϵ as a short-distance cutoff. Although a detailed discussion of the regularization will be in Appendix B, we will just state here that our regularization amounts to a point-splitting where the Dirac δ-function is replaced by

$$\sigma(\vec{x},\vec{y};\epsilon) = \frac{e^{-|\vec{x}-\vec{y}|^2/\epsilon}}{\pi\epsilon} \tag{7.11}$$

This shows that the regularization parameter ϵ is essentially a short-distance cutoff. We recover the δ-function as $\epsilon \to 0$. This has to be augmented by certain factors involving K_{ab} to preserve various invariances, as discussed later. The terms displayed in (7.10) are the finite regularized terms, with $\mathcal{O}(\epsilon)$ indicating terms which are negligible as the cutoff $\epsilon \to 0$.

The two terms appearing in the expression for T are of some interest in their own right. The first term is essentially due to the anomaly in the two-dimensional case. We can see this by calculating

$$\begin{aligned} T\ J_a(\vec{x}) &= -\frac{e^2}{2}\int d^2y \frac{\delta^2 J_a(\vec{x})}{\delta\bar{A}^b(\vec{y})\delta A^b(\vec{y})} = \frac{e^2 c_A}{2\pi} M^\dagger_{am}\mathrm{Tr}\left[T^m \bar{D}^{-1}(\vec{y},\vec{x})\right]_{\vec{y}\to\vec{x}} \\ &= m\ J_a(\vec{x}) \end{aligned} \tag{7.12}$$

The coincident point limit of $\bar{D}^{-1}(\vec{y},\vec{x})$ which appears here is exactly the same as in the calculation of the gauge-invariant measure of integration. Calculating it exactly as in that case, i.e., using (5.19), leads to the second line in (7.12). The result in (7.10) then follows by the chain rule for functional differentiation.

The second term involving $\Omega_{ab}(\vec{z},\vec{x})$ gives the singular pole terms in the operator product expansion for the current of the WZW model $S_{\mathrm{wzw}}(H)$, from a conformal field theory point of view. Its appearance is again very natural.

There is another way to obtain the result (7.10) for the operator T, which is also illuminating in some ways. For this we first write the Gauss law operator, defined in (2.23) as

$$\begin{aligned} G_0(\theta) &= \int d\mu\, \theta^a \frac{(D\cdot E)^a}{e^2} = \int d\mu\, \theta^a I^a \\ I^a &= \frac{(D\cdot E)^a}{e^2} = \frac{2}{e^2}(D\bar{E} + \bar{D}E)^a \end{aligned} \tag{7.13}$$

The idea then is to regard $\bar{E}^a$ and I^a as independent invariables and eliminate E^a. We

can solve for E in terms of $(\bar{E}^a, I^a)$ as

$$E(\vec{x}) = \int_y (\bar{D}^{-1})_{x,y} \left(\frac{e^2}{2} I - D\bar{E} \right) \tag{7.14}$$

The fundamental commutation rules are

$$\begin{aligned} [E^a(\vec{x}), \bar{A}^b(\vec{y})] &= [\bar{E}^a(\vec{x}), A^b(\vec{y})] = -\frac{ie^2}{2}\delta^{ab}\delta(\vec{x}-\vec{y}) \\ [I^a(\vec{x}), A^b(\vec{y})] &= -iD_x^{ab}\delta(\vec{x}-\vec{y}) \end{aligned} \tag{7.15}$$

It is easy to check that this is consistent with the solution for E, so that we may take (7.14) as an operator identity. We can thus write the kinetic energy operator as

$$T = \frac{2}{e^2}\int_x E^a(\vec{x})\bar{E}^a(\vec{x}) \; = \frac{2}{e^2}\int_{x,y} \left[(\bar{D}^{-1})^{ab}(\vec{x},\vec{y}) \left(\frac{e^2}{2} I - D\bar{E} \right)^b (\vec{y}) \right] \; \bar{E}^a(\vec{x}) \tag{7.16}$$

The idea is that we can now move the Gauss law operator to the right end of this expression; this gives

$$\begin{aligned} \frac{1}{2}\int_y (\bar{D}^{-1})^{ab}(\vec{x},\vec{y}) I^b(\vec{y})\bar{E}^a(\vec{x}) &= \frac{1}{2}\int_y (\bar{D}^{-1})^{ab}(\vec{x},\vec{y})\bar{E}^a(\vec{x}) I^b(\vec{y}) \\ &\quad -\frac{i}{2}\int_y (\bar{D}^{-1})^{ab}(\vec{x},\vec{y}) f^{abc}\bar{E}^c(\vec{y})\delta(\vec{x}-\vec{y}) \\ &= \frac{1}{2}\int_y (\bar{D}^{-1})^{ab}(\vec{x},\vec{y})\bar{E}^a(\vec{x}) I^b(\vec{y}) \\ &\quad -\frac{1}{2}\mathrm{Tr}\left[T^c (\bar{D}^{-1})(\vec{x},\vec{y}) \right]_{\vec{y}\to\vec{x}} \bar{E}^c(\vec{x}) \end{aligned} \tag{7.17}$$

Notice that, once agin, the coincident point involved is exactly what we had for the calculation of the volume element and in (7.12) as well. We can evaluate it as done previously to write

$$-\frac{1}{2}\mathrm{Tr}\left[T^c (\bar{D}^{-1})(\vec{x},\vec{y}) \right]_{\vec{y}\to\vec{x}} \bar{E}^c(\vec{x}) = \frac{ic_A}{2\pi}\left(A - M^{\dagger -1}\partial M^\dagger \right)^c \tag{7.18}$$

We can now write T from (7.16) as

$$\begin{aligned} T &= 2im\int_x (A - M^{\dagger-1}\partial M^\dagger)^a(\vec{x})\bar{E}^a(\vec{x}) \; - \; 2e^2\int_{x,y} \left[(\bar{D}^{-1})(\vec{x},\vec{y}) D\bar{E}(\vec{y}) \right]^a \bar{E}^a(\vec{x}) \\ &\quad + e^2\int_{x,y} (\bar{D}^{-1})^{ab}(\vec{x},\vec{y})\bar{E}^a(\vec{x}) I^b(\vec{y}) \end{aligned} \tag{7.19}$$

where $m = e^2 c_A/(2\pi)$. On physical states annihilated by the Gauss law, the last term gives zero. This simplification for T did not require any wave function, and is valid for

both the E-diagonal and A-diagonal representation. We can reduce this expression for T further if we choose wave functions in the A-representation. Towards this, write the parametrization of the fields as

$$A = M^{\dagger -1}(-\partial H H^{-1})M^{\dagger} + M^{\dagger -1}\partial M^{\dagger}, \quad \bar{A} = M^{\dagger -1}\bar{\partial} M^{\dagger} \tag{7.20}$$

This displays our parametrization for $(A, \bar{A})$ as a complex gauge transformation of $(-\partial H H^{-1}, 0)$, by the $SL(N, \mathbb{C})$ group element $M^{\dagger}$. We may therefore take the wave function $\Psi(A, \bar{A})$ as a function of $J = (c_A/\pi)\partial H H^{-1}$ and $M^{\dagger}$. Since a change of $M^{\dagger}$ is equivalent to an $SL(N, \mathbb{C})$ gauge transformation, we may write, for infinitesimal θ,

$$\Psi(M^{\dagger} e^{\theta}, J) \approx \Psi(M^{\dagger}, J) + \int \theta^a I^a \ \Psi(M^{\dagger}, J) \tag{7.21}$$

This shows that, even though θ is complex, the Gauss law condition is enough for us to conclude that

$$\Psi(M^{\dagger} e^{\theta}, J) = \Psi(M^{\dagger}, J) \tag{7.22}$$

We see that, by a sequence of such transformations, we can set $M^{\dagger}$ to the identity. (In two spatial dimensions, all configurations $M^{\dagger}$ are homotopic to the identity, since $\Pi_2(SL(N, \mathbb{C})) = 0$. So there is no obstruction to this procedure.) In other words, we can take the physical wave functions to be functions of J. In this case, we can take $A = -\partial H H^{-1}$, $\bar{E} \sim (\delta/\delta J)$ and T simplifies to the expression given in (7.10).

It is a simpler task to write the potential energy term in terms of the current J. From the structure of the parametrization of fields as in (7.20), we see that

$$B^a t_a = M^{\dagger -1} \left[-\frac{2\pi}{c_A} \bar{\partial} J^a t_a \right] M^{\dagger} \tag{7.23}$$

so that we have

$$\int \frac{B^a B^a}{2e^2} = \frac{\pi}{m c_A} \int_x : \bar{\partial} J^a(\vec{x}) \bar{\partial} J^a(\vec{x}) : \tag{7.24}$$

The normal-ordering indicates the subtraction of the short-distance singularity. We can write this more explicitly as

$$\int \frac{B^a B^a}{2e^2} = \frac{\pi}{m c_A} \left[\int_{x,y} \sigma(\vec{x}, \vec{y}; \epsilon) \bar{\partial} J_a(\vec{x}) \left[K(x, \bar{y}) K^{-1}(y, \bar{y}) \right]_{ab} \bar{\partial} J_b(\vec{y}) - \frac{c_A \mathrm{dim} G}{\pi^2 \epsilon^2} \right] \tag{7.25}$$

Finally, we can combine the expression for T from (7.10) and the potential energy from (7.24) to write the full Hamiltonian as

$$\boxed{\begin{aligned} \mathcal{H} = m \left[\int J_a(\vec{z}) \frac{\delta}{\delta J_a(\vec{z})} + \int_{z,w} \Omega_{ab}(\vec{z}, \vec{w}) \frac{\delta}{\delta J_a(\vec{w})} \frac{\delta}{\delta J_b(\vec{z})} \right] \\ + \frac{\pi}{m c_A} \int_x : \bar{\partial} J^a(\vec{x}) \bar{\partial} J^a(\vec{x}) : + \mathcal{O}(\epsilon) \end{aligned}} \tag{7.26}$$

8 *A propagator mass for the gluon

We have obtained the Hamiltonian in terms of the current J. We also have the volume element for the gauge orbit space, which is what defines the inner product for wave functions. Thus we are now in a position to write down the Schrödinger equation and solve it, in some suitable approximation. However, before we do that, we will discuss the theory from the perturbative limit as it can provide some useful insights. From standard perturbation theory in terms of Feynman diagrams, the effective action Γ (which is the generating function for one-particle irreducible vertices), calculated to one-loop order, has the form

$$\Gamma = -\frac{1}{4e^2}\int F^a_{\mu\nu}(x)\left[1 - \frac{7e^2 c_A}{64}\frac{1}{\sqrt{-\nabla^2}}\right]_{x,y} F^{a\mu\nu}(y) + \cdots \tag{8.1}$$

There is no renormalization of the coupling constant, so predictions for string tension, masses, etc. can be made without worrying about the scale at which e^2 is to be defined. Secondly, the correction shows clearly that the expansion parameter is $e^2/\sqrt{-\nabla^2} \sim (e^2/k)$ where k is the momentum of the field $F_{\mu\nu}$. Thus in a Fourier decomposition, the modes of the field for momenta high compared to e^2 can be treated perturbatively, while the low momentum modes with $k \ll e^2$ have to be treated nonperturbatively. There is no real expansion parameter for the theory as a whole, e^2 is only a marker to signify which modes can be, and which modes cannot be, treated perturbatively.

Based on our Hamiltonian, we can take this a step further and consider an improved perturbation theory where a partial resummation has been carried out. (This has been discussed in [12, 13].) Towards this, we write $H = e^{t^a\varphi^a}$ in terms a set of fields φ^a. Then we have

$$\begin{aligned} J &= -\frac{c_A}{\pi}\partial H\, H^{-1} = -\frac{c_A}{\pi}\int_0^1 d\alpha\, e^{\alpha t\cdot\varphi}(t\cdot\partial\varphi)e^{-\alpha t\cdot\varphi} \\ &\approx -\frac{c_A}{\pi}t_a\left[\partial\varphi^a + \frac{i}{2}f^{abc}\varphi^b\partial\varphi^c + \cdots\right] \end{aligned} \tag{8.2}$$

In perturbation theory, interaction vertices arise from commutators and carry factors of f^{abc}. With this in mind, we can consider a simplification of the Hamiltonian where we keep only the leading term in (8.2). With $J^a \simeq \frac{c_A}{\pi}\partial\varphi^a$, the Hamiltonian has the expansion

$$\mathcal{H} \simeq m\left[\int \varphi_a\frac{\delta}{\delta\varphi_a} + \frac{\pi}{c_A}\int \Omega(\vec{x},\vec{y})\frac{\delta}{\delta\varphi_a(\vec{x})}\frac{\delta}{\delta\varphi_a(\vec{y})}\right]$$

$$+\frac{c_A}{m\pi}\int \partial\varphi_a(-\partial\bar{\partial})\bar{\partial}\varphi_a \ +\mathcal{O}(\varphi^3) \tag{8.3}$$

$$\Omega(\vec{x},\vec{y}) \ = \ -\int \frac{d^2k}{(2\pi)^2}\, e^{ik\cdot(x-y)}\frac{1}{k\bar{k}}$$

The first term in the Hamiltonian, namely, $\int \varphi_a \delta/\delta\varphi_a$ shows that every φ in a wave function will get a contribution of m to the energy. This is basically the origin of the mass gap. To the same order, with $H = e^{t_a\varphi^a} \approx 1 + t_a\varphi^a$, the volume element becomes

$$d\mu(\mathcal{C}) = d\mu(H)\, e^{2c_A S_{\rm wzw}(H)} \simeq [d\varphi]\, e^{-\frac{c_A}{2\pi}\int \partial\varphi^a\bar{\partial}\varphi^a}\ \big(1\ +\mathcal{O}(\varphi^3)\big) \tag{8.4}$$

The exponential factor with the WZW action, can be absorbed into the wave function by defining

$$\Phi = e^{c_A S_{\rm wzw}(H)}\Psi \simeq e^{-\frac{c_A}{4\pi}\int \partial\varphi\bar{\partial}\varphi}\Psi \tag{8.5}$$

In terms of the wave functions Φ, the inner product is given by

$$\langle 1|2\rangle \approx \int [d\varphi]\ \Phi_1^*(H)\,\Phi_2(H) \tag{8.6}$$

We defined $\mathcal{H}$ to act on the Ψ's. For the wave functions Φ, the Hamiltonian should be

$$\boxed{\begin{aligned}\mathcal{H}' &= e^{c_A S_{\rm wzw}(H)}\,\mathcal{H}\, e^{c_A S_{\rm wzw}(H)} \simeq e^{-\frac{c_A}{4\pi}\int \partial\varphi\bar{\partial}\varphi}\,\mathcal{H}\, e^{\frac{c_A}{4\pi}\int \partial\varphi\bar{\partial}\varphi} \\ &\simeq \frac{1}{2}\int_x\left[-\frac{\delta^2}{\delta\phi_a^2(\vec{x})} + \phi_a(\vec{x})\big(m^2-\nabla^2\big)\phi_a(\vec{x})\right]+\cdots\end{aligned}} \tag{8.7}$$

where $\phi_a(\vec{k}) = \sqrt{c_A k\bar{k}/(2\pi m)}\ \varphi_a(\vec{k})$. This is exactly the free part of a Hamiltonian for a field of mass $m = e^2c_A/2\pi$. Thus, to this order, the gauge-invariant version of the gluons are represented by ϕ_a and behave as a field of mass m. It is then straightforward to realize that the propagator corresponding to ϕ_a is[5]

$$\langle \mathcal{T}\phi_a(x)\,\phi_b(y)\rangle = \delta_{ab}\int \frac{d^3k}{(2\pi)^3}e^{-ik\cdot(x-y)}\frac{i}{k^2-m^2+i\epsilon} \tag{8.8}$$

Since $m = (e^2c_A/2\pi)$, this is not the result at the lowest order in the usual perturbation theory. We must expand this in powers of m to make the comparison. The terms of order $(m^2)^n$ in such an expansion may be viewed as arising from the diagrams of order $(e^2)^n$ in perturbation theory, so that (8.8) can be taken to be the result of a selective resummation of the perturbation expansion, where a set of specific terms (and, in fact, a particular kinematic limit of such terms) are summed up.

[5]Here $\mathcal{T}$ denotes the usual time-ordering.

Thus in setting up perturbation theory using our Hamiltonian and expanding H in powers of φ to any order, what we get is an "improved" perturbation theory, where a selective resummation has been done even at the lowest order. The theory at this lowest order is a free scalar field theory of mass m. This does give a useful starting point for some calculations. In fact, we will use this version later to calculate the Casimir energy for a parallel plate geometry in the nonabelian theory and compare with lattice simulations.

9 *The Schrödinger equation: An expansion scheme

In this section, we shall return to the version of the Hamiltonian in terms of the currents, write down the Schrödinger equation and develop a recursive scheme for solving it for the vacuum wave function. $A\ priori$ this is a difficult task since there is no natural expansion parameter in the theory. As explained earlier, the modes of, say, J with momenta $k \ll m$ can be considered low momentum modes and those with $k \gg m$ can be considered as high momentum modes, with the coupling constant e^2 only serving to separate the modes into these two domains. Our aim will be to focus on the nonperturbative part due to the low k modes. Towards this, we will adopt the following strategy to set up the expansion scheme. We will consider an extension of the theory defined by the Hamiltonian as in (7.26) with m and e considered as independent parameters. This will require a rescaling of the current as explained below. We can then develop a series expansion for the vacuum wave function, writing $\Psi_0 = e^{\frac{1}{2}\mathcal{F}}$ where $\mathcal{F}$ is a power series in e. Mathematically, this framing of the problem, with m and e treated as independent parameters, gives us a way to systematize the solution for the vacuum wave function. At the end, we will set $m = (e^2 c_A/2\pi)$ to regain the gauge theory of interest. (The solution for the vacuum wave functional to the lowest order was given in [16] and used to calculate the string tension. The systematic expansion scheme and the solution with the first set of corrections were given in [17].)

It is worth emphasizing again that this is very different from perturbation theory since m is included exactly in the lowest order result for $\mathcal{F}$. Further in the present case, we are not expanding J in terms of φ either. The resulting recursive procedure will still be some sort of resummed theory. The resummation involves collecting $A, \bar{A}$ in an appropriate series to define J and then including m at the lowest order which is another series. Getting to details, we first do a scale transformation on J as $J \to (ec_A/2\pi)J$. In

terms of the new J, we can write

$$\begin{aligned}
\mathcal{H} &= \mathcal{H}_0 + \mathcal{H}_1 \\
\mathcal{H}_0 &= m\int J_a(\vec{z})\frac{\delta}{\delta J_a(\vec{z})} + \frac{2}{\pi}\int_{w,z}\frac{1}{(z-w)^2}\frac{\delta}{\delta J_a(\vec{w})}\frac{\delta}{\delta J_a(\vec{z})} \\
&\quad +\frac{1}{2}\int_z :\bar{\partial}J^a(z)\bar{\partial}J^a(z): \\
\mathcal{H}_1 &= +ie\int_{w,z} f_{abc}\frac{J^c(w)}{\pi(z-w)}\frac{\delta}{\delta J_a(\vec{w})}\frac{\delta}{\delta J_b(\vec{z})}
\end{aligned} \tag{9.1}$$

As stated before, in the expression for $\mathcal{H}$, we take m and e to be independent parameters for now. The interaction term $\mathcal{H}_1$ is to be treated as a perturbation. In the vacuum wave function $\Psi_0 = e^{\frac{1}{2}\mathcal{F}}$, $\mathcal{F}$ is an arbitrary functional of J. Therefore it can, in general, be taken to be of the form

$$\begin{aligned}
\mathcal{F} &= \int f^{(2)}_{a_1a_2}(x_1,x_2)\ J^{a_1}(x_1)J^{a_2}(x_2) \ +\ \frac{e}{2}\int f^{(3)}_{a_1a_2a_3}(x_1,x_2,x_3)\ J^{a_1}(x_1)J^{a_2}(x_2)J^{a_3}(x_3) \\
&\quad +\ \frac{e^2}{4}\int f^{(4)}_{a_1a_2a_3a_4}(x_1,x_2,x_3,x_4)\ J^{a_1}(x_1)J^{a_2}(x_2)J^{a_3}(x_3)J^{a_4}(x_4)\ +\ \cdots
\end{aligned} \tag{9.2}$$

In accordance with the idea of treating $\mathcal{H}_1$ perturbatively, each of the coefficient functions will also be taken to have an expansion in powers of e^2, so that we can write

$$\begin{aligned}
f^{(2)}_{a_1a_2}(x_1,x_2) &= f^{(2)}_{0\ a_1a_2}(x_1,x_2) + e^2 f^{(2)}_{2\ a_1a_2}(x_1,x_2) + \cdots \\
f^{(3)}_{a_1a_2a_3}(x_1,x_2,x_3) &= f^{(3)}_{0\ a_1a_2a_3}(x_1,x_2,x_3) + e^2 f^{(3)}_{2\ a_1a_2a_3}(x_1,x_2,x_3) + \cdots \\
f^{(4)}_{a_1a_2a_3a_4}(x_1,x_2,x_3,x_4) &= f^{(4)}_{0\ a_1a_2a_3a_4}(x_1,x_2,x_3,x_4) + \cdots
\end{aligned} \tag{9.3}$$

The Schrödinger equation for the vacuum wave function takes the expected form

$$(\mathcal{H}_0 + \mathcal{H}_1)\,\Psi_0 = 0 \tag{9.4}$$

We can now substitute for Ψ_0 with $\mathcal{F}$ as in (9.2) into the Schrödinger equation (9.4) and equate the coefficients of terms with similar powers of J to obtain a set of recursion relations. The term with zero powers of J is a constant which can be removed by a suitable normal-ordering of the Hamiltonian. In fact, we have already taken account of this as indicated by the normal-ordering of the potential energy term. Terms with only one power of J will vanish by color contractions. The lowest nontrivial relation pertains to $f^{(2)}_{a_1a_2}(x_1,x_2)$; it is given by

$$2m\ f^{(2)}_{a_1a_2}(x_1,x_2) + 4\int_{x,y} f^{(2)}_{a_1a}(x_1,x)(\bar{\Omega}^0)_{ab}(x,y)f^{(2)}_{ba_2}(y,x_2) + V_{ab}$$

$$+e^2\left[6\int_{x,y} f^{(4)}_{a_1a_2ab}(x_1,x_2,x,y)(\bar{\Omega}^0)_{ab}(x,y)+3\int_{x,y} f^{(3)}_{a_1ab}(x_1,x,y)(\bar{\Omega}^1)_{aba_2}(x,y,x_2)\right]=0 \tag{9.5}$$

where, for brevity, we have used the definitions

$$\begin{aligned}(\bar{\Omega}^0)_{ab}(x,y) &= \delta_{ab}\partial_y\bar{G}(x,y)=\delta_{ab}\,\frac{1}{\pi(x-y)^2}\\ (\bar{\Omega}^1)_{abc}(x,y,z) &= -\frac{i}{2}\,f^{abc}\left[\delta(z-y)+\delta(z-x)\right]\bar{G}(x,y)\\ V_{ab}(x,y) &= \delta_{ab}\int_z \bar{\partial}_z\delta(z-x)\;\bar{\partial}_z\delta(z-y)\end{aligned} \tag{9.6}$$

We have also used (5.16) for $\bar{G}(x,y)$. For the higher point functions, the recursion relation is given by

$$\begin{aligned}&mpf^{(p)}_{a_1\cdots a_p}+\sum_{n=2}^{p} n(p+2-n)f^{(n)}_{a_1\cdots a_{n-1}a}(\bar{\Omega}^0)_{ab}f^{(p-n+2)}_{ba_n\cdots a_p}\\ &+\sum_{n=2}^{p-1} n(p+1-n)f^{(n)}_{a_1\cdots a_{n-1}a}(\bar{\Omega}^1)_{aba_p}f^{(p-n+1)}_{ba_n\cdots a_{p-1}}\\ &+e^2\left[\frac{(p+1)(p+2)}{2}\;f^{(p+2)}_{a_1\cdots a_pab}(\bar{\Omega}^0)_{ab}+\frac{p(p+1)}{2}\;f^{(p+1)}_{a_1\cdots a_{p-1}ab}(\bar{\Omega}^1)_{aba_p}\right]=0\end{aligned} \tag{9.7}$$

This applies for $p\geq 3$. We must solve (9.5) and this set of equations (9.7) to calculate the vacuum wave functional in our scheme.

9.1 The lowest order solution

With each f having a series expansion in powers of e^2, the lowest order solution to (9.5) is

$$\begin{aligned}f^{(2)}_{a_1a_2}(x_1,x_2)\approx f^{(2)}_{0\;a_1a_2}(x_1,x_2) &= \delta_{a_1a_2}\left[-\frac{\bar{q}^2}{m+E_q}\right]_{x_1,x_2}\\ &= -\delta_{a_1a_2}\int\frac{d^2q}{(2\pi)^2}\left[\frac{\bar{q}^2}{m+E_q}\right]e^{i\vec{q}\cdot(\vec{x}_1-\vec{x}_2)}\end{aligned} \tag{9.8}$$

where $E_q=\sqrt{m^2+q^2}$ and $\bar{q}=\frac{1}{2}(q_1-iq_2)$. Using this expression, we get the vacuum wave function to this order as [16]

$$\boxed{\Psi_0\approx\mathcal{N}\,\exp\left[-\frac{1}{2}\int_{x,y}\bar{\partial}J^a(x)\left(\frac{1}{m+\sqrt{m^2-\nabla^2}}\right)_{x,y}\bar{\partial}J^a(y)\right]} \tag{9.9}$$

where $\mathcal{N}$ is a normalization factor. Even with this lowest order result, we can extract some predictions regarding physical quantities. This will be taken up in the next section, but for now, we will give the first set of corrections to this expression.

9.2 The first order corrections to the vacuum wave functional

For the first order corrections to Ψ_0, we will need the lowest order results for $f^{(3)}$ and $f^{(4)}$. Then using them, we can get $f^{(2)}_{1\ a_1a_2}(x_1,x_2)$, which is the term in $f^{(2)}_{a_1a_2}(x_1,x_2)$ at order e^2.

The expressions for the kernels $f^{(3)}$ and $f^{(4)}$ obtained by solving the recursion rules (9.7) to the lowest order are

$$f^{(3)}_{0\ a_1a_2a_3}(k_1,k_2,k_3) = -\frac{f^{a_1a_2a_3}}{24}\,(2\pi)^2\delta(k_1+k_2+k_3)\;g^{(3)}(k_1,k_2,k_3) \tag{9.10}$$

$$f^{(4)}_{0\ a_1a_2;b_1b_2}(k_1,k_2;q_1,q_2) = \frac{f^{a_1a_2c}f^{b_1b_2c}}{64}\,(2\pi)^2\delta(k_1+k_2+q_1+q_2)\;g^{(4)}(k_1,k_2;q_1,q_2) \tag{9.11}$$

where

$$g^{(3)}(k_1,k_2,k_3) = \frac{16}{E_{k_1}+E_{k_2}+E_{k_3}}\left\{\frac{\bar{k}_1\bar{k}_2(\bar{k}_1-\bar{k}_2)}{(m+E_{k_1})(m+E_{k_2})}+\text{cycl. perm.}\right\} \tag{9.12}$$

$$\begin{aligned} g^{(4)}(k_1,k_2;q_1,q_2) =& \frac{1}{E_{k_1}+E_{k_2}+E_{q_1}+E_{q_2}}\times \\ & \Bigg\{g^{(3)}(k_1,k_2,-k_1-k_2)\;\frac{k_1+k_2}{\bar{k}_1+\bar{k}_2}\;g^{(3)}(q_1,q_2,-q_1-q_2) \\ & -\left[\frac{(2\bar{k}_1+\bar{k}_2)\,\bar{k}_1}{m+E_{k_1}}-\frac{(2\bar{k}_2+\bar{k}_1)\,\bar{k}_2}{m+E_{k_2}}\right]\frac{4}{\bar{k}_1+\bar{k}_2}\;g^{(3)}(q_1,q_2,-q_1-q_2) \\ & -\;g^{(3)}(k_1,k_2,-k_1-k_2)\;\frac{4}{\bar{q}_1+\bar{q}_2}\left[\frac{(2\bar{q}_1+\bar{q}_2)\,\bar{q}_1}{m+E_{q_1}}-\frac{(2\bar{q}_2+\bar{q}_1)\,\bar{q}_2}{m+E_{q_2}}\right]\Bigg\} \end{aligned} \tag{9.13}$$

We have displayed the kernels in terms of their Fourier transforms

$$\begin{aligned} f^{(3)}_{a_1a_2a_3}(x_1,x_2,x_3) &= \int d\mu(k)_3\exp\left(i\sum_i^3 k_ix_i\right)\;f^{(3)}_{a_1a_2a_3}(k_1,k_2,k_3) \\ f^{(4)}_{a_1a_2a_3a_4}(x_1,x_2,x_3,x_4) &= \int d\mu(k)_4\exp\left(i\sum_i^4 k_ix_i\right)\;f^{(4)}_{a_1a_2a_3a_4}(k_1,k_2,k_3,k_4) \end{aligned} \tag{9.14}$$

$$d\mu(k)_n \;=\; \frac{d^2k_1}{(2\pi)^2}\cdots\frac{d^2k_n}{(2\pi)^2} \tag{9.15}$$

Note also that $f^{(4)}_{a_1a_2;b_1b_2}(k_1,k_2;q_1,q_2)$ as defined in (9.11,9.13) is symmetric under independent exchange of the first and second pairs of indices as well as under the simultaneous exchange $(\{a_1,k_1\},\{a_2,k_2\}) \leftrightarrow (\{b_1,q_1\},\{b_2,q_2\})$. It could have been made completely symmetric but it is notationally simpler to leave it as it is for now.

Finally, using the expressions (9.10)-(9.13) for $f^{(3)}_0$, $f^{(4)}_0$ in the recursion rule (9.5), the term of order e^2 in $f^{(2)}$ is given by [17]

$$f^{(2)}_2(q) = \frac{m}{E_q}\left(\int \frac{d^2k}{32\pi}\,\frac{1}{\bar{k}}\,g^{(3)}(q,k,-k-q) \;+\; \int \frac{d^2k}{64\pi}\,\frac{k}{\bar{k}}\,g^{(4)}(q,k;-q,-k)\right) \tag{9.16}$$

This completes the calculation of $\mathcal{F}$ to order e^2. The kernels $f^{(n)}$, $n \geq 5$, are zero to this order, becoming nonzero starting only at the next order in e^2.

To summarize, to the lowest order in our expansion scheme, the vacuum wave function is given in (9.9). Equations(9.10-9.13) and (9.16) then give the first set of corrections to the wave function, i.e., to order e^2.

9.3 Another route to the vacuum wave functional

We have already seen in section 8 how we can define an improved perturbation theory where the lowest order result gives the Hamiltonian for a free massive scalar field. The Hamiltonian given in (8.3) has the term $m\varphi^a(\delta/\delta\varphi^a)$, which assigns a mass m to each power of φ. The existence of this term is directly related to the integration measure

$$d\mu(\mathcal{C}) = d\mu(H)\,e^{2c_A S_{\rm wzw}(H)} \simeq [d\varphi]\,e^{-\frac{c_A}{2\pi}\int \partial\varphi^a\bar{\partial}\varphi^a}\;\left(1 \;+\mathcal{O}(\varphi^3)\right) \tag{9.17}$$

Given this integration measure, the term $m\varphi^a(\delta/\delta\varphi^a)$ is necessary for self-adjointness of the Hamiltonian. We can now use this to give an alternative approach to the wave functional (9.9).

Since it corresponds to a free massive scalar field, the Hamiltonian $\mathcal{H}'$ from (8.7) has the vacuum wave functional

$$\Phi_0 \sim \exp\left(-\frac{1}{2}\int \phi^a\sqrt{m^2-\nabla^2}\,\phi^a\right), \qquad \phi_a(\vec{k}) = \sqrt{c_A k\bar{k}/(2\pi m)}\;\varphi_a(\vec{k}) \tag{9.18}$$

Converting this back to $\Psi = e^{\frac{c_A}{4\pi}\int \partial\varphi\bar{\partial}\varphi}\Phi$, we find

$$\Psi_0 \sim \exp\left[-\frac{c_A}{\pi m}\int (\partial\bar{\partial}\varphi^a)\left(\frac{1}{m+\sqrt{m^2-\nabla^2}}\right)(\partial\bar{\partial}\varphi^a)\right] \tag{9.19}$$

The key argument is then the following. We know, from other considerations, that the wave functions can be taken to be functionals of the current J. For small φ, $H \approx 1 + \varphi$ and $J \approx \frac{c_A}{\pi}\partial\varphi$. So we ask: Is there a functional of J which reduces to the form (9.19) for $J \approx \frac{c_A}{\pi}\partial\varphi$? There is a unique answer, it is given by (9.9),

$$\Psi_0 \approx \mathcal{N} \exp\left[-\frac{1}{2}\int_{x,y} \bar{\partial}J^a(x)\left(\frac{1}{m+\sqrt{m^2-\nabla^2}}\right)_{x,y}\bar{\partial}J^a(y)\right] \tag{9.20}$$

The measure of integration for the inner product is exact, being determined by an anomaly calculation. This in turn fixes the form of $\mathcal{H}$ for the small φ version of the Hamiltonian, and gives the wave function (9.19). The requirement that Ψ be a function of the current then ties it down to the form (9.20). This argument shows that there is a certain robustness to the form of Ψ_0 in (9.20).

10 *Analytic results, comparison with numerics

In the previous section we have obtained the solution of the Schrödinger equation for the vacuum wave function up to the lowest two orders in our expansion scheme. We have also identified an approximate description of the gluons by a massive scalar field. The generation of mass is a nonperturbative effect, even though we expect this approximate description to be valid in a kinematic regime towards the high momentum range. In this section, we will use these results to calculate certain quantities of physical interest and compare with numerical studies.

10.1 String tension

Since confinement has been the aspirational goal of many attempts at the nonperturbative analysis of gauge theories, first, we will consider the calculation of the string tension σ_R for the representation R. As explained in section 3, this is related to the vacuum expectation value of the Wilson loop operator as

$$\langle W_R(C)\rangle \approx \mathcal{N} e^{-\sigma_R A_C} \tag{10.1}$$

Since we are interested in loops of large area, we will consider the vacuum wave function for the low momentum modes of the fields. From (9.9), the lowest order result for this is

$$\Psi_0 \approx \mathcal{N} \exp\left[-\frac{1}{2}\int_{x,y} \bar{\partial}J^a(x)\left(\frac{1}{m+\sqrt{m^2-\nabla^2}}\right)_{x,y}\bar{\partial}J^a(y)\right]$$

$$\approx \quad \mathcal{N} \exp\left[-\frac{1}{4m}\int_x \bar{\partial}J^a\,\bar{\partial}J^a\right] = \mathcal{N}\exp\left[-\frac{1}{8me^2}\int_x F^2\right] \tag{10.2}$$

where, in going to the second line, we have simplified the kernel as it applies to the low momentum modes. The expectation value of the Wilson loop operator $W_R(C)$, where C is purely spatial, can be written as

$$\begin{aligned}\langle W_R(C)\rangle &= \int d\mu(\mathcal{C})\,\Psi_0^*\Psi_0\,W_R(C) \\ &= \mathcal{N}'\int d\mu(\mathcal{C})\,W_R(C)\,\exp\left[-\frac{1}{4g^2}\int F^2\right]\end{aligned} \tag{10.3}$$

where $g^2 = me^2$. This is exactly the Euclidean path-integral version of the expectation value in a two-dimensional Yang-Mills theory, with a coupling constant g^2. By the arguments presented in section 3, we can also calculate this as the interaction energy for a heavy particle-antiparticle pair in the 1+1 dimensional Yang-Mills theory. Using ϕ and χ to represent the heavy particles as in section 3, the action we need is

$$S = \int d^2x\left[-\frac{1}{4g^2}F^2 + i\phi^\dagger D_0\phi + i\chi^\dagger D_0\chi\right] \tag{10.4}$$

This is a fairly trivial theory to investigate. Since the canonical momentum Π_0^a is zero, we can choose $A_0^a = 0$ as the conjugate constraint and eliminate the pair. There is no magnetic field in 1+1 dimensions, so the Hamiltonian in the $A_0^a = 0$ gauge is

$$\mathcal{H} = \frac{1}{2g^2}\int dx\,E^aE^a \tag{10.5}$$

We also have the Gauss law constraint

$$(DE)^a + g^2(\phi^\dagger t^a\phi - \chi^\dagger\tilde{t}^a\chi) = 0 \tag{10.6}$$

(This is the same as (3.6).) As the conjugate constraint, we can take $\partial_x A = 0$. If we take A to vanish at spatial infinity, the only solution is $A = 0$. The Gauss law (10.6) then constrains E^a in terms of the charge densities. Thus there are no propagating degrees of freedom associated to the Yang-Mills field. The solution of the Gauss law condition is $E^a = \partial f^a$, with

$$f^a(x) = -\frac{g^2}{2}\int_y |x-y|\,(\phi^\dagger t^a\phi - \chi^\dagger\tilde{t}^a\chi)(y) \tag{10.7}$$

The Hamiltonian now becomes

$$\mathcal{H} = -\frac{g^2}{4}\int_{x,y}(\phi^\dagger t^a\phi - \chi^\dagger\tilde{t}^a\chi)(x)\,|x-y|\,(\phi^\dagger t^a\phi - \chi^\dagger\tilde{t}^a\chi)(y) \tag{10.8}$$

Acting on the state $|0, L\rangle = \phi_i^\dagger(0)\chi_i^\dagger(L)\,|0\rangle$ we find

$$\begin{aligned}\mathcal{H}\,|0,L\rangle &= \frac{g^2}{2}L\,(\phi^\dagger t^a)_i(\chi^\dagger \tilde{t}^a)_i\,|0\rangle = \frac{g^2 c_R}{2}L\,|0,L\rangle \\ &= \frac{e^4 c_A c_R}{4\pi}\,L\;|0,L\rangle \end{aligned} \tag{10.9}$$

where c_R is the value of the quadratic Casimir operator for the representation R. The string tension can now be read off from this result as

$$\boxed{\sigma_R = \frac{e^4 c_A c_R}{4\pi}} \tag{10.10}$$

This is an analytic prediction for the string tension for the Wilson loop operators in any representation [16].

An interesting observation regarding this result is that we have not used any simplification of the gauge theory that might arise from the large N approximation. However, the final result (10.10) is consistent with large N expectations. For example, for the fundamental representation of $SU(N)$, we find

$$\sigma_F = \frac{(e^2N)^2}{4\pi}\frac{N^2-1}{2N^2} \to \frac{\lambda^2}{8\pi}, \quad \text{as } N\to\infty \tag{10.11}$$

where $\lambda = e^2 N$ is the 't Hooft coupling constant.

10.2 Comparison of string tension with numerical estimates

Even though we obtained the result (10.10) for the string tension using the wave function to the lowest order in our expansion scheme, it is useful at this stage to pause and compare the values given by (10.10) with numerical simulations. In the Table 1, we show the results for a number of different gauge groups and representations carried out by Teper and collaborators. It is clear that the values are very close to the predictions from (10.10), the difference being less than 3%. In addition to these, there has been a high precision calculation for the fundamental representation ($k = 1$) of $SU(2)$ which gives value of $0.33576(24)$ for $\sqrt{\sigma}/e^2$ [21]. Again this compares favorably with our value of 0.3455. An independent numerical estimate of the large N result has also been carried out in [22], giving a value of $0.1964\,N$.

An especially fascinating group is G_2, since all representations of this group are screenable. Lattice-based calculations of the string tension for the representations $\underline{7}$, $\underline{14}$, $\underline{27}$, $\underline{64}$, $\underline{77}$, $\underline{77'}$, $\underline{182}$ and $\underline{189}$ have been carried out in [23]. They have verified the

Group	Representations					
	k=1 Fund.	k=2 antisym	k=3 antisym	k=2 sym	k=3 sym	k=3 mixed
$SU(2)$	0.345 0.335					
$SU(3)$	0.564 0.553					
$SU(4)$	0.772 0.759	0.891 0.883		1.196 1.110		
$SU(5)$	0.977 0.966					
$SU(6)$	1.180 1.167	1.493 1.484	1.583 1.569	1.784 1.727	2.318 2.251	1.985 1.921
$SU(N)$ $N\to\infty$	0.1995 N 0.1976 N					

Table 1: Comparison of $\sqrt{\sigma}/e^2$ as predicted by (10.10) (upper entry) and lattice estimates (lower entry, in red) from [18, 19, 20]. k is the rank of the representation.

relation $\sigma_R = \sigma_7(c_R/c_7)$ (which follows from (10.10)) to within 1%. The value of $\sqrt{\sigma_7}$ itself agrees with (10.10) to within 1.8%.

The fact that the predictions from (10.10) and the results of the numerical calculations do match rather well is very nice, but one could ask whether there are corrections and, if so, whether they do remain small so as not to vitiate the agreement we find here. We will consider the corrections due to the terms of the next order (i.e., to order e^2) in the wave function and show that the corrections are indeed small. Since these calculations are rather long, and all too technical, we will defer this to Appendix C. For now, we will make some comments regarding the string tension and then move on to the propagator masses and the Casimir effect.

10.3 Comments regarding string tension

There are a couple of interesting and important comments to be made about the string tension.

As mentioned earlier, G_2 is a group for which all representations are screenable. The

fundamental representation of G_2 is 7-dimensional while the adjoint representation is 14-dimensional. The product ($7 \times \text{Adjoint} \times \text{Adjoint} \times \text{Adjoint}$) contains a singlet or the trivial representation, ensuring that all representations are screenable. Generally the form of the potential for static sources in screenable representations will show a linear increase with distance up to a certain critical value R_b and will become flat for $r > R_b$. The distance R_b is referred to as the string-breaking distance. The lattice estimate of the string tension for G_2 (and for screenable representations for other groups) is the slope of the linearly rising part, before the flattening, i.e., for $r < R_b$. These are the values for which we make the comparison for the screenable cases.

In a larger context, we can ask whether it makes sense to consider the string picture of confinement in a situation where the string can eventually break. The lattice simulation in [24] considered 3d Yang-Mills theory coupled to a number of scalar fields in the fundamental representation, so that all representations are screenable (by suitable binding with the scalar fields). The results show that an effective string description is still valid for the confining part of the potential; even boundary terms and higher order corrections from the Nambu-Goto string action can be correctly reproduced by the simulation.

The second comment is about Casimir scaling versus the sine-law for the string tension, an issue which took some time to be clarified. Here one considers the k-string corresponding to the antisymmetric rank k representation of $SU(N)$. The value of the quadratic Casimir operator for this representation is easily calculated as

$$c_k = \frac{N+1}{2N}\, k(N-k) \tag{10.12}$$

If the string tensions are proportional to c_k (as we found in (10.10)), and as argued by others as well, then

$$\frac{\sigma_k}{\sigma_F} = \frac{k(N-k)}{N-1} \tag{10.13}$$

This is the Casimir scaling law. An important feature is that in a large N expansion, we have

$$\frac{\sigma_k}{\sigma_F} \approx k\left(1 - \frac{k-1}{N} - \frac{k+1}{N^2} + \cdots\right) \tag{10.14}$$

Thus one can get odd powers of $1/N$ in this case.

The sine-law for the k-string is the statement that

$$\frac{\sigma_k}{\sigma_F} = \frac{\sin(\pi k/N)}{\sin(\pi/N)} \approx k\left(1 - \frac{(k^2-1)\pi^2}{6N^2} + \cdots\right) \tag{10.15}$$

In this case, we have only even powers of $1/N$, evident from the symmetry of the ratio of the sines under $N \to -N$.

The sine-law was recognized as a possibility that one needs to consider following the work of Douglas and Shenker who derived it in $\mathcal{N} = 1$ supersymmetric Yang-Mills (SYM) theory in 4 dimensions [25]. This theory can be obtained by adding a supersymmetry breaking term to the $\mathcal{N} = 2$ SYM theory whose nonpertrubative analysis was carried out by Seiberg and Witten, and who obtained the exact low energy effective action [26]. A similar result was obtained in [27] using a 5-brane construction in M-theory, the so-called MQCD. Within the context of holography, one can obtain the k-string tension as the value of the Hamiltonian for a classical supergravity configuration in the holographic dual description. The sine-law is then obtained for the 4d SYM for the Maldacena-Nunez dual and an approximate sine-law for the Klebanov-Strassler background [28].

While these results were obtained for the supersymmetric theory using the gravity dual, a general argument for the sine-law was suggested in [29], see [30] for a review. The basic argument is the following. Since representations with zero N-ality can all be screened, the asymptotic formula for the string tension should depend only on the N-ality of the representation. The rank k antisymmetric representation and the rank $(N-k)$ antisymmetric representation are conjugates of each other. Therefore we should further expect the tensions to be invariant under $k \to N - k$. This means that the ratio σ_k/σ_F, which is a dimensionless function depending only on k and N, should be a function of $|\sin(\pi k/N)|$; we can represent it as a power series of the form

$$\frac{\sigma_k}{\sigma_F} = c_1|\sin(\pi k/N)| + c_2|\sin(\pi k/N)|^2 + \cdots \tag{10.16}$$

Further, we know that counting powers of N in terms of diagrams in perturbation theory show that at fixed k, one should only have even powers of $1/N$. The limit $N \to \infty$ with fixed k should also exist, as we expect confinement at large N with a finite and nonzero string tension. (This is after everything is expressed in terms of the 't Hooft coupling e^2N as in (10.11).) These properties require that $c_{2n+1} \sim N^1$ and $c_{2n} \sim N^0$. The terms with odd powers have the property that, in the limit $k \to \infty$, $N \to \infty$ with k/N fixed, σ_k/k is a function of $x = \pi k/N$. The authors of [29] refer to this as the saturation property. Keeping only such terms, one ends up with an odd series in $|\sin(\pi k/N)|$, with $c_{2n+1} \sim N$. By comparison with the gravity dual arguments and fitting to some numerical data, one can then argue that a single power of $|\sin(\pi k/N)|$ suffices. The emerging suggestion from this line of reasoning was that Casimir scaling should be ruled out as not being compatible with the $(1/N)$-expansion of Yang-Mills theory.

However, the data from lattice simulations were fairly decisively in favor of Casimir scaling. This follows from the results of [18]-[22] and also from the specific check of Casimir scaling done for G_2 in [23]. Simulations done for the high temperature (T) limit of 4d Yang-Mills theory, which should reduce to the zero-temperature 3d theory with a redefined coupling $e^2 = g^2\mathsf{T}$, also shows Casimir scaling [31]. A calculation using the gravity dual for the 3d SYM also supports Casimir scaling [32]. Detailed analyses with the gravity dual, for the string tension and for the Luscher term, were carried out in [33], [34]; the results seem to lie in between the Casimir law and the sine-law, and close to both cases.

It would seem from the previous two paragraphs that there is a possible conflict between the standard $(1/N)$-expansion and Casimir scaling (which seems to hold for a number of cases and which can include odd powers of $(1/N)$). However, this is not the case, there is a loophole in the arguments presented in [29], as shown by [35]. The essence of this resolution is that, for the string tension, one is calculating a matrix element of the form $\langle 0|F\ e^{-HT}F^\dagger|0\rangle$, as shown in section 3. Using a complete set of energy eigenstates, we can write this in the form

$$\langle 0|F\ e^{-HT}F^\dagger|0\rangle = \sum_\alpha \mathcal{C}_\alpha\, e^{-E_\alpha T} \tag{10.17}$$

where $\mathcal{C}_\alpha$ and E_α are functions of the coupling constant, N, etc. Generally there is also a representation dependence arising from the choice of F. Consider now the $(1/N)$-expansion of various terms in the sum. It is possible for the individual $\mathcal{C}_\alpha$ and E_α to have odd powers of $(1/N)$. When we expand in $(1/N)$ at finite T, there can be cancellation of the odd powers between different terms in the sum, thus rendering the $(1/N)$-expansion of the correlator consistent with expectations from the diagrammatic side. However, if we take large T first, then the term which dominates is the term with the lowest energy, say, e^{-E_0T}. This is what is done both in our analytic calculation and in the lattice simulations, with the string tension extracted from E_0. As mentioned earlier, E_0 can have odd powers of $(1/N)$, but in taking the large T limit first, the possible cancellants of the odd powers of $(1/N)$ from higher $\mathcal{C}_\alpha$, E_α are discarded, so the odd powers in E_0 are retained. This argument shows that there may be no contradiction between Casimir scaling and the $(1/N)$-expansion.

The point is that the two limits, namely large T and large N, do not necessarily commute. (In [35] the authors give a specific example of how such a scenario can be realized, with the cancellation of the odd powers in the correlator, while retaining odd powers in E_0, in a lattice model in the strong coupling expansion.) The conclusion is that Casimir scaling is compatible with the expectations from the $(1/N)$-expansion in

terms of diagrams.

10.4 Casimir effect: Calculation

In section 8 we argued that our analysis leads to an "improved" perturbation theory where, at the lowest order the gluon is described by a scalar field ϕ^a with mass $m = (e^2 c_A/2\pi)$. The Hamiltonian for this was given in (8.7) and it corresponds to the action

$$S = \int d^3x \, \frac{1}{2} \left[\dot{\phi}^a \dot{\phi}^a - (\nabla\phi^a)(\nabla\phi^a) - m^2 \phi^a \phi^a \right] + \cdots \tag{10.18}$$

We can now use this to calculate the Casimir energy for the nonabelian gauge theory, in the usual classic set-up of two parallel conducting plates or, rather, wires since we are in two spatial dimensions. We consider the fields in a square box of side L, with two parallel wires separated by a distance R. Eventually, we can take $L,\, b_1,\, b_2 \to \infty$ keeping R fixed. The relevant geometry is shown in Fig. 1. In the small φ^a expansion, the gauge potentials have the form

$$A_i^a \approx \frac{1}{2} \left[-\partial_i \theta^a + \epsilon_{ij} \, \partial_j \varphi^a + \cdots \right], \qquad M = \exp\left(-\frac{i}{2} t_a (\theta^a + i\varphi^a) \right) \tag{10.19}$$

(The field ϕ^a is related to φ^a as $\phi_a(\vec{k}) = \sqrt{c_A k\bar{k}/(2\pi m)}\ \ \varphi_a(\vec{k})$.) The boundary condition appropriate to perfectly conducting wires is that the tangential component of the electric field should vanish; i.e.,

$$\epsilon_{ij} \, n_i F_{0j}^a = 0, \tag{10.20}$$

where n_i is the unit vector normal to the wire. For small φ^a, we see that this is equivalent to the condition

$$n_i \epsilon_{ij} \epsilon_{jk} \partial_k \dot{\varphi}^a = -n_i \partial_i \dot{\varphi}^a = 0 \tag{10.21}$$

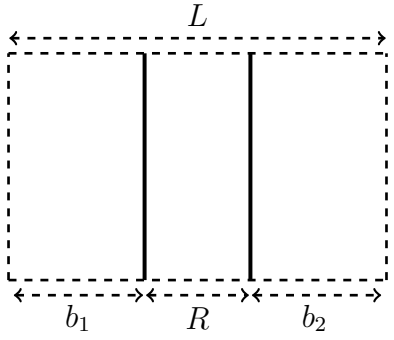


Figure 1: The set-up for Casimir effect

Since the time-derivative does not affect the spatial boundary conditions, this can be satisfied by imposing the Neumann boundary condition $n \cdot \partial\varphi^a = 0$ on the scalar field φ^a or, equivalently, on ϕ^a. This gives us a simple strategy for calculating the Casimir energy within our improved perturbation theory: We just calculate the Casimir energy of a free massive scalar field, of mass m, with Neumann boundary conditions on the wires. (It may be worth re-emphasizing that, even though we use a free field theory, interactions and some nonperturbative effects are folded in since there is a nonzero mass m.) Accordingly, the field in the region between the wires has the expansion

$$\phi^a = \int \frac{dk}{2\pi} \sum_{n=0}^{\infty} C^a_{n,k} \sqrt{\frac{2}{R}} \cos\left(\frac{n\pi x_1}{R}\right) e^{ikx_2}, \tag{10.22}$$

consistent with the Neumann boundary conditions. The action is then obtained as

$$S = \int dt \, \frac{dk}{2\pi} \sum_n \frac{1}{2} \left[\dot{C}^a_{n,k}\dot{C}^a_{n,k} - \Omega^2_{n,k}\, C^a_{n,k} C^a_{n,k}\right] + \cdots \tag{10.23}$$

where $\Omega^2_{n,k} = k^2 + (n\pi/R)^2 + m^2$. Here n is an integer ≥ 1. (Notice that $n = 0$ will not give and R-dependent term.) The unrenormalized zero-point energy can be easily read off as

$$\begin{aligned} \mathcal{E} &= \frac{L}{2}\mathrm{dim}G \int \frac{dk}{2\pi} \sum_n \Omega_{n,k} \\ &= \frac{L}{2}\mathrm{dim}G \int \frac{dk}{2\pi} \sum_n \left(\frac{\partial^2}{\partial x_0^2}\right) \int \frac{dk_0}{\pi} \frac{e^{ik_0 x_0}}{k_0^2 + \Omega^2_{n,k}}\Big]_{x_0=0} \end{aligned} \tag{10.24}$$

The summation can be done using the formula

$$\sum_{n=1}^{\infty} \frac{1}{k_0^2 + \Omega^2_{n,k}} = -\frac{1}{2\omega^2} + \frac{R}{2\omega} + \frac{R}{\omega}\frac{1}{e^{2\omega R} - 1} \tag{10.25}$$

where $\omega^2 = k_0^2 + k^2 + m^2$. Thus $\mathcal{E}$ splits into three terms. The contribution from the first term on the right hand side of (10.25) is independent of R and will disappear when we take $\mathcal{E}(R) - \mathcal{E}(R \to \infty)$ to obtain the renormalized energy. As for the second term, there will be similar contributions from the regions of extent b_1 and b_2, so that together we get $(R + b_1 + b_2)/(2\omega) = L/(2\omega)$. So its contribution is also independent of R. The expression for the energy now becomes

$$\begin{aligned} \mathcal{E} &= -\frac{LR}{4\pi}\mathrm{dim}G \int_0^{\infty} dp \, \frac{p^3}{\sqrt{p^2 + m^2}} \frac{1}{e^{2R\sqrt{p^2+m^2}} - 1} \\ &= -\mathrm{dim}G \, \frac{L}{4\pi R^2}(mR)^3 \int_1^{\infty} dz \frac{(z^2 - 1)}{e^{2mRz} - 1} \end{aligned}$$

$$= \quad -\mathrm{dim}G\,\frac{L}{16\pi R^2}\left[2mR\,\mathrm{Li}_2(e^{-2mR}) + \mathrm{Li}_3(e^{-2mR})\right] \tag{10.26}$$

In going from (10.24), (10.25) to the first line of this equation, we have carried out the angular integrations, and p in this expression is given as $\sqrt{k_0^2 + k^2}$. By using $p = m\sinh q$, and $z = \cosh q$ we get to the second line. The expansion of this in powers of e^{-2mRz} leads to the last line of (10.26) in terms of the polylogarithms,

$$\mathrm{Li}_s(w) = \sum_1^\infty \frac{w^n}{n^s} \tag{10.27}$$

Using the expression (10.10) for the string tension, we can re-express (10.26) in terms of $x = \sqrt{\sigma_F}\,R$ as

$$\boxed{\frac{\mathcal{E}}{L\sigma_F} = -\frac{\mathrm{dim}G}{16\pi}\left[\frac{2\sqrt{c_A/\pi c_F}}{x}\,\mathrm{Li}_2\left(e^{-2\sqrt{c_A/\pi c_F}\,x}\right) + \frac{1}{x^2}\mathrm{Li}_3\left(e^{-2\sqrt{c_A/\pi c_F}\,x}\right)\right]} \tag{10.28}$$

10.5 Casimir effect: Comparison with lattice data

The formula for the Casimir energy given in (10.28) is in a form that can be compared to the lattice simulations. In fact, for the case of $G = SU(2)$, such a simulation and calculation of the Casimir energy for the parallel wire geometry have been carried out in [36]. Using the appropriate values of c_A and c_F, the specialization of formula (10.28) to $SU(2)$ is

$$\frac{\mathcal{E}}{L\sigma_F} = -A\,\frac{\mathrm{dim}G}{16\pi}\left[\frac{1.84}{x}\,\mathrm{Li}_2\left(e^{-1.84\,x}\right) + \frac{1}{x^2}\mathrm{Li}_3\left(e^{-1.84\,x}\right)\right] \tag{10.29}$$

We have also included a prefactor A. The motivation for a possible change of the prefactor (from the value of $A = 1$ as in (10.28)) is the following. The prefactor is really a measure of the number of degrees of freedom. This is clear from the $\mathrm{dim}G$ factor. However, lattice simulations of Yang-Mills theories have shown that the number of degrees of freedom do not quite reach a value corresponding to a gas of free gluons even at very high temperatures, where we expect a deconfined gluon plasma. This has known for a fairly long time. (A recent review which gives updated results is [37]; in particular, see figure 4 of this reference.) There could also be higher order effects (interactions among the φ^a fields, corrections to the wave function, etc.) which could contribute to A. The exponential fall-off is however controlled by the mass m. So we do not tamper with that; the value from our analysis, namely $m = (e^2 c_A/2\pi)$, has been used in (10.29).

As for the comparison with lattice data, the authors of [36] fitted the data points to a phenomenologically motivated formula

$$\frac{\mathcal{E}}{L\sigma_F} = -\dim G\,\frac{\zeta(3)}{16\pi}\,x^{-\nu}\,e^{-M\,x/\sqrt{\sigma_F}}, \tag{10.30}$$

the best fit values being $\nu = 2.05$ and $M = 1.38\,\sqrt{\sigma_F}$. in Fig. 2, we show the curve corresponding to (10.30) as the dashed red line, using the best fit values quoted above. It is very clear that our formula (10.29) is at least as good a fit to the lattice data as the phenomenological formula (10.30). We have used only one fitting parameter, namely A. Its best fit value is $A = 0.9715$. If we used (10.28) without allowing for a change of the prefactor (which means $A = 1$), the agreement would still be rather good, since the deviation is only about 3%. Notice that the exponential factors are just as predicted from (10.28). Even a small error in the mass m could give a significant deviation since it is in the exponent.

Why is our result for the Casimir energy so accurate considering that it is obtained using the "free theory", albeit including the mass which is nonperturbatively generated? Obtaining a lattice estimate of the Casimir energy at large separations is problematic because of the exponential damping. The numerical values are lost in the noise. At the other end, for short distance, lattice artifacts get in the way. So the lattice estimates are by necessity confined to a certain range (roughly between $x = 0.1$ and $x = 0.7$ in the graph). This is the kinematic regime which we might expect to be more or less accessible by perturbation theory, but improved to incorporate a mass which is necessary to include the exponential fall-off.

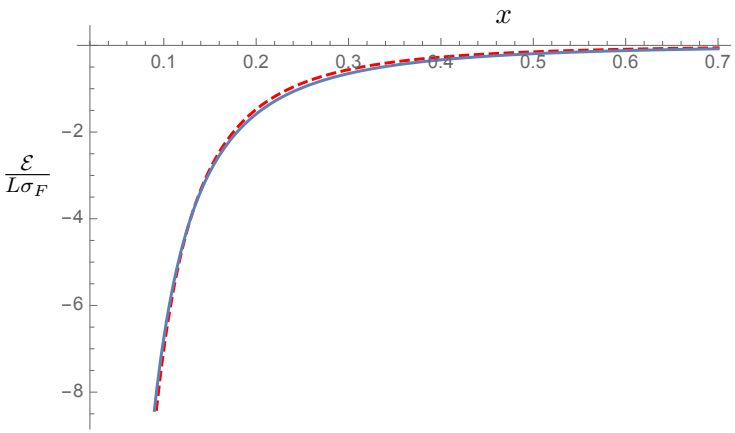


Figure 2: Comparison of (10.30) (dashed red line) and (10.29) (solid blue line).

10.6 Propagator mass: Alternate approaches

The Casimir effect, as discussed above, may be the most accurate way to test the prediction about the propagator mass for gluons. But there are a few other ways to attempt the calculation or the numerical estimate of this quantity.

First of all, since we argued that our analysis, in the high momentum regime, could be viewed as a resummation of a select series of Feynman diagrams, one could attempt a direct resummation within standard covariant perturbation expansion. In such an approach, the difficulty we might face is that the selection of the terms to be resummed has to respect gauge invariance or BRST invariance. This means that the chosen set of terms should form a closed set with respect to the relevant Ward-Takahashi identities. Ensuring this feature can be cumbersome in practice. However, since we are primarily interested in the mass, not full-blown off-shell amplitudes, there is a simpler method we can use. The idea is to first construct a possible gauge-invariant mass term for the gauge fields. This will be of the form

$$S_{\rm mass} = \frac{1}{2}\int \frac{d^3k}{(2\pi)^3} A_i^a(-k) A_j^a(k) \left(\delta_{ij} - \frac{k_i k_j}{\vec{k}^2}\right) + \mathcal{O}(A^3) \tag{10.31}$$

The quadratic term shows that this is truly a mass term for the transverse gauge fields, but is not gauge-invariant. However, one can add to it a suitable series involving A's to get a gauge-invariant completion of this term. Since the completion is not uniquely defined, we can have different possible choices for $S_{\rm mass}$. All such mass terms are necessarily sums of nonlocal monomials of the fields. Once we have chosen a mass term $S_{\rm mass}$, we consider the action[6]

$$S = S_{\rm YM} + M^2 S_{\rm mass} - \Delta\ S_{\rm mass} \tag{10.32}$$

where $S_{\rm YM}$ is the usual Yang-Mills action. We take Δ to have a loop expansion starting at the 1-loop order, writing

$$\Delta = \sum_1 \hbar^n\ \Delta^{(n)} \tag{10.33}$$

After adding gauge fixing and ghost terms, we can calculate the effective action Γ. This will have the form

$$\Gamma = S_{\rm YM} + M^2 S_{\rm mass} - \sum_1 \hbar^n \Delta^{(n)}\ S_{\rm mass} + \hbar\Gamma^{(1)}(\hbar) + \hbar^2\Gamma^{(2)}(\hbar) + \cdots \tag{10.34}$$

[6]For the rest of this section, we use the Euclidean theory in keeping with the fact that this signature is used in diagrammatic perturbation theory.

Notice that while $\Gamma^{(1)}$ is obtained from terms which are graphically of the 1-loop topology, it contains terms of arbitrary order in $\hbar$ since the propagators involve Δ and are of the form $(k^2 + M^2 - \Delta)^{-1}$. Thus $\Gamma^{(1)}$ has the expansion

$$\Gamma^{(1)} = \Gamma_0^{(1)} + \sum_1 \hbar^n \, \Gamma_n^{(1)} \tag{10.35}$$

In a similar way, additional $\hbar$ dependence arises for the 2-loop vertices. The terms of order $\hbar^2$ in $\Gamma^{(2)}$ are due to
a) 2-loop graphs with the propagators $\sim (k^2 + M^2)^{-1}$
b) terms from $\Gamma_1^{(1)}$.
Collectively we will denote these terms by $\Gamma_0^{(2)}$. Just to clarify, we may note that the second set of terms can arise from 1-loop diagrams with vertices from $S_{\rm YM} + M^2 S_{\rm mass}$ (plus similar ghost vertices) and propagators expanded to $\Delta^{(1)}$ order or from 1-loop diagrams with propagators $\sim (k^2 + M^2)^{-1}$, but with vertices from $\Delta^{(1)} S_{\rm mass}$. Similar reasoning will hold for higher order terms. The expansion (10.34) thus takes the form

$$\begin{aligned} \Gamma \;\; = \;\; & S_{\rm YM} + M^2 S_{\rm mass} + \hbar \left(\Gamma_0^{(1)} - \Delta^{(1)} S_{\rm mass} \right) + \hbar^2 \left(\Gamma_0^{(2)} - \Delta^{(2)} S_{\rm mass} \right) + \cdots \\ & + \text{gauge-fixing terms} + \text{ghost terms} \end{aligned} \tag{10.36}$$

We take M^2 to be the propagator mass; this is so at the tree level with the pole appearing at $k^2 + M^2 = 0$. It can be kept at the same point to order $\hbar$ by choosing $\Delta^{(1)}$ to cancel any shift of the pole induced by $\Gamma_0^{(1)}$. This will determine $\Delta^{(1)}$ as a function of M. Likewise, we choose $\Delta^{(2)}$ to cancel any shift of the pole at order $\hbar^2$, etc. Finally, we are not seeking to change the theory, it should still be the usual Yang-Mills theory. So we should at the end set

$$M^2 = \Delta = \Delta^{(1)} + \Delta^{(2)} + \cdots \, , \tag{10.37}$$

so that the starting action (10.32) is just the Yang-Mills action. This equation becomes a gap equation for the theory determining M in terms of the coupling constant.

The procedure we have outlined gives a systematic and gauge-invariant way to carry out a resummation of the select set of terms and identify the propagator mass. It can also be implemented order by order; for example, $M^2 = \Delta^{(1)}$ will be the 1-loop gap equation, $M^2 = \Delta^{(1)} + \Delta^{(2)}$ is the 2-loop gap equation, etc. The series of terms which are selected to be resummed is determined by the choice of $S_{\rm mass}$, with different choices corresponding to different series being resummed. This method of obtaining the gap equation has been explained in some detail in [38, 39].

Calculations along these lines have been carried out for several different choices of $S_{\rm mass}$. In [40], a Higgs field was used to generate a gauge-invariant mass term, in a way similar to how the Higgs mechanism gives a mass to vector bosons. In [38], we have used a different mass term inspired by the 2d-WZW action and also by the Debye screening mass term in 4d-Yang-Mills theory at high temperatures. This mass term also has an interesting geometrical side to it, in terms of Sasakian structures on three-dimensional spaces [41]. Jackiw and Pi have used a more conventional mass term of the form $F(1/D^2)F$ [39]. A Chern-Simons term, although parity-violating, has also been used [42]. In the references cited so far, the calculations were done to 1-loop order and the resulting gap equation was then solved to obtain the value of the propagator mass. The calculations of [40] have also been extended to obtain the 2-loop gap equation in [43, 44]. The values of the propagator mass obtained using these mass terms are given in Table 2.

A second method is to use the Schwinger-Dyson formulation of the theory [45]. This is effectively a reorganization and resummation of the perturbation expansion and so it is ideologically related to the resummation method discussed above. By combining the Schwinger-Dyson equation with the pinch technique [46, 47], it is possible to get gluon propagators and identify the mass. (The pinch technique is a way of adding a certain kinematic limit (the pinching limit) of some Feynman diagrams to other n-point functions to obtain gauge-invariant n-point functions. See the references quoted to see the details of how this can be implemented.) The Schwinger-Dyson equations are nonlinear and in the end, in this approach, some numerical work is needed to solve them. The result seems to give a value close to what is obtained by the other methods, as seen from Table 2.

Yet another approach is to use lattice simulations again. Arguably, the best feature of lattice gauge theory is the ease of preserving manifest gauge invariance. However, to define the gluon propagators and the propagator mass one needs to do gauge-fixing. This will also bring in issues like the Gribov problem. Nevertheless, calculations have been done using the Landau gauge, the maximal Abelian gauge and for what we shall refer to as the $\lambda_3 = 2$ gauge. (See the quoted reference for details.)

Finally, we point out that Philipsen [52] has also calculated the propagator mass by considering the correlation length and fall-off of gauge-invariant partonic correlators in the 4d-Yang-Mills theory at high temperatures (which is equivalent to the 3d-theory with a redefinition of the coupling constant).

Group	m/e^2	Method
$SU(2)$	0.38	Resummation, 1-loop [38]
	0.28	Resummation, 1-loop [40, 39]
	0.35	Lattice, common factor for glueball masses [40]
	0.34	Two-loop gap equation [43]
	0.33	Two-loop gap equation [44]
	0.25	Resummation of perturbation theory [48]
	0.51	Lattice, maximal abelian gauge [49]
	0.52	Lattice, Landau gauge [49]
	0.44	Lattice, $\lambda_3 = 2$ gauge [49]
	0.456	Lattice, Landau gauge [50]
	0.37	Gauge-invariant lattice definition [51]
	0.36	Gauge-invariant correlation length [52]
	0.32	Calculation via our Hamiltonian method
$SU(3)$	0.568	Resummation, 1-loop [38]
	0.42	Resummation, 1-loop [40, 39]
	0.515	Lattice, Landau gauge [53]
	0.482	Lattice, Landau gauge [54]
	0.48	Calculation via our Hamiltonian method

Table 2: Comparison of propagator mass calculations

11 Screenable representations and string breaking

Screenable representations are representations R such that $R \times \text{Adjoint} \times \text{Adjoint} \cdots$ contains the trivial (or singlet) representation. For $G = SU(N)$, these are representations of zero N-ality. As discussed in section 3, if we consider a particle-antiparticle pair in a screenable representation, the potential between them will flatten out at some point as the separation between the two is increased. Thus, for screenable R, the formula for the string tension should apply only up to this critical separation.

The process of the flattening out of the potential can be visualized as follows. Since $R \times \text{Adjoint} \times \text{Adjoint} \cdots$ contains a singlet, we could have a bound state of zero charge made of the particle with a certain number of gluons. This bound state is usually referred to as a glue lump. Similarly one can have a glue lump for the antiparticle with some gluons. As we increase the separation between the particle and the antiparticle, at some point, it becomes energetically favorable to convert the interaction energy

into creating a glue lump pair. Once this is achieved, there is no further energy cost to separating the lumps, since each of them has zero charge. The energy of the pair remains what it was at the point of the glue lump formation. This is the flattening, which can also be thought of as the breaking up of the string connecting the particle and antiparticle.

The ideal scenario theoretically would be to see this directly in the calculation of the expectation value of the Wilson loop. This has not been possible so far using the vacuum wave function we have obtained, namely, the result to the lowest two orders within our expansion scheme. Presumably higher order effects will be important for string breaking. However, what we can do is an approximate calculation of the ground state energy of a glue lump and then argue that the energy needed to create a glue lump pair in their ground state is the critical value V_* of the potential energy at the string-breaking point.

11.1 A Schrödinger equation for the glue lump

We will first outline the calculation of the glue lump ground state energy. Recall that we can consider the Wilson loop as a process involving the propagation of a heavy particle-antiparticle pair, each, say, of mass M. The simplest case to consider is when the representation R for the particle is the adjoint one, so that the glue lump is the bound state of this particle (or antiparticle) with the gluon [55]. If ϕ^a is field representing the heavy particle, then the gauge-invariant version, denoted by Φ^a, is given by

$$\Phi^a t_a = M^\dagger \, \phi^a t_a \, M^{-1} \tag{11.1}$$

The wave function of a glue lump state will then have the form

$$\begin{aligned} \Psi_G &= \int_{x,y} f(\vec{x},\vec{y}) \, \bar{\partial} J^a \widetilde{W}^{ab}(x,y) \Phi^b \, \Psi_0 \\ \widetilde{W}^{ab}(x,y) &= \left[K(x,\bar{y}) K^{-1}(y,\bar{y})\right]^{ab} \end{aligned} \tag{11.2}$$

Here Φ^a represents the particle and J^a the gluon. $f(\vec{x},\vec{y})$ is the two-body wave function for the gluon and the heavy particle. Ψ_0 is the wave function for the ground state, which is given by

$$\Psi_0 = \mathcal{N} \exp\left(-\frac{1}{2}\mathcal{F}\right) \exp\left(-\frac{1}{2} M \int \Phi^a \Phi^a\right) \tag{11.3}$$

with $\mathcal{F}$ as given in section 9. The second exponential is the ground state wave function for a heavy particle of mass M.

Our aim now is to act on this wave function with the Hamiltonian. In some approximation, as explained below, this will lead to an ordinary two-body Schrödinger equation for $f(\vec{x},\vec{y})$. We can then estimate the ground state energy as we do in quantum mechanics. The Hamiltonian is given by $\mathcal{H}$ from section 7, or the expression (9.1), with the Hamiltonian for the scalar field added to it. The result is $\mathcal{H}=\mathcal{H}_{\rm YM}+\mathcal{H}_\Phi$, with

$$\begin{aligned}
\mathcal{H}_{\rm YM} &= m\int J_a(\vec{z})\frac{\delta}{\delta J_a(\vec{z})}+\frac{2}{\pi}\int_{w,z}\frac{1}{(z-w)^2}\frac{\delta}{\delta J_a(\vec{w})}\frac{\delta}{\delta J_a(\vec{z})}\\
&\quad +\frac{1}{2}\int_z :\bar\partial J^a(z)\bar\partial J^a(z): +ie\int_{w,z} f_{abc}\frac{J^c(w)}{\pi(z-w)}\frac{\delta}{\delta J_a(\vec{w})}\frac{\delta}{\delta J_b(\vec{z})}\\
\mathcal{H}_\Phi &= \int :\left[-\frac{1}{2}\frac{\delta^2}{\delta\Phi^a\delta\Phi^a}+\left(\frac{2\pi}{c_A}\bar\partial\Phi^a(\mathcal{D}\Phi)^a+\frac{M^2\Phi^a\Phi^a}{2}\right)\right]\\
&\quad +im\int_{z,w}\Lambda_{cd}(\vec{w},\vec{z})f^{abc}\Phi^a(\vec{w})\frac{\delta}{\delta\Phi^b(\vec{w})}\frac{\delta}{\delta J^d(\vec{z})}\ : \qquad (11.4)
\end{aligned}$$

where

$$\begin{aligned}
\mathcal{D}_{w\ ab} &= \frac{c_A}{\pi}\partial_w\delta_{ab}+if_{abc}J_c(\vec{w})\\
\Lambda_{cd}(w,z) &= -\partial_z\left[\int_x\bar{\mathcal{G}}_{ac}(\vec{x},\vec{w})K_{ab}(\vec{x})\mathcal{G}_{bs}(\vec{x},\vec{z})\right]\ K^{-1}_{sd}(\vec{z})\\
\bar{\mathcal{G}}_{ma}(\vec{x},\vec{y}) &= \frac{1}{\pi(x-y)}\Big[\delta_{ma}-e^{-|\vec{x}-\vec{y}|^2/\epsilon}\big(K(x,\bar y)K^{-1}(y,\bar y)\big)_{ma}\Big]\\
\mathcal{G}_{ma}(\vec{x},\vec{y}) &= \frac{1}{\pi(\bar x-\bar y)}\Big[\delta_{ma}-e^{-|\vec{x}-\vec{y}|^2/\epsilon}\big(K^{-1}(y,\bar x)K(y,\bar y)\big)_{ma}\Big] \qquad (11.5)
\end{aligned}$$

The propagators given in the last two lines of (11.5) are the regularized form of the corresponding propagators. $\mathcal{H}_\Phi$ is normally ordered so that there is no zero-point energy, as indicated in (11.4).

We can now consider the action of $\mathcal{H}$ on Ψ_G. Before we give details, we will make some observations which are useful in understanding the genesis of various terms in the resulting Schrödinger equation. We have already seen that $TJ^a=mJ^a$ in (7.12). When we include Ψ_0 as well, we find

$$T\,J^a(x)\Psi_0=\left(m-\frac{\nabla^2}{2m}\right)J^a(x)\Psi_0+\frac{2i\pi}{mc_A}f_{abc}J^b\bar\partial J^c\Psi_0+\cdots \qquad (11.6)$$

If we ignore the terms of order J^2, this is like an eigenvalue equation, the eigenvalue itself being $(m-\nabla^2/2m)$, which is the nonrelativistic version of $\sqrt{m^2+k^2}$.[7]

[7]Actually one can recover the fully relativistic expression $\sqrt{m^2+k^2}$ by summing up a series in $1/m$. The situation is very similar to what happens with quantum fluctuations around a static soliton. A series of terms produced by the zero modes of the Hessian has to be summed up to get the relativistic formula.

Notice that we also have

$$\int : \left[-\frac{1}{2}\frac{\delta^2}{\delta\Phi^a\delta\Phi^a} + \frac{1}{2}M^2\Phi^a\Phi^a \right] : \Phi^b\Psi_0 = M\,\Phi^b\Psi_0 \tag{11.7}$$

The action of the operator on the left also produces a term $\frac{1}{2}M\,\Phi^b\Psi_0$, but a similar result is obtained for just Ψ_0 as well; it is the vacuum energy which is removed by the normal ordering. From these two statements and the fact that Ψ_G contains $\bar{\partial}J^a$ and Φ^b, we expect that the action of the Hamiltonian will produce a contribution of $(M+m-\nabla^2/2m)$ to the eigenvalue. (One could also have $-\nabla^2/2M$, but this is negligible as we take M very large.) There will also be the energy of the interaction between the charged factors $\bar{\partial}J^a$ and Φ^b in Ψ_G. These expectations are born out by the explicit calculations, with the result

$$\mathcal{H}\,\Psi_G = \int \left[M + m - \frac{\nabla_x^2}{2m} + \sigma_A|\vec{x}-\vec{y}| \right] f(\vec{x},\vec{y})\;\bar{\partial}J^a(\vec{x})\widetilde{W}^{ab}(\vec{x},\vec{y})\Phi^b(\vec{y})\Psi_0 \;+\cdots \tag{11.8}$$

Here σ_A is the string tension for the adjoint representation and the ellipsis stands for a number of terms we have neglected. The approximations involved in arriving at this equation are the following.

1. First of all, there is the obvious approximation of using the leading solution (9.9) for the vacuum wave function; i.e., $\mathcal{F}$ in (11.3) is taken to be the leading kernel for the quadratic term in the J's.

2. There are terms which correspond to new operator structures, i.e., they are not of the form of Ψ_G and have more powers of J. These are possible new states in the theory. The glue lump state $\Psi_G = \bar{\partial}J\widetilde{W}\Phi\,\Psi_0$ given in (11.2) can have overlap with such states since they have the same quantum numbers.

   The fact that the Hamiltonian acting on Ψ_G can produce these other structures shows that there can be mixing. But, as is well known in quantum mechanics, the effect of nondiagonal terms in $\mathcal{H}$ comes with energy denominators and can be taken to be small if the differences between the energies of the lowest glue lump state (which is what we are interested in) and higher states are large enough. We expect this to be the case at large enough coupling, since the differences must go like m. But ultimately it is to be justified *a posteriori*.[8] For more details on these issues, see [55].

[8]The transition amplitude from the Wilson loop without string-breaking to the glue lumps is seen to be very small from lattice data, even for $SU(2)$ [56, 57]. This is another indication for the expectation that the off-diagonal elements mentioned above should be small.

Accepting the caveats mentioned above, we can now see that the glue lump state (11.2) will be an eigenstate of the Hamiltonian if $f(\vec{x},\vec{y})$ obeys the ordinary Schrödinger equation

$$\left[M + m - \frac{\nabla_x^2}{2m} + \sigma_A|\vec{x}-\vec{y}|\right] f(\vec{x},\vec{y}) = E\ f(\vec{x},\vec{y}) \tag{11.9}$$

Removing the center of mass motion (which is zero as $M \to \infty$), we see that this equation is the obtained by minimizing the energy functional

$$\begin{aligned} E &= M + m + \frac{1}{\mathcal{N}}\int d^2x\ \left[\frac{|\nabla f|^2}{2m}\ + \sigma_A|\vec{x}|\ |f|^2\right] \\ \mathcal{N} &= \int d^2x\ |f|^2 \end{aligned} \tag{11.10}$$

The simplest way to proceed further is to use a variational procedure. We consider an ansatz of the form

$$f = \exp(-\beta|x|^\mu) \tag{11.11}$$

where β and μ are to be treated as variational parameters. Calculating $E(\beta,\mu)$ and extremizing it we find that the minimum occurs at $\beta = \beta_*$, with

$$\begin{aligned} E(\beta_*,\mu) &= M + m + \left[\frac{2^{-(2\mu+1)/\mu}}{2m\Gamma(2/\mu)}\left(\frac{2^{(\mu-3)/\mu}2m\sigma_A\Gamma(3/\mu)}{\mu^2}\right)^{-1/3}\right. \\ &\qquad \left.\times\left(8m\sigma_A\Gamma(3/\mu) + 8^{1/\mu}\mu^2\left(\frac{2^{(\mu-3)/\mu}2m\sigma_A\Gamma(3/\mu)}{\mu^2}\right)\right)\right] \\ \beta_* &= \left(\frac{2^{(\mu-3)/\mu}(2m\sigma_A)\Gamma(3/\mu)}{\mu^2}\right)^{\mu/3} \end{aligned} \tag{11.12}$$

The minimization with respect to μ has to be done numerically.

A seemingly more general ansatz for f is

$$f = \frac{e^{-\beta|x|^\mu}}{(1+|x|)^\nu} \tag{11.13}$$

with β, μ, ν as variational parameters. This may seem better motivated since the solutions of the Schrödinger equation with a linear potential involve the Airy functions Ai and we have

$$\frac{\mathrm{Ai}\left((2m\sigma_A)^{1/3}|x|\right)}{\sqrt{|x|}} \approx \frac{\exp\left[-\frac{2}{3}\left((2m\sigma_A)^{1/3}|x|\right)^{3/2}\right]}{|x|^{3/4}} \tag{11.14}$$

for large $|x|$. However the minimization of $E(\beta,\mu,\nu)$ shows that the minimum occurs at $\nu = 0$, so, in the end, this is equivalent to (11.11).

11.2 Comparison with lattice simulations

The calculation of string-breaking given above is admittedly rather crude. As stated, there are several terms which have been neglected. Even after this, the Schrödinger equation (11.9) has only the linear potential $\sigma - A|\vec{x}-\vec{y}|$. At short distances one should expect a Coulomb potential (logarithmic in 2+1 dimensions) as in perturbation theory. In the glue lump, since the wave function has some nonzero probability at short separation between the constituents, the Coulomb potential can have an impact on the energy. There is also the Luscher term $(\pi/24|\vec{x}-\vec{y}|)$ for the potential energy. Finally we are only carrying out a variational estimate of the ground state energy. So, all things considered, the calculation presented above should be viewed as primarily being of qualitative value, demonstrating the possibility of string-breaking. Nevertheless, it is interesting to compare with lattice estimates, keeping in mind all the caveats mentioned above.

Minimizing the energy in the formula (11.12) with respect to μ and using the value of the adjoint string tension given by (10.10), namely, $\sigma_A = (e^4 c_A^2/4\pi) = \pi m^2$, we find $E(\beta_*, \mu_*) = M + 3.958m$. Since we had two masses M initially, the extra energy needed to create a glue lump pair is $2 \times 3.958m = 7.916m$. Thus we should expect that the string breaking should occur when the interaction energy is $V_{*\mathrm{cal}} = 7.916m$.

A point of internal consistency of the calculation is the following. Taking the value $7.916m$ for $V_{*\mathrm{cal}}$, the separation between the constituents of the glue lump at the point of breaking is $r_* = (V_*/\sigma_A) \gtrsim 2.52/m$ and so, the typical momentum at this separation $\lesssim (m/2.52)$. While this does not make an unassailable case for the nonrelativistic treatment, it is not inconsistent.

Turning to the lattice numbers, an early estimate of string-breaking in $SU(2)$ by Philipsen and Wittig [56] gave a value of the breaking separation R_b as $13.6/M_g$ where M_g is the mass of the lightest glueball. Taking this to be the 0^{++} with a mass of $5.17\,m$ as given by [18], this translates to $V_{*\mathrm{lat}} = 8.26m$. The calculated value agrees with the lattice estimate up to $|V_{*\mathrm{lat}} - V_{*\mathrm{cal}}|/V_{*\mathrm{lat}} = 4.16\%$.[9]

In the lattice estimate by Kratochvila and de Forcrand [57], the value of V_* for $G = SU(2)$ was obtained as $V_{*\mathrm{lat}} = 2.063a^{-1}$, while the fundamental string tension is $\sigma_F = 0.0625a^{-2}$, where a is a lattice spacing used in the simulation. This implies

[9] String-breaking was also clearly demonstrated in [58], but the authors considered the theory with quarks in the fundamental representation. This is different from our case of adjoint static charges, so a comparison is not obtained.

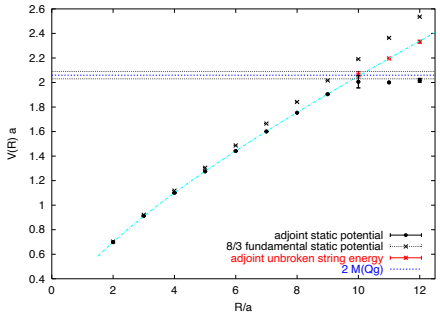


Figure 3: The adjoint and fundamental static potentials $V(R)$ (the latter multiplied by the Casimir factor $8/3$) versus R using Wilson loops only. The adjoint static potential remains approximately constant for $R \geq R_b = 10a$ proving string breaking. The unbroken-string state energy is also drawn. The horizontal line at $2.06(3)a^{-1}$ represents twice the mass of a glue lump. This is graph is from [57].

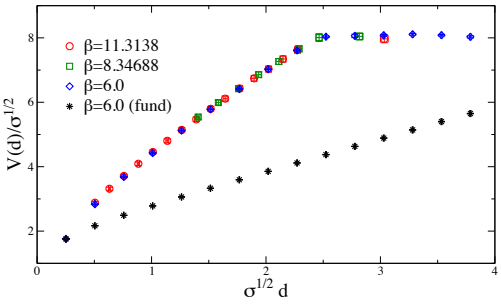


Figure 4: Continuum scaling of the adjoint static potential: data have been obtained on lattices with $L = 32(\beta = 6.0)$, $L = 48(\beta = 8.34688)$ and $L = 64(\beta = 11.3138)$. The potential between two fundamental charges for $\beta = 6.0$ is also reported for comparison. This is graph is from [61].

$V_{*\mathrm{lat}} = 2.06m\sqrt{(\pi c_F/0.0625c_A)} = 8.68m$, fairly close to the value obtained by [56]. The deviation of this from the calculated value of $7.916m$ is approximately 8.76%. However, the form of the potential which emerges from the simulations in [57] do display a Coulomb term and the Lüscher term in addition to the linear potential, so a direct comparison with (11.12) is not really appropriate. A more meaningful quantity might be the distance at which breaking occurs, signaled by the flattening of the potential. This happens at about $10a$ for $\beta = 6$, where β is the parameter which appears in the lattice action (Wilson action) and a is the lattice spacing. These are related to the coupling e^2 as $a = (2N/\beta e^2)$. We then find $\sigma_A \times (10a) = 6.67m$. This is about 19% below our estimate. (The flattening of the potential is very clearly seen in many lattice simulations. We display two examples, just to illustrate this point, in Fig. 3 and Fig. 4.)

Another recent estimate is the one reported in [59]. These authors show that the breaking occurs at a value of 1.3 (which is twice the mass of a "constituent gluon" in their terminology). The units used are such that the 0^{++} glueball has a mass of 1.198. Taking $M_{0^{++}} = 5.17m$ as in [18], we find that $V_{*\mathrm{lat}} = (1.3/1.198) \times 5.17m = 5.61m$.

We also mention [60], where the authors again demonstrate string-breaking for a bound state of a quark and an antiquark. As in the case of [58], the quarks being in the fundamental representation, we are not able to obtain a comparison with the value

calculated here.

An interesting recent simulation for $SU(2)$ lattice gauge theory considered the longitudinal and transverse chromoelectric fields as the distance between the particle and antiparticle static sources is increased [61].[10] The longitudinal fields suddenly drop to almost zero as the separation approaches the string-breaking value. This happens at about $10a$ for $\beta = 6$, in agreement with [57]. So, as in that case, $\sigma_A \times (10a) = 6.67m$. It is however worth mentioning that the focus of [61] was not so much on the value of the energy. The key result is about how the breaking occurs, signaled by the rapid decay of of the longitudinal contribution to the energy (for which they find clear evidence).

12 Alternate candidates for the vacuum wave functional

The procedure we have described in section 9 gives a systematic expansion scheme for solving the Schrödinger equation. The solution was then used in arriving at the formula for the string tension. But there have been a few other suggestions regarding the vacuum wave functional for the 3d Yang-Mills theory. We will briefly review some of them here.

In a couple of very interesting papers, Leigh, Minic and Yelnikov (LMY) considered an alternate method of solving the Schrödinger equation [62]. The starting observation was to note that the kernel given in the solution (9.9) can be expanded in powers of $-\nabla^2$ as

$$\left(\frac{1}{m+\sqrt{m^2-\nabla^2}}\right) = \frac{1}{2m}\left[1 - \frac{1}{4}\left(\frac{-\nabla^2}{m^2}\right) + \frac{1}{8}\left(\frac{-\nabla^2}{m^2}\right)^2 + \cdots\right] \tag{12.1}$$

As a result, the exponent of the wave functional (9.9) can be viewed as a sum of terms involving monomials of the form

$$\mathcal{O}'_n = \int \bar\partial J^a (\partial\bar\partial)^n \bar\partial J^a \tag{12.2}$$

Based on this, LMY introduce a set of operators [62]

$$\mathcal{O}_n = \int \bar\partial J^a \left[(\mathcal{D}\bar\partial)^n\right]^{ab} \bar\partial J^b, \quad \mathcal{D}^{ab} = \frac{c_A}{\pi}\partial\,\delta^{ab} + i f^{abc} J^c \tag{12.3}$$

(The operator $\mathcal{D}^{ab}$ was previously introduced in (11.5).) The motivation for this has to do with invariance under holomorphic transformations of H, namely, under $H \to V(z) H \bar V(\bar z)$. The operator $\mathcal{D}$ is a covariant derivative for this and leads to manifest

[10] I thank Claudio Bonati for a useful comment and for sharing some of their data.

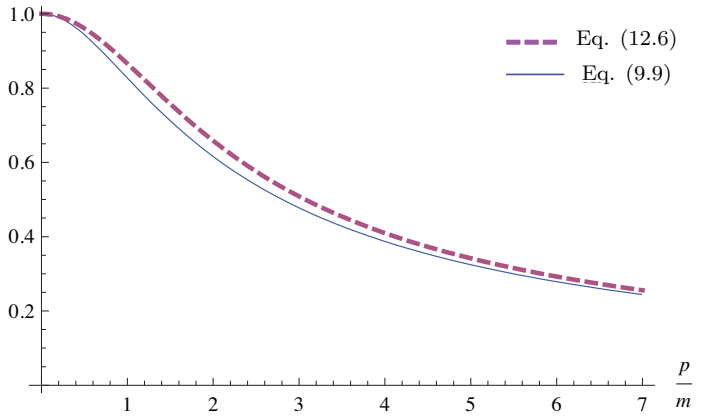


Figure 5: Comparison of the kernel from (9.9) and (12.6) for the Gaussian term of the vacuum wave functional as a function of the momentum p

invariance for $\mathcal{O}_n$ under the holomorphic transformations. The next step in [62] was to postulate an ansatz for the wave functional of the form $\Psi_0 = e^{\frac{1}{2}\mathcal{F}}$, with

$$\mathcal{F} = \sum_n c_n \mathcal{O}_n \tag{12.4}$$

To evaluate the action of the Hamiltonian, they assumed that $\mathcal{O}_n$ are eigenvectors of the kinetic operator,

$$T\,\mathcal{O}_n = (2+n)m\,\mathcal{O}_n \tag{12.5}$$

This relation can actually be proved for $n = 0, 1$, but there are additional terms in general for higher n. Nevertheless, if one neglects any correction to (12.5), one can solve the Schrödinger equation and arrive at a wave functional

$$\Psi_0 = \exp\left[-\frac{1}{4m}\int \bar{\partial}J\ K[L]\ \bar{\partial}J\right], \qquad K[L] = \frac{J_2(4\sqrt{L})}{\sqrt{L}\ J_1(4\sqrt{L})} \tag{12.6}$$

where $L = -\mathcal{D}\bar{\partial}/m^2$ and $J_1,\ J_2$ are Bessel functions of orders 1 and 2 respectively. Given the possibility of extra terms in (12.5), we may take this as a good approximate solution of the Schrödinger equation.

The kernel K which appears in (12.6) is, despite appearances, very close to the kernel in our solution (9.9). In Fig. 5, we show a comparision of the two kernels.[11].

[11] I thank A. Yelnikov for this comparison graph.

The value of the string tension does not change compared to (9.9), since the low momentum limit of the kernel is the same. However, LMY were able to use this wave functional to calculate a number of glueball masses as well. These were obtained from the two-point function for different color-singlet composite operators, characterized by spin, parity and charge conjugation properties (J^{PC}-notation). The results, in units of $\sqrt{\sigma_F}$, are shown in Table 3. Again, there is reasonable agreement with the lattice data of references [18, 19].

Since two-dimensional Yang-Mills theory leads to an area law for the Wilson loop, it has long been suspected that a similar form for the wave functional might be applicable in higher dimensions [63]. The form of the wave functional we find, namely (9.9), is in accordance with this. We find $\mathcal{F} \sim F^2$ for modes of the field with low momenta, while $\mathcal{F} \sim F(1/\sqrt{-\nabla^2})F$ at high momenta, in agreement with the expected perturbative behavior. An interpolation between these two limiting behaviors, different from our result (9.9), was suggested by Samuel [64] and used to estimate the 0^{++} glueball mass. Essentially the same form (with a small variation) was suggested more recently as a candidate variational ansatz for the wave functional [65]. Specifically it reads $\Psi_0 = e^{\frac{1}{2}\mathcal{F}}$, with

$$\mathcal{F} = -\frac{1}{2}\int_{x,y} F^a_{ij}(x)\left(\frac{1}{\sqrt{-D^2-\lambda_0+M^2}}\right)^{ab}_{x,y} F^b_{ij}(y) \tag{12.7}$$

Here D^2 is the square of the covariant derivative in the adjoint representation, λ_0 is the

State	LMY Calculation	Lattice
0^{++}	4.098	4.065 ± 0.055
0^{++*}	5.407	6.18 ± 0.13
0^{++**}	6.716	7.99 ± 0.22
0^{++***}	7.994	9.44 ± 0.38
0^{--}	6.15	5.91 ± 0.25
0^{--*}	7.46	7.63 ± 0.37
0^{--**}	8.77	8.96 ± 0.65
2^{++}	6.72	6.88 ± 0.16
2^{++*}	7.99	8.62 ± 0.38
2^{++**}	9.26	9.22 ± 0.32
2^{+-}	8.76	8.04 ± 0.50
2^{--}	8.76	7.89 ± 0.35
2^{+-*}	10.04	9.97 ± 0.91
2^{--*}	10.04	9.46 ± 0.46

Table 3: Comparison of glueball mass estimates from [62] and lattice calculations from [18, 19].

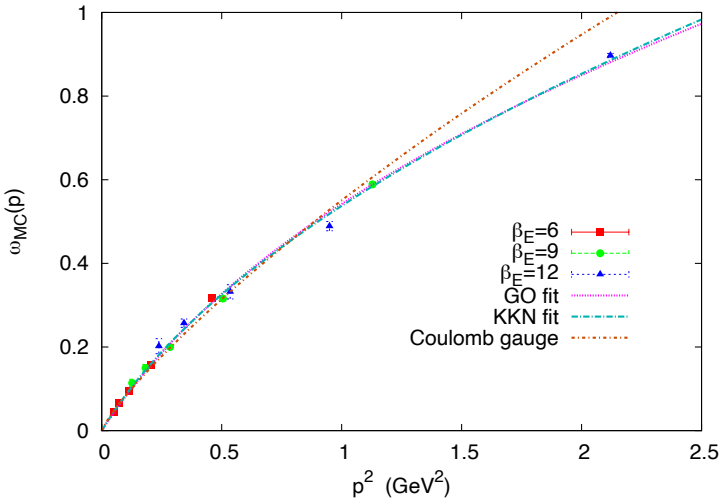


Figure 6: Comparison of cumulative data for ω_{MC} versus p^2 on lattices of extensions $L = 16, 24, 32, 40, 48$ and Euclidean lattice couplings $\beta_E = 6, 9, 12$. The curve labeled "GO fit" refers to the ansatz in [65] while the curve labeled "KKN-fit" refers to our result from (9.9). This graph is from [66].

lowest eigenvalue of $-D^2$ and M is a parameter with the dimension of mass, treated as a variational parameter. A number of quantities can be calculated after fixing M by minimizing the ground state energy. In [66], Monte Carlo simulations of the kernel in a Gaussian ansatz for the wave functional were carried out and then compared against kernels in our wave functional (9.9) and the variational ansatz (12.7), see Fig. 6. For the two cases, the kernel for the quadratic Gaussian part may be written in terms of momentum variable p as

$$\begin{aligned} \omega_{\rm KKN} &= \frac{1}{e^2}\frac{p^2}{\sqrt{p^2+m^2}+m}, \qquad \text{(from (9.9)} \\ \omega_{\rm GO} &= \frac{1}{e^2}\frac{p^2}{\sqrt{p^2+M^2}} \qquad\qquad \text{(from [65])} \end{aligned} \tag{12.8}$$

The fit to the Monte Carlo data is designated as $\omega_{\rm MC}$. From the figure, it seems clear that both agree rather well with the simulations.

The close match between the wave functional we calculated, namely (9.9), and the other candidate functionals is rather nice, but the variational approach comes with a word of caution. The exact vacuum wave functional has zero energy, for reasons of Lorentz invariance, as already explained in the Introduction. For the variational approaches, one calculates the expectation value of the energy and minimizes it with

respect to the variational parameters. This is almost always nonzero, unless one is so lucky as to hit on the exact vacuum functional as the guess for the variational ansatz. Recall that the variational estimate is an upper bound on the true ground state energy, so we generally have $E_{\rm var,min} > 0$. One could try to subtract this out by using $\mathcal{H} - E_{\rm var,min}$ as our notion of a "normally ordered" Hamiltonian, but such a Hamiltonian, if used in the Schrödinger equation and we are able to solve for Ψ_0, we would end up with a negative energy for the true vacuum. In this sense, $\mathcal{H} - E_{\rm var,min}$ will not be the true normally-ordered Hamiltonian to be used in the Schrödinger equation. So, basically, this means that we have to live with a nonzero vacuum energy in the variational approach. Then the question is: Is this acceptable? When we set up a relativistic field theory, the aim is to solve for the vacuum state preserving all the isometries of the spacetime on which it is defined, namely, full Poincaré invariance in flat space. So, with variational ansätze, there is a contradiction between the premise and the end result. (One cannot view $E_{\rm var,min}$ as a cosmological constant or anything of that sort; that would require a spacetime with a different group of isometries, such as the de Sitter or anti-de Sitter space. This would again imply a discord between the starting spacetime and the one consistent with the final results.) This is not to say the variational approach is useless, but it should be used with caution.

13 Extensions of the Yang-Mills theory and some comments

13.1 An intuitive argument for the mass gap

The emergence of the mass gap $m = (e^2 c_A/2\pi)$ is an important feature which allowed for the expansion scheme for solving the Schrödinger equation. It is therefore useful to understand this, not just in technical terms, but in some intuitive way. The key ingredient for this is the measure of integration (5.31). In some ways, its role is already evident in how the propagator mass emerged in section 8. To bring out the connection of the volume element (5.31) and the mass gap, consider the Hamiltonian of the theory written simply as

$$\mathcal{H} = \frac{1}{2e^2} \int \left[(E^a)^2 + (B^a)^2 \right] \tag{13.1}$$

There is a simple argument based on the uncertainty principle which helps us to get a sense of the low lying excitations of this Hamiltonian. The basic commutation rule we need is $[E_i^a, A_j^b] = -i\delta_{ij}\delta^{ab}$. Let us first consider the Abelian case of electrodynamics, with $B = \nabla \times A$. In terms of components E_k and B_k of wave vector (or momentum) k,

this becomes $[E_i, B] = -e^2\epsilon_{ij}k_j$, so that the uncertainty principle reads $\Delta E_k\ \Delta B_k \sim e^2 k$, where $\Delta E_k,\ \Delta B_k$ stand for the root mean square of the fluctuations of the electric field E_k and the magnetic field B_k. The expectation value of the Hamiltonian for a state with wave function Ψ with momentum k is then

$$\begin{aligned}\mathcal{E} &= \frac{1}{2e^2}\langle\Psi_k|\left[E^2+B^2\right]|\Psi_k\rangle = \frac{1}{2e^2}\left[(\Delta E_k^2)+(\Delta B_k^2)\right] \\ &\sim \frac{1}{2}\left(\frac{e^2k^2}{\Delta B_k^2}+\frac{\Delta B_k^2}{e^2}\right)\end{aligned} \tag{13.2}$$

For low lying states, we must minimize this $\mathcal{E}$ with respect to ΔB^2, which gives $\Delta B^2_{k,\min}$ $\sim e^2 k$, giving $\mathcal{E} \sim k$. This corresponds to the familiar photon of the Abelian theory.

For the nonabelian theory, this is inadequate since the expectation value $\langle\mathcal{H}\rangle = \int \Psi^*\,\mathcal{H}\,\Psi$ involves the factor $e^{2\,c_A S_{\rm wzw}(H)}$. In fact,

$$\langle\mathcal{H}\rangle\ \sim \frac{1}{2}\int d\mu(H) e^{2\,c_A\,S_{\rm wzw}(H)}\ (e^2E_k^2 + B_k^2/e^2) \tag{13.3}$$

In terms of B, the behavior of the WZW action is

$$2\,c_A\,S_{\rm wzw}(H) \approx -\frac{c_A}{2\,\pi}\int_k B_{-k}\left(\frac{1}{k^2}\right)B_k + \cdots \tag{13.4}$$

We see that, in the integration measure in (13.3), we have a Gaussian distribution for B with a width of $\Delta B_k^2 \approx \pi k^2/c_A$, for small values of k. Evidently, this Gaussian dominates near small k, since it becomes narrower and narrower as $k \to 0$, giving $\Delta B_k^2 \sim k^2(\pi/c_A)$. Another way to see this is to notice that $B^2 \sim \bar\partial J\,\bar\partial J$ and the currents in the WZW theory obey

$$\langle J^a(x)J^b(y)\rangle \sim \delta^{ab}\partial\frac{1}{(x-y)} \tag{13.5}$$

This translates to $\langle J^a(k)J^b(-k)\rangle \sim (k/\bar k)$, leading to $\Delta B_k^2 \sim k\bar k(\pi/c_A) \sim k^2(\pi/c_A)$ again. What this means is that, even though $\mathcal{E}$ in (13.2) is minimized around $\Delta B_k^2 \sim k$, probability is concentrated around $\Delta B_k^2 \sim k^2(\pi/c_A)$. For the expectation value of the energy, we then find $\langle\mathcal{H}\rangle \sim e^2c_A/2\pi + \mathcal{O}(k^2)$. Thus the kinetic term, in combination with the measure factor $e^{2c_AS_{\rm wzw}(H)}$, leads to a mass gap of order $e^2c_A/2\pi$. The argument is admittedly not rigorous, but does capture the essential physics. The key point is that the volume element (5.31) cuts off the low momentum modes. This suggests that the calculation of the volume element in extensions of the theory with matter content can, by itself, shed light on the issue of the mass gap. In fact, we shall briefly analyze some extensions of the Yang-Mills theory from this point of view. In the case of the Yang-Mills theory modified by the addition of a Chern-Simons term, which will be considered

next, in subsection 13.2, we can carry out the simplification of the Hamiltonian and see how this is indeed realized. We will also consider some supersymmetric extensions in subsection 13.3.

13.2 Yang-Mills-Chern-Simons theory

We consider the Yang-Mills theory modified by the addition of a Chern-Simons (CS) term, so that the action we start with is

$$S = -\frac{1}{4e^2}\int d^3x\ F^a_{\mu\nu}F^{a\mu\nu} - \frac{k}{4\pi}\int d^3x\ \mathrm{Tr}\left(A_\mu\partial_\nu A_\alpha + \frac{2}{3}A_\mu A_\nu A_\alpha\right)\epsilon^{\mu\nu\alpha} \tag{13.6}$$

The second term in S is the integral of the Chern-Simons (CS) 3-form. Although it had been known in the mathematics literature in the context of secondary characteristic classes, it was initially introduced in physics literature as a possible mass term for gauge fields in three dimensions [67, 68]. The parameter k is known as the level number of the Chern-Simons term.

Under a gauge transformation, the CS term changes as

$$CS(A^g) = CS(A) - \frac{k}{4\pi}\int d\mathrm{Tr}(g^{-1}dg\,A) - \frac{k}{12\pi}\int \mathrm{Tr}(g^{-1}dg)^3 \tag{13.7}$$

For transformations g such that $g \to 1$ at the spacetime boundary, we see that the total derivative will integrate to zero. The boundary condition means that $g(x)$ is equivalent to a map from S^3 to G. Such maps can be homotopically nontrivial since $\Pi_3(G)$ is $\mathbb{Z}$ for a simple Lie group, see (6.3). The last term is then 2π times the winding number of this map. For the invariance of the theory under all such gauge transformations, we need invariance of e^{iS}. This requires that the level number k should be an integer. (There are other related ways to understand the quantization of the level number, see for example [69].)

The CS term is odd under parity and time-reversal. Its role as a mass term is made clear by considering the propagator for the theory. In a gauge where $\partial_\mu A_\mu = 0$, we find

$$\langle 0|\,\mathcal{T}A^a_i(x)A^b_j(y)\,|0\rangle = \delta^{ab}\int\frac{d^3p}{(2\pi)^3}e^{-ip(x-y)}\frac{i}{p^2-\mu^2+i\epsilon}\left(\delta_{ij}+i\mu\frac{\epsilon_{ijk}p^k}{p^2}\right) \tag{13.8}$$

where $\mu = (e^2k/4\pi)$. This shows that perturbatively μ is the mass of the gauge particle.

As in the case of the pure Yang-Mills theory, we shall now use the $A_0 = 0$ gauge to set up the Hamiltonian formalism [70]. The canonical momenta can be easily identified

from the action (13.6) and are related to the electric fields $\dot{A}$ as

$$\begin{aligned} E^a &= \frac{e^2}{2}\Pi^a + \frac{ie^2k}{8\pi}A^a = -\frac{ie^2}{2}\frac{\delta}{\delta\bar{A}^a} + \frac{ie^2k}{8\pi}A^a \\ \bar{E}^a &= \frac{e^2}{2}\bar{\Pi}^a - \frac{ie^2k}{8\pi}\bar{A}^a = -\frac{ie^2}{2}\frac{\delta}{\delta A^a} - \frac{ie^2k}{8\pi}\bar{A}^a \end{aligned} \tag{13.9}$$

The commutation rule for $E^a, \bar{E}^a$ is given by

$$[\bar{E}^a(\vec{x}), E^b(\vec{y})] = \frac{e^4k}{8\pi}\delta^{ab}\delta(\vec{x}-\vec{y}) \tag{13.10}$$

The Gauss law operator $G_0(\theta)$ is given by

$$G_0(\theta) = \int \theta^a \left[(D\bar{\Pi} + \bar{D}\Pi)^a + \frac{ik}{4\pi}(\partial\bar{A}^a - \bar{\partial}A^a)\right] \tag{13.11}$$

As before, we can parametrize the fields A, $\bar{A}$ in terms of M and $M^\dagger$, with wave functions taken as functionals of M and $M^\dagger$. The Gauss law operator generates gauge transformations on the argument $(M, M^\dagger)$ of the wave functionals. The Gauss law condition (2.24) for physical states is then equivalent to

$$\Psi(hM, M^\dagger h^{-1}) = \left[1 + \frac{k}{2\pi}\int \mathrm{Tr}\left(M^{\dagger -1}\bar{\partial}M^\dagger\,\partial\theta + \bar{\partial}\theta\,\partial M M^{-1}\right)\right]\,\Psi(M, M^\dagger) \tag{13.12}$$

where $h(x) \approx 1 + \theta(x)$, $\theta = -it^a\theta^a$. The general form of the wave function obeying this condition can be written as

$$\begin{aligned} \Psi(M, M^\dagger) &= \exp\left[\frac{k}{2}\left(S_{\mathrm{wzw}}(M^\dagger) - S_{\mathrm{wzw}}(M)\right)\right]\,\chi(H) \\ &\equiv e^{i\omega(M,M^\dagger)}\chi(H) \end{aligned} \tag{13.13}$$

where χ is gauge-invariant, depending on $M, M^\dagger$ only via the combination $H = M^\dagger M$. $S_{\mathrm{wzw}}(M)$ is the familiar WZW action for M, see 5.21). The Chern-Simons term may be viewed as a "velocity-dependent" potential, since it involves the time-derivative of the A's. The appearance of a phase factor, $e^{i\omega(M,M^\dagger)}$ in (13.13) is in accordance with the fact that the wave functions must carry phase factors when we have velocity-dependent potentials [71]. The gauge-invariant volume element is still as given in section 5, so that the inner product is

$$\langle 1|2\rangle = \int d\mu(H)\, e^{2c_A S_{\mathrm{wzw}}(H)}\,\chi_1^*\,\chi_2 \tag{13.14}$$

We can formulate the Schrödinger equation in terms of $\chi(H)$. But it will turn out to be simpler to use another wave function $\Phi(H) = e^{-\frac{1}{2}kS_{\rm wzw}(H)}\ \chi(H)$. The original wave function Ψ is related to $\Phi(H)$ as

$$\begin{aligned}\Psi &= e^{i\omega(M,M^\dagger)} \exp\left[\tfrac{1}{2}kS_{\rm wzw}(H)\right]\ \Phi(H) \\ &= \exp\left[\frac{k}{2}\left(S_{\rm wzw}(M^\dagger) - S_{\rm wzw}(M) + S_{\rm wzw}(H)\right)\right]\ \Phi(H) \\ &= \exp\left[kS_{\rm wzw}(M^\dagger) - \frac{k}{4\pi}\int A^a\bar{A}^a\right]\ \Phi(H) \end{aligned} \tag{13.15}$$

In going to the last line of this equation from the second, we have used the Polyakov-Wiegmann identity (5.22). The inner product of the states, expressed in terms of Φ's, is given by

$$\langle 1|2\rangle = \int d\mu(H) e^{(k+2c_A)S_{\rm wzw}(H)}\ \Phi_1^*\Phi_2 \tag{13.16}$$

This inner product agrees with what is obtained for the pure Chern-Simons theory as well. Notice that, compared to the pure YM case, the key difference in the integration measure is in the coefficient of the WZW action $S_{\rm wzw}(H)$, with $2c_A \to k + 2c_A$.

The Chern-Simons term is the integral of a differential form. Being independent of the spacetime metric, it does not contribute to the energy-momentum tensor and the Hamiltonian. Thus $\mathcal{H}$ is still of the form

$$\mathcal{H} = \frac{1}{2e^2}\int\left[(E^a)^2 + (B^a)^2\right] \tag{13.17}$$

However, the action of this on the wave functionals is altered as the electric fields have additional terms when expressed in terms of functional derivatives as in (13.9).

There are several steps involved in working out the Hamiltonian or the expression for the kinetic energy as a functional differential operator. First of all, we express the derivatives with respect to A, $\bar{A}$ in terms of the translation operators p_a, $\bar{p}_a$ on M, $M^\dagger$, as in (7.5). In E, $\bar{E}$, we also have the additional terms proportional to A, $\bar{A}$. Finally, for the action on $\Phi(H)$, we need $\mathcal{H} \to e^{-\frac{1}{2}kS_{\rm wzw}(H) - i\omega}\mathcal{H}e^{i\omega + \frac{1}{2}kS_{\rm wzw}(H)}$. Since

$$\bar{p}_a S_{\rm wzw}(M^\dagger) = -\frac{i}{2\pi}\bar{\partial}(\partial M^\dagger M^{\dagger -1})_a, \quad p_a S_{\rm wzw}(M) = -\frac{i}{2\pi}\partial(M^{-1}\bar{\partial}M)_a, \tag{13.18}$$

we find

$$e^{-i\omega}\,T\,e^{i\omega} = \frac{e^2}{2}\int K_{ab}(\vec{x})\left[\int_y \bar{G}(\vec{x},\vec{y})\bar{p}(\vec{y}) - \frac{ik}{4\pi}(\partial H H^{-1})(\vec{x})\right]^a$$

$$\times \left[\int_u G(\vec{x},\vec{u}) p(\vec{u}) + \frac{ik}{4\pi} (H^{-1}\bar{\partial} H)(\vec{x}) \right]^b \tag{13.19}$$

Including the $e^{\frac{1}{2}kS_{\rm wzw}(H)}$ factors, we then obtain

$$\begin{aligned}
\tilde{T} &= \frac{e^2}{2} \int K_{ab}(\vec{x}) \left(\int_y \bar{\mathcal{G}}(\vec{x},\vec{y})\, \bar{p}_a(\vec{y}) - \frac{ik}{2\pi} (\partial H H^{-1})_a \right) \int_u \mathcal{G}(\vec{x},\vec{u})\, p_b(\vec{u}) \\
&= \frac{e^2}{2} \int K_{ab}\, e^{-kS_{\rm wzw}(H)}\, \bar{\mathcal{G}} \bar{p}_a\, e^{kS_{\rm wzw}(H)}\, \mathcal{G} p_b \qquad (13.20) \\
\tilde{T} &\equiv e^{-\frac{1}{2}kS_{\rm wzw}(H)} e^{-i\omega}\, T\, e^{i\omega} e^{\frac{1}{2}kS_{\rm wzw}(H)}
\end{aligned}$$

It is possible to write this in a more symmetric form, similar to the expression (7.7) for the pure Yang-Mills case, but it will not be important for us at this stage. (The relevant expressions are given in [70].)

In the Yang-Mills theory, the observables are all obtained in terms of the current J^a. However, in the YMCS theory, there are additional observables. Notice that the inner product (13.16) expresses matrix elements of operators in terms of a hermitian WZW model of level $(k+2c_A)$. The correlators of the hermitian WZW model are the analytic continuation of the correlators of the level k $SU(N)$ WZW-model with $\kappa = k + c_A$ replaced by $-\kappa = -(k+c_A)$. The level k $SU(N)$ WZW model has integrable primary operators (of finite norm) other than the identity. However, we expect that the vacuum wave function can still be expressed in terms of the currents. For a wave functional Φ which depends on J, rather than H in general, the expression for the Hamiltonian can be written in terms of functional derivatives with respect to J. The result is

$$\begin{aligned}
\tilde{T} &= T_{\rm YM} + \frac{e^2 k}{4\pi} \int J^a \frac{\delta}{\delta J^a} \\
T_{\rm YM} &= \frac{e^2 c_A}{2\pi} \left[\int_u J^a(\vec{u}) \frac{\delta}{\delta J^a(\vec{u})} + \int \Omega_{ab}(\vec{u},\vec{v}) \frac{\delta}{\delta J^a(\vec{u})} \frac{\delta}{\delta J^b(\vec{v})} \right] \\
\Omega_{ab}(\vec{u},\vec{v}) &= \frac{c_A}{\pi^2} \frac{\delta_{ab}}{(u-v)^2} - i \frac{f_{abc} J^c(\vec{v})}{\pi(u-v)} + \mathcal{O}(\epsilon) \qquad (13.21)
\end{aligned}$$

We see that the coefficient of the $\int J\delta/\delta J$-term is $(k+2c_A)e^2/4\pi$, giving a mass of this value to every factor of J^a. The perturbative mass μ gets corrected by the addition of $(e^2 c_A/2\pi)$. This is also consistent with the shift $2c_A \to k+2c_A$ in the integration measure in (13.16), and also in accordance with the intuitive argument for the mass gap given in subsection 13.1.

The expression (13.21) for $\tilde{T}$ can be rewritten as

$$\begin{aligned}\tilde{T} &= \tilde{m}\int_u J^a(\vec{u})\frac{\delta}{\delta J^a(\vec{u})} \;+\; m\int \Omega_{ab}(\vec{u},\vec{v})\frac{\delta}{\delta J^a(\vec{u})}\frac{\delta}{\delta J^b(\vec{v})} \\ &= \tilde{m}\left[\int_u \xi^a(\vec{u})\frac{\delta}{\delta \xi^a(\vec{u})} \;+\; \int \Omega_{ab}(\vec{u},\vec{v})\frac{\delta}{\delta \xi^a(\vec{u})}\frac{\delta}{\delta \xi^b(\vec{v})}\right] \end{aligned} \tag{13.22}$$

where $\tilde{m} = (k+2c_A)e^2/4\pi$ and $\xi = \sqrt{\tilde{m}/m}\; J$. The potential energy $\int B^2$ is as it was in the pure YM case.

Given the similarity of these expressions to the what we obtained for the YM theory, we can use the expansion scheme outlined in section 9 and work out the lowest order vacuum wave functional as

$$\begin{aligned}\Phi_0 &= \exp\left[-\frac{\pi}{\tilde{m}c_A}\int \bar{\partial}\xi\left(\frac{1}{\tilde{m}+\sqrt{\tilde{m}^2-\nabla^2}}\right)\bar{\partial}\xi\right] \\ &= \exp\left[-\frac{\pi}{mc_A}\int \bar{\partial}J\left(\frac{1}{\tilde{m}+\sqrt{\tilde{m}^2-\nabla^2}}\right)\bar{\partial}J\right] \\ &\approx \exp\left(-\frac{1}{4g^2}\int F^2\right)\end{aligned} \tag{13.23}$$

where $g^2 = \tilde{m}e^2 = e^4(k+2c_A)/4\pi$. In the last line, we give the expression for the modes of low momentum or long wave length. For an observable $\mathcal{O}$ (involving long wave length modes of the fields), the expectation value is thus

$$\begin{aligned}\langle\mathcal{O}\rangle &\approx \int d\mu(H) e^{(k+2c_A)S_{\rm wzw}(H)}\; e^{-\frac{1}{4g^2}\int F^2}\;\mathcal{O} \approx \int d\mu\,(\mathcal{C})\;\; e^{kS_{\rm wzw}(H)} e^{-\frac{1}{4g^2}\int F^2}\;\mathcal{O} \\ &\approx \int [dQ] d\mu\,(\mathcal{C})\;\; e^{-\mathcal{S}}\;\mathcal{O} \\ \mathcal{S} &= \int d^2x\;\left[\frac{1}{4g^2}F^a_{\mu\nu}F^{a\mu\nu} + \sum_{i=1}^{k}\bar{Q}^i\gamma\cdot DQ^i\right]\end{aligned} \tag{13.24}$$

where we have used the fact that $e^{kS_{\rm wzw}(H)}$ can also be expressed in terms of integration over fermions (in two dimensions) as

$$\begin{aligned}e^{S_{\rm wzw}(H)} &= \det(-D\bar{D}) = \int [dQ] e^{-\int(\bar{Q}_L D Q_L + \bar{Q}_R \bar{D} Q_R)} \\ &= \int [dQ]\; e^{\int \bar{Q}\gamma\cdot DQ}\end{aligned} \tag{13.25}$$

Here Q's are fermions in the fundamental representation of $SU(N)$ and we use k flavors to get the factor $e^{kS_{\rm wzw}(H)}$.

Effectively, we have the expectation value to be computed in a two-dimensional theory for the YM theory coupled to k flavors of fermions in the fundamental representation. (We are not saying that there are fermions in the spectrum of the theory; (13.24) is just a useful way of expressing expectation values.) These fermions can screen charges in any representation and hence the the expectation value of the Wilson loop will not display an area law. Already, at the level of perturbation theory, we have seen that the Chern-Simons term in the action (13.6) acts as a mass term for the gluons. Therefore we should expect that the interaction energy between charges cannot be of long range. The fact that we do not obtain an area law for the Wilson loop is entirely in accordance with this expectation.

13.3 *Supersymmetric theories

In this subsection we will make a few observations about supersymmetric Yang-Mills theories. (This analysis is heavily based on [72, 73].) In subsection 13.1 we have seen that the mass gap is closely related to the integration measure used for the inner product of the wave functionals. We have also seen that this correlation holds also for the YMCS theory where the inner product involves the integral of $e^{(k+2c_A)S_{\rm wzw}(H)}$, see (13.16), and the mass $\tilde{m} = e^2(k+2c_A)/4\pi$. Further, the measure (13.16) is the same for the pure CS theory and for the YMCS theory. With these observations in mind, we see that we can make some statements regarding supersymmetric theories via the following strategy. We calculate the integration measure for the supersymmetric CS theory, taking $k \to 0$ to obtain the YM case. It is also possible to identify the integration measure without detailed calculations by using a set of indirect, although slightly intricate, arguments. For this, we will first consider Hamiltonian analysis of the level k CS theory coupled to a set of point charges, the charge matrices being t_a is some representation r of the group G, following [9]. The action is given by

$$\begin{aligned} S &= -\frac{k}{4\pi}\int d^3x\ \mathrm{Tr}\left(A_\mu\partial_\nu A_\alpha + \frac{2}{3}A_\mu A_\nu A_\alpha\right)\epsilon^{\mu\nu\alpha} \\ &\quad -i\sum_r \int dt(-it_a)_r(A_0^a(x_r) + A_i^a\dot{x}_r^i) \end{aligned} \tag{13.26}$$

The canonical commutation rules are given by

$$[\bar{A}^a(x), A^b(y)] = \frac{2\pi}{k}\delta^{ab}\delta^{(2)}(x-y) \tag{13.27}$$

We take functionals of A to describe the states, with

$$\bar{A}^a(x)\,\Psi(A) = \frac{2\pi}{k}\frac{\delta}{\delta A^a(x)}\,\Psi(A) \tag{13.28}$$

The equation of motion for A_0^a or the Gauss law is given by

$$\left[\frac{ik}{2\pi}(\partial\bar{A} - \bar{\partial}A + [A,\bar{A}])^a - \sum_r (t_a)_r\,\delta^{(2)}(x-x_r)\right]\Psi = 0 \tag{13.29}$$

This can be rewritten as

$$\delta_\theta\,\Psi(A) = \int \theta^a\,\left[\frac{k}{2\pi}\bar{\partial}A + \sum_r (-it_a)_{(r)}\delta^{(2)}(x-x_r)\right]\Psi \tag{13.30}$$

where $\delta_\theta\,\Psi(A)$ denotes the change of $\Psi(A)$ under an infinitesimal gauge transformation. In the absence of charges, the solution of this equation is

$$\Psi_0 = N_0\,e^{kS_{\rm wzw}(M)} \tag{13.31}$$

where the normalization factor N_0 can be identified from[12]

$$|N_0|^2\int d\mu(H)\,\exp\left[(k+2\,c_A)S_{\rm wzw}(H)\right] = 1 \tag{13.32}$$

Now consider two point-charges, conjugate to each other, at positions $\vec{x}_1$ and $\vec{x}_2$. The solution to the Gauss law condition is

$$\Psi = N(z_1,z_2)\,M(x_1)\,M^{-1}(x_2)\,\Psi_0 \tag{13.33}$$

where $N(z_1,z_2)$, which depends on the positions of the charges, is determined by the requirement that Ψ should obey the Schrödinger equation. The Hamiltonian corresponding to (13.26) is given as

$$\mathcal{H} = \sum_r (t_a)_r\left[\frac{2\pi}{k}\dot{\bar{z}}_r\frac{\delta}{\delta A^a(x_r)} + \dot{z}_r A^a(x_r)\right] \tag{13.34}$$

When $\mathcal{H}$ acts on Ψ we will encounter singular terms due to terms like $\delta M(x_1)/\delta A(x_1)$. The properly regularized version leads to two interesting features. First, k in the expression (13.34) is shifted to $\kappa = k + c_A$. (In the Hamiltonina framework, this shift has

[12]The action (13.26) also shows that the canonical two-form $(ik/2\pi)\int \delta\bar{A}\,\delta A$ is a Kähler form associated with the Kähler potential $K = (ik/2\pi)\int \bar{A}A$. The normalization of Ψ in the coherent state basis we are using then requires a factor of e^{-K} in the integration measure. (For more details, see [9].) This term combined with $e^{kS_{\rm wzw}(M)}$ and its conjugate (from Ψ_0 and Ψ_0^*) leads to the $e^{kS_{\rm wzw}(H)}$ factor in (13.32).

been obtained in [74, 9].) The Schrödinger equation is then identical to the Knizhnik-Zamolodchikov (KZ) equation [75] for the chiral blocks of the WZW theory with parameter κ. Thus $N(z_1, z_2)$ becomes a chiral block of the level k $SU(N)$-WZW theory.

Finally, we consider the normalization of the state Ψ in (13.33). We can write the required integral as

$$|N(z_1, z_2)|^2 \int d\mu(H)\, e^{\tilde{k}\, S_{\rm wzw}(H)}\; H(x_1)\, H(x_2)^{-1} = 1 \tag{13.35}$$

We have put in the measure factor with an arbitrary coefficient $\tilde{k}$ for $S_{\rm wzw}(H)$ to show how we can determine it. There are two points we can make about this normalization:

1. The integral in (13.35) will yield the correlator $\langle H(x_1)\, H(x_2)^{-1}\rangle$ of the hermitian WZW model (for $SL(N, \mathbb{C})/SU(N)$) of level $\tilde{k}$. The (z_1, z_2)-dependence of this correlator must exactly cancel the similar dependence of $|N(z_1, z_2)|^2$ to allow for a proper normalization of (13.33). So $\langle H(1)\, H(2)^{-1}\rangle$ should be given by the solution of the KZ equation with $\kappa = (k + c_A) \to -\kappa = -(k + c_A)$.

2. At the same time, we also know that the correlator $\langle H(x_1)\, H(x_2)^{-1}\rangle$ of the hermitian theory (of level $\tilde{k}$) is the same as the corresponding correlator of the $SU(N)$ theory of level $-\tilde{k}$ [11].

These two statements together imply that

$$-\bar{k} \;+\; c_A = -\,(k + c_A) \tag{13.36}$$

We see that this determines $\tilde{k}$ to be $k + 2c_A$, as expected.

This illustrates how we can obtain the measure in more general cases. First we calculate the shift in the level number k to identify the KZ parameter κ. This can be done via the Hamiltonian method, or in an even simpler way, by straightforward use of Feynman diagrams [76]. Once this is done, the compatibility of the two requirements given above for $\langle H(x_1)\, H(x_2)^{-1}\rangle$ will be

$$\left.\begin{array}{c}\text{KZ parameter of } SU(N) \\ \text{WZW theory of level } -\,\tilde{k}\end{array}\right\} = \left\{\begin{array}{l}\text{KZ parameter of } SU(N) \\ \text{WZW theory of level } k\end{array}\right. \tag{13.37}$$

We are now ready to look at supersymmetric theories. From diagrammatic calculation (for which the supersymmetric Yang-Mills term may be viewed as a regulator), the

KZ parameters are given as [77]

$$k \to \begin{cases} k + c_A & \mathcal{N} = 0 \\ k + \frac{1}{2} c_A & \mathcal{N} = 1 \\ k & \mathcal{N} \geq 2 \end{cases} \tag{13.38}$$

The normalization of the wave functionals for the supersymmetric YMCS theories are thus given by

$$\langle 1|2\rangle = \int d\mu(H)\, \exp[\tilde{k}\, S_{\rm wzw}(H)]\, d\mu[\text{Fermions}]\, \Psi_1^*\, \Psi_2 \tag{13.39}$$

$$\tilde{k} = \begin{cases} k + 2\, c_A & \mathcal{N} = 0 \\ k + c_A & \mathcal{N} = 1 \\ k & \mathcal{N} \geq 2 \end{cases} \tag{13.40}$$

For $\mathcal{N} = 0$, we can take $k = 0$ and obtain the result for the pure YM case. For $\mathcal{N} = 1$, we cannot take $k = 0$ since there is a parity anomaly, so we need $k = 1$ as the minimal choice for a consistent theory [78, 79]. In this case, the value of $\tilde{k}$ suggest that there will be a mass gap, of a magnitude different form the case of $\mathcal{N} = 0$. For $\mathcal{N} \geq 2$, we can take $k = 0$. For these cases, we should expect that there will be no mass gap.

These statements are in accordance with expectations from other analyses. For the $\mathcal{N} = 4$ case, constraints of unbroken supersymmetry prevent a mass term [78], but a partial spontaneous breaking of the gauge symmetry is possible. For $\mathcal{N} = 2$ theories, no mass gap is expected, but there may be no stable supersymmetric vacuum [78, 80, 81]. The absence of mass gap for $\mathcal{N} = 2$ has also been analyzed by different methods in [82, 83].

While the arguments presented above for the measure bypassed direct calculations, one can ask whether the same result is obtained in a straightforward Hamiltonian formulation of the supersymmetric theories. This is indeed the case, as discussed in some detail in [72]. Here we will briefly indicate the steps to highlight a subtle point in obtaining the Hamiltonian. The classical action for the $\mathcal{N} = 1$ theory is given by

$$S = -\frac{1}{4\, e^2} \int F^a_{\mu\nu} F^{a\mu\nu} - \frac{i}{2e^2} \int \bar{\psi}^a (\gamma^\mu D_\mu \psi)^a \tag{13.41}$$

The supersymmetry transformation is given by

$$\delta_\epsilon A^a_\mu = -i\, \bar{\epsilon}\, \gamma_\mu \psi^a, \quad \delta_\epsilon \psi^a = \frac{1}{2} F^a_{\mu\nu} \gamma^{\mu\nu} \epsilon \tag{13.42}$$

The action is invariant under this transformation with the supercharges given by

$$Q^\dagger = \int (i\psi^\dagger \gamma^i \frac{\delta}{\delta A^i} + \frac{1}{e^2} \psi^\dagger B), \quad Q = \int (i\gamma^i \psi \frac{\delta}{\delta A^i} + \frac{1}{e^2} \psi B) \tag{13.43}$$

Q is a two-component spinor, and we make the identification $Q^1 = q$, $Q^2 = q^\dagger$. As mentioned before, the parity anomaly will make the partition function of this theory vanish, rendering it trivial or inconsistent [78, 79]. To get a consistent theory, we must include a supersymmetric Chern-Simons term

$$S_{\rm SCS} = -\frac{k}{4\pi}\int d^3x\ {\rm Tr}\left[\left(A_\mu\partial_\nu A_\alpha - \frac{2}{3}A_\mu A_\nu A_\alpha\right)\epsilon^{\mu\nu\alpha} + ie^2\bar\psi\psi\right] \tag{13.44}$$

The full action is thus $S_{\rm SYM} = S + S_{\rm SCS}$. Being a supersymmetric theory, the Hamiltonian can be obtained as the anticommutator of supercharges. Towards this, we first define the gauge-invariant wave function $\Phi(H)$ as in (13.15),

$$\Psi = e^{i\omega(M,M^\dagger)}\exp\left[\tfrac{1}{2}kS_{\rm wzw}(H)\right]\ \Phi(H) \tag{13.45}$$

The supercharge in terms of its action on Φ is given by

$$q' = i\int \chi^{\dagger a}(\mathcal{G}p)^a - \frac{1}{e^2}\frac{2\pi}{c_A}\int \chi^a\bar\partial J^a \tag{13.46}$$

where $\mathcal{G}$ is the regularized version of the Green's function $G = \partial^{-1}$ and χ is the gauge-invariant version of the fermion field defined by $\chi^b = (M^{-1})^{ab}\psi^a$, $\chi^{b\dagger} = \psi^{a\dagger}M^{ab}$.

We take the integration measure for the inner product of the Φ's to be given in terms of

$$d\mu = d\mu(H)\ \exp\left[(k + (2-n)c_A)\,S_{\rm wzw}(H)\right] \tag{13.47}$$

For the present case, $n = 1$, but we will leave it arbitrary for now. The adjoint of the supercharge, which is consistently the adjoint with (13.47) defining the integration measure, is

$$q'^\dagger = -i\int \chi^a\left((\bar{\mathcal{G}}\bar p)^a - i\frac{k}{2\pi}(\partial HH^{-1})^a + i\frac{nc_A}{2\pi}(\partial HH^{-1})^a\right) - \frac{1}{e^2}\frac{2\pi}{c_A}\int(\chi^a\bar\partial J^a)^\dagger \tag{13.48}$$

Recall that, by virtue of the physical states being annihilated by the Gauss law operator, we were able to eliminate E^a in favor of $\bar E^a$ and the currents, in the simplification of the kinetic energy operator, see equations (7.13) to (7.22). Equivalently, we can eliminate $\bar p^a$ in favor of p^a. Effectively, this amounts to the statement

$$\bar p^a = (Kp)^a + \frac{1}{e^2}f^{alm}(K\chi^\dagger)^l\chi^m \tag{13.49}$$

When this is used in (13.48), we have to move χ^a to the right end to obtain normal ordering. This results in a singular term $\bar{\mathcal{G}}(x,x)$, exactly the same kind of term we encountered in section 7. Evaluating it as before, we end up with

$$-i\int\chi^a(x)(\bar{\mathcal{G}}\bar p)^a(x) = -\,i\int\chi^a(x)(\bar{\mathcal{G}}Kp)^a(x) - \int\chi^a(x)J^a(x)$$

$$+\frac{i}{e^2}\int \bar{\mathcal{G}}^{ab}(x,y)f^{blm}(K\chi^\dagger)^l(y)\chi^a(x)\chi^m(y) \tag{13.50}$$

The $\chi^a J^a$-term arises from the normal ordering mentioned above. When this is used for $q'^\dagger$ in (13.48) and the anticommutator is taken, we find the gauge-invariant form of the Hamiltonian for the supersymmetric theory as

$$\begin{aligned}\mathcal{H}=\frac{1}{2}\{q',q'^\dagger\}=&\frac{e^2c_A}{2\pi}\left(\int J^a\frac{\delta}{\delta J^a}+\int\Omega^{ab}(x,y)\frac{\delta}{\delta J^a(x)}\frac{\delta}{\delta J^b(y)}\right)+\frac{2\pi^2}{e^2c_A^2}\int(\bar\partial J^a\bar\partial J^a)\\&+\frac{e^2(k-nc_A)}{4\pi}\int J^a\frac{\delta}{\delta J^a}-\frac{ic_A}{2\pi}\int f^{abc}\bar G(x,y)K^{cs}(y)\chi^{s\dagger}(y)\chi^b(y)\frac{\delta}{\delta J^a(x)}\\&-\frac{1}{e^2}\int(\chi^\dagger\bar{\mathcal{D}}_J\chi^\dagger-\chi\mathcal{D}_J\chi)+\left(\frac{e^2}{4\pi}(k+2c_A-nc_A)\right)\int\frac{1}{e^2}\chi^{\dagger a}(K^{-1})^{ab}\chi^b\end{aligned} \tag{13.51}$$

Notice the equality of the masses for the J's and the χ's, as expected for a supersymmetric theory.[13]

14 Entanglement in Yang-Mills (2+1)

Entanglement is a property of the state and can be characterized by a reduced density matrix obtained by integrating $\Psi^*[\varphi]\Psi[\varphi']$ over fields in some subregion of space. So it would seem that if there is any feature of the quantum theory for which wave functions provide a better framework than manifestly covariant methods, it would be entanglement. And this is indeed the case, although, for ease of calculation a path integral with a cut on (the unintegrated) part of space is often used (with a replica trick as well). In the case of gauge theories, this led to the identification of an extra term in the entanglement entropy, known as the contact term (or Kabat term) [84], compared to what is expected for matter fields. Here we will consider the contact term for YM(2+1) in a Hamiltonian formulation and relate it to something familiar in mathematics literature known as the BFK gluing formula [85]. (This analysis is basically taken from [86].)

14.1 Entanglement in Maxwell theory

It is simpler and conceptually more clarifying to consider the Maxwell theory first. The Gauss law condition takes the form $G_0 = \nabla\cdot E = 0$. We will choose a conjugate constraint $\chi = \nabla\cdot A$. In general, if we have constraints ζ_i and conjugate constraints χ_j,

[13] It may be useful to keep in mind that the anticommutator is given as $\{\chi^a,\chi^{b\dagger}\}=e^2\delta^{ab}\delta^{(2)}(x-y)$.

the Hamiltonian path integral is given by

$$Z = \int [d\mu]\, \delta(\zeta)\delta(\chi)\, \det[\{\zeta_i, \chi_j\}]\, e^{iS} \tag{14.1}$$

Here $d\mu$ is the phase space measure of integration, the constraints are enforced by δ-functions and we also need the determinant of the Poisson brackets of the constraints. S is the action expressed in terms of the phase space variables. For the Maxwell theory we thus get

$$Z = \int [d\mu]\, \delta(\nabla \cdot E)\delta(\nabla \cdot A)\ \det[-\nabla^2]\ e^{iS} \tag{14.2}$$

We have already set $A_0 = 0$, so that the phase space variables are E_i and A_i, $i = 1, 2$. We consider the theory in some region of space V with a boundary $\partial\mathrm{V}$. The fields can be parametrized as

$$A_i = \partial_i\theta + \epsilon_{ij}\partial_j\phi, \qquad E_i = \dot{A}_i = \partial_i\sigma + \epsilon_{ij}\partial_j\Pi \tag{14.3}$$

We separate the fields into a bulk part and a boundary part by writing

$$\begin{aligned}
\theta(x) &= \tilde{\theta}(x) + \oint_{\partial\mathrm{V}} \theta_0(y)\, n\cdot\partial G(y,x), \qquad \phi(x) = \quad \tilde{\phi}(x) + \oint_{\partial\mathrm{V}} \phi_0(y)\, n\cdot\partial G(y,x) \\
\sigma(x) &= \tilde{\sigma}(x) + \oint_{\partial\mathrm{V}} \sigma_0(y)\, n\cdot\partial G(y,x), \qquad \Pi(x) = \tilde{\Pi}(x) + \oint_{\partial\mathrm{V}} \Pi_0(y)\, n\cdot\partial G(y,x)
\end{aligned} \tag{14.4}$$

The tilde-fields all obey Dirichlet conditions, vanishing on $\partial\mathrm{V}$. The values of the fields on the boundary are designated with a subscript 0 and are continued into the interior of V via Laplace's equation, i.e.,

$$\nabla_x^2 \oint_{\partial\mathrm{V}} \theta_0(y)\, n\cdot\partial G(y,x) = 0 \tag{14.5}$$

The Green's function $G(y,x)$ for the Laplace operator also obeys Dirichlet conditions. The decomposition of fields as in (14.4) follows from Green's theorem. The canonical one-form is given by $\mathcal{A} = \int E_i \delta A_i$ and by direct substitution of (14.4), we find

$$\begin{aligned}
\mathcal{A} &= \int E_i\, \delta A_i \\
&= \int_{\mathrm{V}} \Big[(-\nabla^2\tilde{\sigma})\, \delta\tilde{\theta} + \tilde{\Pi}\, \delta B\Big] + \oint \mathcal{E}\, \delta\theta_0(x) + \oint Q\, \delta\phi_0(x)
\end{aligned} \tag{14.6}$$

where $B = -\nabla^2\tilde{\phi}$ is the magnetic field. Also $\mathcal{E}$ and Q are given by

$$\mathcal{E}(x) \;\; = \;\; \oint_y \sigma_0(y) M(y,x) + \partial_\tau \Pi_0(x)$$

$$Q(x) = \oint_y \Pi_0(y) M(y,x) - \partial_\tau \sigma_0(x) \tag{14.7}$$

$$M(x,y)_{\rm I} = n\cdot\partial_x\, n\cdot\partial_y G(x,y)\Big|_{x,\ y\ on\ \partial \rm V} \tag{14.8}$$

$M(x,y)$ is what is usually referred to the Dirichlet-to-Neumann operator. and $\partial_\tau = n_i\epsilon_{ij}\partial_j$ denotes the tangential derivative on the boundary. $\mathcal{E}$ and Q are not independent, but are related by

$$\mathcal{C} = \partial_y \oint \mathcal{E}_{\rm I}(x) M^{-1}(x,y) + Q_{\rm I}(x) = 0 \tag{14.9}$$

In the sense of Dirac's theory of constraints, $\mathcal{C}$ is of the first class; one can choose a conjugate constraint $\phi_0 = 0$ and eliminate the pair, so that

$$\mathcal{A} = \int_{\rm V} \Big[(-\nabla^2\tilde\sigma)\,\delta\tilde\theta + \tilde\Pi\,\delta B\Big] + \oint_{\partial \rm V} \mathcal{E}\,\delta\theta_0 \tag{14.10}$$

The phase volume associated with this canonical structure is[14]

$$d\mu = [d\tilde\sigma d\tilde\theta]\,[d\mathcal{E}\,d\theta_0]\,[d\tilde\Pi dB]\ \det(-\nabla^2) \tag{14.11}$$

The constraints entering the path integral (14.2) can be written out as

$$\delta(\nabla\cdot E) = (\det(-\nabla^2))^{-1}\,\delta(\tilde\sigma), \quad \delta(\nabla\cdot A) = (\det(-\nabla^2))^{-1}\,\delta(\tilde\theta) \tag{14.12}$$

Since $\tilde\theta$ vanishes on ∂V, we are imposing the Gauss law with test functions vanishing on ∂V. The value of θ on the boundary and its conjugate $\mathcal{E}$ represent physical degrees of freedom. Using (14.11) and (14.12), we see that we can set $\tilde\sigma = \tilde\theta = 0$, and all factors of $\det(-\nabla^2)$ cancel out, so that Z in (14.2) becomes

$$Z = \int [d\mathcal{E}\,d\theta_0]\,[d\tilde\Pi dB]\ e^{iS} \tag{14.13}$$

The action S also involves only the fields Π, B, $\mathcal{E}$, θ_0. We see that we have a theory of the bulk fields Π, B, which constitute a single bulk field, with "edge modes" described by $\mathcal{E}$, θ_0.

Now consider going through this procedure for a region of space divided into two with an interface (dashed line) as shown in Fig. 7. We construct the theory on the whole space and then in $\rm V_1$ and $\rm V_2$ separately, put them together and compare the results. For the theory on the whole space, the result is basically as we have already discussed, with

$$Z_{\rm whole} = \int [d\sigma d\theta]\,[d\Pi dB]\,[\det -\nabla^2]\,\delta(\nabla\cdot E)\,\delta(\nabla\cdot A)\ \det[-\nabla^2]\ e^{i\mathcal{S}}$$

[14]The determinants of $-\nabla^2$ are calculated with Dirichlet boundary conditions.

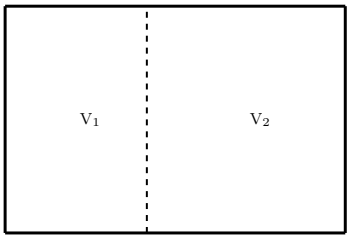


Figure 7: Showing division of space into two regions for entanglement considerations

$$= \int [d\Pi dB]\; e^{iS} \tag{14.14}$$

We do not consider any edge modes for the boundary of the whole space since our focus will be on entanglement across the interface. (They can be included without changing the essence of the argument.)

Now consider building the theory separately in V_1 and V_2. The fields on the interface can be continued into V_1 and V_2 again using Laplace's equation, so that we have

$$\theta(x) = \begin{cases} \tilde{\theta}_1(x) + \oint_{\partial \mathrm{V}_1} \theta_0(y)\; n\cdot\partial G_1(y,x) & \text{in } \mathrm{V}_1 \\ \tilde{\theta}_2(x) + \oint_{\partial \mathrm{V}_2} \theta_0(y)\; n\cdot\partial G_2(y,x) & \text{in } \mathrm{V}_2 \end{cases} \tag{14.15}$$

with similar expressions for the other fields. G_1 and G_2 are Green's functions for the Laplacian for regions V_1 and V_2, respectively, vanishing on the interface. The phase volume has the form

$$d\mu_{\text{split}} = [d\tilde{\sigma} d\tilde{\theta}]_1\, [d\tilde{\sigma} d\tilde{\theta}]_2\; \det(-\nabla^2)_1\; \det(-\nabla^2)_2\, [d\mathcal{E} d\theta_0] \times [d\mu_{\Pi,B}\text{-part}] \tag{14.16}$$

The key issue is about the constraints. Using $f,\ h$ for test functions, with boundary values on the interface designated as $f_0,\ h_0$, respectively, the constraints are

$$\begin{aligned} \int \partial_i f\, E_i &= \int_{\mathrm{V}_1} \tilde{f}_1(-\nabla^2\tilde{\sigma}_1) + \int_{\mathrm{V}_2} \tilde{f}_2(-\nabla^2\tilde{\sigma}_2) + \oint f_0\, \mathcal{E} \approx 0 \\ \int \partial_i h\, A_i &= \int_{\mathrm{V}_1} \tilde{h}_1(-\nabla^2\tilde{\theta}_1 + \int_{\mathrm{V}_2} \tilde{h}_2(-\nabla^2\tilde{\theta}_2) + \oint h_0\, (M_1 + M_2)\, \theta_0 \approx 0 \end{aligned} \tag{14.17}$$

$M_1,\ M_2$ are the Dirichlet-to-Neumann operators for G_1 and G_2, respectively. For the theory on the full space, θ-dependence is eliminated everywhere including the interface, so each term in (14.17 must vanish separately and the δ-functions for the constraints must be interpreted as

$$\delta(\nabla\cdot E)\; \delta(\nabla\cdot A) = \delta[-\nabla^2\tilde{\sigma}_1]\, \delta[-\nabla^2\tilde{\sigma}_2]\; \delta[-\nabla^2\tilde{\theta}_1]\, \delta[-\nabla^2\tilde{\theta}_2]$$

$$\times\delta[\mathcal{E}]\,\delta[(M_1+M_2)\theta_0] \tag{14.18}$$

We also have the BFK gluing formula [85][15]

$$\det(-\nabla^2) = \det(-\nabla^2)_1\ \det(-\nabla^2)_2\ \det(M_1+M_2) \tag{14.19}$$

If we use results (14.18) and (14.19), we get back to (14.14) as expected; splitting the fields is only a more involved way of writing the path integral for the full space.

Consider now integrating out fields in V_2. Since the interface is a boundary to V_2, from the point of view of the theory in V_2, we can only impose the Gauss law with test functions which vanish on the interface. The edge modes $\mathcal{E}$, θ_0 are physical degrees of freedom. Thus we must take $f_0 = h_0 = 0$, and the constraints become

$$\delta(\nabla\cdot E)\ \delta(\nabla\cdot A) = \delta[-\nabla^2\tilde{\sigma}_1]\,\delta[-\nabla^2\tilde{\sigma}_2]\ \ \delta[-\nabla^2\tilde{\theta}_1]\,\delta[-\nabla^2\tilde{\theta}_2] \tag{14.20}$$

The determinant $\det(M_1+M_2)$ is not canceled out and the reduced theory in V_1 takes the form

$$Z_{\mathrm{red}} = \det(M_1+M_2)\int [d\mathcal{E}\,d\theta_0]\,d\mu_{\Pi,B}\,e^{i\mathcal{S}} \tag{14.21}$$

There is an extra factor $\det(M_1+M_2)$; since this is part of the phase volume, it is to be considered as a degeneracy factor. Thus if we define a reduced density matrix, it takes the form

$$\rho = \frac{\mathbb{1}}{\det(M_1+M_2)}\ (\rho_{\mathrm{bulk}})_{\mathrm{red}} \tag{14.22}$$

where $(\rho_{\mathrm{bulk}})_{\mathrm{red}}$ refers to the reduced density matrix for all the remaining physical degrees of freedom and $\mathbb{1}$ is a matrix such that $\mathrm{Tr}\,\mathbb{1} = \det(M_1+M_2)$.

This determinant $\log\det(M_1+M_2)$ is Kabat's contact term. Its origin is due to the simple fact that in the full space Gauss law eliminates θ, $\mathcal{E}$, but for the theory in each region, these are not eliminated. The contact term can also be identified as the interface term in the BFK gluing formula.

14.2 The case of YM(2+1)

It is now straightforward to consider the situation for the Yang-Mills theory. Since we phrased the discussion given above in the language of gauge-fixing, the simplest way

[15] This formula tells us that if a Riemannian manifold is separated into V_1, V_2, etc. by suitable hypersurfaces, then the determinant of the Laplacian for the full space can be obtained as the product of similar determinants with Dirichlet boundary conditions in each of the regions V_1, V_2, etc. times a set of interface contributions which are the determinants of the Dirichlet-to-Neumann operators.

for us is to eliminate E from our considerations using the Gauss law as we did in section 7, see equations (7.13 to (7.22). This is like a complex gauge-fixing, since $M^\dagger$ gets set to 1. The canonical one-form is $\mathcal{A} = \int E_i^a \delta A_i^a = -4 \int \mathrm{Tr}(\bar{E}\,\delta A + E\,\delta \bar{A})$ and the Gauss law takes the form

$$\mathcal{G}^a = 2(\bar{D}E + D\bar{E})^a \tag{14.23}$$

As the conjugate constraint, we take $\chi^a = (D\bar{A})^a$. Eliminating E, the canonical one-form can then be written as

$$\mathcal{A} = -4 \int \mathrm{Tr}\Big[\bar{E}\,\delta A + \mathcal{G}(x)\,(-D\bar{D})^{-1}_{x,y}\,\delta\chi(y)\Big] \tag{14.24}$$

The corresponding phase volume is

$$d\mu = \det[(-D\bar{D})^{-1}]\;[d\bar{E}dA]\,[d\mathcal{G}d\chi] \tag{14.25}$$

We see, in a way similar to what happens in the Maxwell case, that we will get $\det[(-D\bar{D})]_1$ and $\det[(-D\bar{D})]_2$ for V_1 and V_2, and $\det[(-D\bar{D})]_{\mathrm{V}_1\cup\mathrm{V}_2}$ for the full space $\mathrm{V} = \mathrm{V}_1 \cup \mathrm{V}_2$. The contact term is then given by

$$S_{\mathrm{contact}} = \log\left[\frac{\det(-D\bar{D})_{\mathrm{V}_1\cup\mathrm{V}_2}}{\det(-D\bar{D})_{\mathrm{V}_1}\,\det(-D\bar{D})_{\mathrm{V}_2}}\right] \tag{14.26}$$

Unlike the Abelian case, this expression depends on the fields. So one has to carry out an averaging over the physical fields, i.e., do the integration over H, to calculate the entropy. If we ignore the field dependence, the contribution of (14.26) is the same as the result for (dim$G$ copies of) the Abelian theory.

A noteworthy point is the following. The operator $(-D\bar{D})$ which comes into the contact term is independent of the mass, even though the theory does have a mass gap. In a massive theory, the entanglement tends to vanish as the mass becomes large. But in the gauge theory the contact term will lead to nonzero entanglement even in the large m limit.

15 Appendix A: Conventions and Notations

Summation over repeated indices is assumed. Greek letters μ, ν, etc. are used to denote spacetime components, taking values 0, 1, 2, 3 in $(3+1)$-dimensional spacetime, and 0, 1, 2 in $(2+1)$-dimensional spacetime. The metric for flat Minkowski space is denoted by $\eta_{\mu\nu}$; the contravariant version is denoted by $\eta^{\mu\nu}$. The components of $\eta_{\mu\nu}$ are given by $\eta_{00} = 1$, $\eta_{ij} = -\delta_{ij}$, and $\eta_{0i} = \eta_{i0} = 0$.

We will also use the abbreviation $\partial_\mu = \frac{\partial}{\partial x^\mu}$. The scalar product of two vectors with components A_μ and B_ν will be written as $A \cdot B = \eta^{\mu\nu} A_\mu B_\nu = A_0 B_0 - A_i B_i$. In some cases, such as in writing $e^{ip\cdot x}$ we often abbreviate the scalar product as just $px \equiv p_0 x_0 - p_i x_i$.

The Levi-Civita symbol in three dimensions is ϵ^{ijk} which is totally antisymmetric under exchange of any two indices and is normalized as $\epsilon^{123} = 1$. $\epsilon^{\mu\nu\alpha\beta}$ is defined in a similar way, with $\epsilon^{0123} = 1$.

The symbol ∂ is also used to denote the boundary of a spatial or spacetime region. Thus ∂V denotes the boundary of the region V. Differential forms will be used for certain discussion and have the usual expression in terms of a coordinate basis. Thus if B denotes a differential p-form, it has the local coordinate expression

$$B = \frac{1}{p!} B_{\mu_1 \mu_2 \cdots \mu_p}\, dx^{\mu_1} \wedge dx^{\mu_2} \cdots \wedge dx^{\mu_p} \tag{A1}$$

with the wedge symbol, as usual signifying the antisymmetrization of the coordinate differentials. The symbol d will be used for the exterior derivative of a differential form,

$$dB = \frac{1}{p!} (\partial_\mu B_{\mu_1 \mu_2 \cdots \mu_p})\, dx^{\mu} \wedge dx^{\mu_1} \wedge dx^{\mu_2} \cdots \wedge dx^{\mu_p} \tag{A2}$$

While a large part of the discussions will use flat space, there will be occasions to discuss some curved manifolds. The appropriate metric will be given as the occasion arises.

For spinors, we will need the Dirac γ-matrices; these are defined by

$$\gamma^\mu \gamma^\nu + \gamma^\nu \gamma^\mu = 2\, \eta^{\mu\nu}\, \mathbb{1} \tag{A3}$$

In the case of nonchiral spinors in four dimensions, γ's can be realized as 4×4 matrices. Thus the $\mathbb{1}$ on the right hand side of (A3) denotes the 4×4 identity matrix. A specific choice for the γ's is

$$\gamma^0 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad \gamma^i = \begin{pmatrix} 0 & \sigma^i \\ -\sigma^i & 0 \end{pmatrix} \tag{A4}$$

Each entry in the matrices in (A4) is a 2×2 matrix. σ^i are the Pauli matrices given as

$$\sigma^1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma^2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma^3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{A5}$$

In three dimensions (or $2+1$ dimensions), the spinors are nonchiral and the γ's can be realized as 2×2 matrices. A specific choice is $\gamma^0 = \sigma^3$, $\gamma^1 = i\sigma^1$, $\gamma^2 = i\sigma^2$.

For the group $SU(N)$, the generators of the Lie algebra in the fundamental (N-dimensional) representation are denoted by t_a, $a = 1, 2, \cdots, \dim G = N^2 - 1$. They are taken to be normalized as $\mathrm{Tr}(t_a t_b) = \frac{1}{2}\delta_{ab}$. The commutation rules are $[t_a, t_b] = i f_{abc} t_c$. The quadratic Casimir operator has the value $c_F = (N^2-1)/2N$ for the fundamental representation and $c_A = N$ for the adjoint representation.

16 Appendix B: Regularization

16.1 The regularized form of the operators

We will now go over some of the issues related to defining the regularized form of the operators for the kinetic and potential energies, and the Hamiltonian. A good regularization procedure must preserve gauge invariance. With our choice of variables, it is also important to preserve the holomorphic invariance. This property was discussed at the end of section 4. The matrices $(M, M^\dagger)$ and $(M', M'^\dagger)$ where $M' = M\bar{V}$, $M'^\dagger = VM^\dagger$ will give the same potentials $(A, \bar{A})$. Here V is a holomoprhic function of the coordinates, $\bar{V}$ is an antiholomoprhic function. Generally, we need this freedom in how we define M and $M^\dagger$, so that configurations can be represented on various coordinate patches without singularities. The calculations we do will involve the Green's functions for $D = \partial + A$ and $\bar{D} = \bar{\partial} + \bar{A}$. These were introduced in section 5 in the form

$$D^{-1}(x,y) = M(x)G(x,y)M^{-1}(y), \quad \bar{D}^{-1}(x,y) = M^{\dagger -1}(x)\bar{G}(x,y)M^\dagger(y) \tag{B1}$$

where G and $\bar{G}$ are the Green's functions for ∂ and $\bar{\partial}$, respectively. For particular coordinate patch, we can take these to be

$$G(x,y) = \frac{1}{\pi(\bar{x}-\bar{y})}, \quad \bar{G}(x,y) = \frac{1}{\pi(x-y)} \tag{B2}$$

Consider the construction of $D^{-1}(x,y)$ and $\bar{D}^{-1}(x,y)$ using M' and $M'^\dagger$. These Green's functions are unchanged if we define

$$G'(x,y) = \bar{V}^{-1}(x)\, G(x,y)\, \bar{V}(y), \quad \bar{G}'(x,y) = V(x)\, \bar{G}(x,y) V^{-1}(y) \tag{B3}$$

Notice that these will still satisfy the required equations

$$\partial_x G(x,y) = \bar{\partial}_x \bar{G}(x,y) = \delta^{(2)}(x-y) \tag{B4}$$

We see that the use of different forms for the matrices M, $M^\dagger$ must be accompanied by the use of different definitions for the Green's functions G and $\bar{G}$.

Our aim is to use a point-splitting regularization for the Green's functions $G(x,y)$ and $\bar{G}(x,y)$ which preserves the transformation property (B3). This can be done by use of the Gaussian approximation to the Dirac δ-function given in (7.11), namely,

$$\sigma(\vec{x},\vec{y};\epsilon) = \frac{e^{-|\vec{x}-\vec{y}|^2/\epsilon}}{\pi\epsilon} \tag{B5}$$

Based on this we define

$$\begin{aligned} \mathcal{G}(\vec{x},\vec{y}) &= \int_u G(\vec{x},\vec{u})\sigma(\vec{u},\vec{y};\epsilon)K^{-1}(y,\bar{u})K(y,\bar{y}) \\ \bar{\mathcal{G}}(\vec{x},\vec{y}) &= \int_u \bar{G}(\vec{x},\vec{u})\sigma(\vec{u},\vec{y};\epsilon)K(u,\bar{y})K^{-1}(y,\bar{y}) \end{aligned} \tag{B6}$$

Since we will be using the matrices M, $M^\dagger$, H in the adjoint representation for most of the calculations, we have given these expressions in the appropriate form. Here $K_{ab} = 2\mathrm{Tr}(t_a H t_b H^{-1})$ is the same as the matrix H but in the adjoint representation. It is easy enough to verify that $\mathcal{G}$ and $\bar{\mathcal{G}}$ have the same transformation as G and $\bar{G}$. By expanding the K's in (B6), it is possible to carry out the integration and reduce these to the form

$$\begin{aligned} \mathcal{G}_{ma}(\vec{x},\vec{y}) &= G(\vec{x},\vec{y})[\delta_{ma} - e^{-|\vec{x}-\vec{y}|^2/\epsilon}\big(K^{-1}(y,\bar{x})K(y,\bar{y})\big)_{ma}] \\ \bar{\mathcal{G}}_{ma}(\vec{x},\vec{y}) &= \bar{G}(\vec{x},\vec{y})[\delta_{ma} - e^{-|\vec{x}-\vec{y}|^2/\epsilon}\big(K(x,\bar{y})K^{-1}(y,\bar{y})\big)_{ma}] \end{aligned} \tag{B7}$$

The coincident point limit of $\bar{\mathcal{G}}$ can be read off from these as

$$\bar{\mathcal{G}}(x,x) = -\frac{\partial K K^{-1}}{\pi} \tag{B8}$$

Correspondingly, we have

$$\bar{D}^{-1}(\vec{x},\vec{x})_{\mathrm{reg}} = -\frac{1}{\pi}M^{\dagger -1}(\vec{x})(\partial K K^{-1})M^\dagger(\vec{x}) = \frac{1}{\pi}(A - M^{\dagger -1}\partial M^\dagger)(\vec{x}) \tag{B9}$$

This reproduces the results in equations (5.19) and (5.20) used in the calculation of the volume element for $\mathcal{C}$.

Turning to the kinetic energy operator, we start with the form given in (7.7),

$$T = \frac{e^2}{4}\int_x e^{-2c_A S_{\rm wzw}(H)}\Big[\bar{\mathcal{G}}\bar{p}_a(\vec{x})K_{ab}(\vec{x})\,e^{2c_A S_{\rm wzw}(H)}\mathcal{G}p_b(\vec{x}) + \mathcal{G}p_a(\vec{x})K_{ba}(\vec{x})\,e^{2c_A S_{\rm wzw}(H)}\bar{\mathcal{G}}\bar{p}_b(\vec{x})\Big] \tag{B10}$$

We have put in the regularized form of the Green's functions. In this expression, we use the abbreviation

$$\mathcal{G}p_b(\vec{x}) = \int_u \mathcal{G}_{bc}(\vec{x},\vec{u})p_c(\vec{u}), \quad \bar{\mathcal{G}}\bar{p}_a(\vec{x}) = \int_u \bar{\mathcal{G}}_{ac}(\vec{x},\vec{u})\bar{p}_c(\vec{u}), \ \text{ etc.} \tag{B11}$$

In (B10), in moving p_a and $\bar{p}_a$ to the right, we encounter the commutators

$$\bar{\mathcal{G}}[\bar{p}_a(\vec{x}), K_{ab}(\vec{x})\,e^{2c_A S_{\rm wzw}(H)}], \quad \mathcal{G}[p_a(\vec{x}), K_{ba}(\vec{x})\,e^{2c_A S_{\rm wzw}(H)}] \tag{B12}$$

These involve the coincident point limit of the Green's functions. Using (B7) we can calculate the commutators and see that they are zero as $\epsilon \to 0$; see [13] for more details. The expression for T can then be brought to the form

$$\begin{aligned} T &= \frac{e^2}{2}\int \Pi_{rs}(\vec{u},\vec{v})\bar{p}_r(\vec{u})p_s(\vec{v}) \\ \Pi_{rs}(\vec{u},\vec{v}) &= \int_x \bar{\mathcal{G}}_{ar}(\vec{x},\vec{u})K_{ab}(\vec{x})\mathcal{G}_{bs}(\vec{x},\vec{v}) \end{aligned} \tag{B13}$$

This is the regularized version of (7.6), thus establishing its equivalence with (7.7) as well.

The action of p and $\bar{p}$ on the current J is given by

$$\begin{aligned} [p_s(\vec{v}),\ J_a(\vec{z})] &= -i\frac{c_A}{\pi}K_{as}(\vec{z})\partial_z\delta(\vec{z},\vec{v}) \\ [\bar{p}_r(\vec{u}),\ J_b(\vec{w})] &= -i(\mathcal{D}_w)_{br}\delta(\vec{w}-\vec{u}), \quad \mathcal{D}_{w\ ab} = \frac{c_A}{\pi}\partial_w\delta_{ab} + if_{abc}J_c(\vec{w}) \end{aligned} \tag{B14}$$

We can then use the form of T from (B13) and work out its action on a functional of the currents; basically this involves using the chain rule and the commutators (B14). The result is

$$T\ \Psi(J) = m\left[\int_z \omega_a(\vec{z})\frac{\delta}{\delta J_a(\vec{z})} + \int_{z,w}\Omega_{ab}(\vec{z},\vec{w})\frac{\delta}{\delta J_a(\vec{z})}\frac{\delta}{\delta J_b(\vec{w})}\right]\Psi(J) \tag{B15}$$

where

$$\omega_a(\vec{z}) = -if_{arm}\big[\partial_z\Pi_{rs}(\vec{u},\vec{z})\big]_{\vec{u}\to\vec{z}}\,K^{-1}_{sm}(\vec{z}) = if_{arm}\Lambda_{rm}(\vec{u},\vec{z})\big|_{\vec{u}\to\vec{z}}$$

$$\begin{aligned}\Omega_{ab}(\vec{z},\vec{w}) &= -\Big[\Big[\frac{c_A}{\pi}\partial_w\delta_{br}+if_{brm}J_m(\vec{w})\Big]\ \partial_z\Pi_{rs}(\vec{w},\vec{z})\Big]K^{-1}_{sa}(\vec{z})\\ &= \mathcal{D}_{w\ br}\Lambda_{ra}(\vec{w},\vec{z}) \qquad\qquad (B16)\\ \Lambda_{ra}(\vec{w},\vec{z}) &= -(\partial_z\Pi_{rs}(\vec{w},\vec{z}))K^{-1}_{sa}(\vec{z}) \qquad\qquad (B17)\end{aligned}$$

Using $\Pi_{rs}(\vec{u},\vec{v})$ from (B13), the expression for Λ_{ra} can be written out as

$$\begin{aligned}\Lambda_{ra}(\vec{w},\vec{z}) &= \int_x \bar{\mathcal{G}}_{mr}(\vec{x},\vec{w})G(\vec{x},\vec{z})e^{-|\vec{x}-\vec{z}|^2/\epsilon}\Big[\frac{\bar{x}-\bar{z}}{\epsilon}K(x,\bar{x})K^{-1}(z,\bar{x})\\ &\qquad +K(x,\bar{x})\partial_z(K^{-1}(z,\bar{x})K(z,\bar{z}))K^{-1}(z,\bar{z})\Big]_{ma} \qquad\qquad (B18)\end{aligned}$$

Because of the exponential $e^{-|\vec{x}-\vec{z}|^2/\epsilon}$, the region $|\vec{x}-\vec{z}|^2\lesssim\epsilon$ is what is relevant for $\omega_a(\vec{z})$. Expanding around z, we get

$$\omega_a(\vec{z}) = J_a(\vec{z})+\mathcal{O}(\epsilon) \qquad\qquad (B19)$$

If we use the expression (B7) for $\bar{\mathcal{G}}_{mr}(\vec{x},\vec{w})$ in (B18), the expression for Λ_{ra} will split into four terms. One can expand the integrands in powers of $x-w$, $\bar{x}-\bar{w}$ and carry out the x-integration to generate an expansion in powers of ϵ. We then find

$$\begin{aligned}\Lambda_{ra}(\vec{w},\vec{z}) &= \frac{1}{\pi(z-w)}\big[\delta_{ra}-\big(K(\vec{w})K^{-1}(z,\bar{w})\big)_{ra}e^{-|z-w|^2/2\epsilon}\big]\\ &\qquad +(\text{terms of higher order in }\epsilon\text{ or }(z-w),\ (\bar{z}-\bar{w}))\\ &\equiv \bar{\mathcal{G}}'_{ra}(\vec{z},\vec{w})+\cdots \qquad\qquad (B20)\end{aligned}$$

Here $\bar{\mathcal{G}}'$ is the transpose of $\bar{\mathcal{G}}$ with ϵ replaced by 2ϵ. We can use this expression in (B16) for $\Omega_{ab}(\vec{z},\vec{w})$ and write the kinetic energy operator as

$$\begin{aligned}T\Psi(J) &= m\left[\int J_a(\vec{z})\frac{\delta}{\delta J_a(\vec{z})}+\int\big(\mathcal{D}_w\bar{\mathcal{G}}'(\vec{z},\vec{w})\big)_{ab}\frac{\delta}{\delta J_a(\vec{w})}\frac{\delta}{\delta J_b(\vec{z})}\right]\Psi(J)+\mathcal{O}(\epsilon)\\ &= m\int_{z,w}\left[\bar{\partial}J_a(\vec{w})\bar{G}(\vec{z},\vec{w})\frac{\delta}{\delta J_a(\vec{z})}+\big(\mathcal{D}_w\bar{\mathcal{G}}'(\vec{z},\vec{w})\big)_{ab}\frac{\delta}{\delta J_a(\vec{w})}\frac{\delta}{\delta J_b(\vec{z})}\right]\Psi(J)+\mathcal{O}(\epsilon)\\ &\qquad\qquad (B21)\end{aligned}$$

It is easy to see that as $\epsilon\to 0$, $\Lambda(\vec{w},\vec{z})\to\bar{G}(\vec{z},\vec{w})$. The first line of (B21) then reproduces the expression (7.10) in text. In simplifying (B16) for ω_a to get (B19), we have cancelled powers of $(z-w)$ against $\bar{G}(\vec{z},\vec{w})$. This can lead to expressions which seemingly do not have the holomorphic invariance. The second line of (B21) shows T in a manifestly holomorphic invariant form.

As for the potential energy, we can do a point-splitting and write

$$\begin{aligned} V_{(\lambda')} &= \frac{\pi}{mc_A}\int_x :\bar\partial J_a(\vec x)\bar\partial J_a(\vec x): \\ &= \frac{\pi}{mc_A}\Big[\int_{x,y}\sigma(\vec x,\vec y;\lambda')\bar\partial J_a(\vec x)(K(x,\bar y)K^{-1}(y,\bar y))_{ab}\bar\partial J_b(\vec y) - \frac{c_A \mathrm{dim}G}{\pi^2\lambda'^2}\Big] \end{aligned} \tag{B22}$$

A priori we have the freedom to choose a different value λ', rather than ϵ, for the width of $\sigma(\vec x,\vec y;\lambda')$, so we have displayed the expression for such a choice. The action of T on V is important for solving the Schrödinger equation. Since we have regularized all operators, it is straightforward to work this out and obtain

$$T_{(\epsilon)}\ V_{(\lambda')} = 2m\left[1+\tfrac{1}{2}\log(\lambda'/2\epsilon)\right]\ V_{(\lambda')}\ +\cdots \tag{B23}$$

Since it is different from λ', we display the regularization parameter ϵ for T as a subscript.

To understand how ϵ and λ' may be related, we first note that since we are using the $A_0=0$ gauge, the Coulomb potential at short distances will obtained from the action of the kinetic term on wave functionals. In 2+1 dimensions, the Coulomb potential is logarithmic and so a subtraction point needs to be chosen to define the zero of the potential. The freedom of choosing this point is also a reflection of the fact that the kinetic operator is scale invariant as $\epsilon\to 0$. In terms of the regularized version, this means that we can define

$$\begin{aligned} T_{(\lambda)} &= T_{(\epsilon)} + \frac{e^2}{2}\log(2\epsilon/\lambda)\ \mathcal{Q} \\ \mathcal{Q} &= \epsilon\int \sigma(\vec u,\vec v;\epsilon)K_{rs}(u,\bar v)\left(\bar p_r(\vec u) - i\bar\partial J_r(\vec u)\right)p_s(\vec v) \end{aligned} \tag{B24}$$

Acting on V, we now get

$$T_{(\lambda)}V_{(\lambda')} = 2m\left[1+\tfrac{1}{2}\log(\lambda'/\lambda)\right]\ V_{(\lambda')}\ +\cdots \tag{B25}$$

The addition of $\mathcal{Q}$ to T can be interpreted as follows. Recall that, in covariant perturbation theory, the addition of local counterterms is equivalent to a change of regularization. The addition of $\mathcal{Q}$ in (B24) may be viewed as the analogous procedure for the Hamiltonian formulation.

16.2 A Lorentz-invariance argument

We have regularized T and V with different parameters, λ (or ϵ) and λ'. Each expression will thus involve fields with modes of momenta larger than $1/\sqrt{\lambda}$ and $1/\sqrt{\lambda'}$,

respectively. Being short-distance regularization parameters, we need them to be much smaller than any physical scales such as $1/(e^2)^2$.

The key missing ingredient in treating T and V separately is Lorentz invariance. Recall that under a Lorentz transformation corresponding to velocity v_i, the electric and magnetic fields transform as

$$\delta E_i \approx -\epsilon_{ij} v_j B, \quad \delta B \approx \epsilon_{ij} v_i E_j, \qquad |v| \ll 1 \tag{B26}$$

The Hamiltonian is the integral of the energy density T_{00}. The momentum and stress densities are given by

$$T_{0i} = \epsilon_{ij} E_j B, \quad T_{ij} = -E_i E_j + \delta_{ij} T_{00} \tag{B27}$$

Under a Lorentz transformation, we have

$$\delta T_{0i} = v_i T_{00} + v_j T_{ij} \tag{B28}$$

If we use the transformation (B26) for the fields,

$$\delta T_{0i} = \delta(\epsilon_{ij} E_j B) = v_i (E^2 + B^2) - v_k E_k \, E_i \tag{B29}$$

This is in agreement with (B28). But it also shows that if we regularize the momentum $P_i = \int \epsilon_{ij} E_j B$ with a parameter λ, Lorentz invariance will require that both terms in the energy (on the right hand side of (B29) should be regularized with the same parameter. (A variant of this argument was given in [87].) Thus, although $a\ priori$ we could use different regularizations for T and V, consistency with Lorentz symmetry requires $\lambda = \lambda'$ (with $1/\sqrt{\lambda} \gg e^2$). In this case, (B25) simplifies to

$$T_{(\lambda)} V_{(\lambda)} = 2m \, V_{(\lambda)} \ + \cdots \tag{B30}$$

At this point, we can also go back and use ϵ in place of λ as we did for the potential energy in (7.25).

Another version of this argument, where the action of T on V and the vacuum wave functional are considered, is given in [87]. We may note that the issue of regularization and how it relates to Lorentz invariance is somewhat tricky. For a discussion with different points of view, see [88, 89].

17 Appendix C: Corrections to string tension

In this Appendix, we will calculate the first set of corrections to the formula (10.10) for the string tension using the first order corrections to the vacuum wave functional

obtained in section 9. From the recursive solution of the Schrödinger equation, the correction to the quadratic kernel; in the wave function is

$$\begin{aligned} e^2\ f_2^{(2)}(q) &= \frac{m}{E_q}\int \frac{d^2k}{32\pi}\ \left(\frac{1}{\bar{k}}\ g^{(3)}(q,k,-k-q)\ +\frac{k}{2\bar{k}}\ g^{(4)}(q,k;-q,-k)\right) \\ &\approx \frac{\bar{q}^2}{2m}\,(1.1308)+\ldots \end{aligned} \tag{C1}$$

In the second line we give the lowest order (quadratic in q^2) term, as this is what is relevant for the calculation of the string tension. Seemingly, this is a 113% correction, but there are important additional terms which should be included. In calculating the the vacuum expectation value of an operator as $\langle\mathcal{O}\rangle = \int \Psi^*\Psi\,\mathcal{O}$, we have to do a functional integration over H, so this can be viewed as a two-dimensional field theory. It is then clear that there are loop corrections, in the 2d field theory sense, to the quadratic kernel. So we start with a procedure for simplifying and systematizing these contributions.

Since the measure of integration has the WZW action, our first step will be to transform the functional integration over $\Psi_0^*\Psi_0 = e^{\mathcal{F}}$ into the integration over a two-dimensional chiral boson field $\varphi,\ \bar{\varphi}$. (Although we use the same letter, this is not the φ we used in parametrizing H as $e^{t_a\varphi^a}$.) The key point is that $\mathcal{F}$ is given in terms of currents, so consider the calculation of the current correlators in just the hermitian WZW theory. We can write

$$\begin{aligned} \frac{1}{\mathcal{Z}}\int d\mu(H) e^{2c_A S_{\rm wzw}(H)} e^{-\frac{c_A}{\pi}\int \bar{C}^a(\partial H H^{-1})^a} &= \frac{1}{\mathcal{Z}}\int d\mu(H) e^{2c_A S_{\rm wzw}(UH)-2c_A S_{\rm wzw}(U)} \\ &= e^{-2c_A S_{\rm wzw}(U)} \end{aligned} \tag{C2}$$

where $\bar{C} = U^{-1}\bar{\partial}U$ and we have used the Polyakov-Wiegmann identity

$$S_{\rm wzw}(H) - \frac{1}{\pi}\int {\rm Tr}(\bar{C}\partial H H^{-1}) = S_{\rm wzw}(UH)\ -\ S_{\rm wzw}(U) \tag{C3}$$

(Here $\mathcal{Z}$ is just a normalization factor; it is the partition function of the hermitian WZW theory.) Since $\exp(-2c_A S_{wzw}(U))$ is the inverse of the chiral Dirac determinant in two dimensions, we can use the formula

$$\exp(-2c_A S_{wzw}(U)) = \int [d\varphi d\bar{\varphi}]\ e^{-\int \bar{\varphi}(\bar{\partial}+\bar{C})\varphi} \tag{C4}$$

The complex boson field φ transforms in the adjoint representation of $SU(N)$. For the YM case, including $\Psi_0^*\Psi_0$, we have

$$\langle\mathcal{O}\rangle = \int d\mu(H)\, e^{2c_A S_{\rm wzw}(H)} e^{\mathcal{F}}(J)\,\mathcal{O}(J)$$

$$= \left[\mathcal{O}(\hat{J})\, e^{F(\hat{J})}\right] \int d\mu(H)\, e^{2c_A S_{\rm wzw}(H)}\, e^{-\frac{c_A}{\pi}\int \bar{C}^a(\partial H H^{-1})^a}\Big]_{\bar{C}=0}$$
$$= \int [d\varphi d\bar{\varphi}]\; e^{-S(\varphi)}\; \mathcal{O}(\sqrt{2\pi/mc_A}\;\bar{\varphi}t^a\varphi) \tag{C5}$$

where $\hat{J}^a = -\sqrt{2\pi/mc_A}\;\frac{\delta}{\delta \bar{C}^a}$ and, after introducing the representation (C4), we have evaluated the action of the $\hat{J}^a$'s and set $\bar{C}$ to zero. The action $S(\varphi)$ is given by

$$S(\varphi) = \int \bar{\varphi}\bar{\partial}\varphi - \mathcal{F}(\sqrt{2\pi/mc_A}\;\bar{\varphi}t^a\varphi) \tag{C6}$$

There is a correction to be made to this formula once we have $\mathcal{F}$ which will introduce additional interactions for the φ-field. If we think of this as a 2d field theory, it is easy to see that we will need renormalization constants (Z-factors) for the chiral boson action. The representation of the determinant which is applicable in the presence of interactions is thus

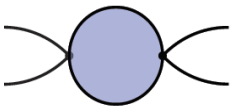
$$\exp(-2c_A S_{wzw}(U)) = \int [d\varphi d\bar{\varphi}]\;\exp\left[-\int \bar{\varphi}(Z_2\bar{\partial} + Z_1\bar{C})\varphi\right] \tag{C7}$$

For the expectation values, we still get the formula (C5), but now with the action

$$S(\varphi) = \int\; (Z_2\bar{\varphi}\bar{\partial}\varphi + Z_1\bar{\varphi}\bar{C}\varphi)\;-\;\mathcal{F}(Z_1\sqrt{2\pi/mc_A}\;\bar{\varphi}t^a\varphi) \tag{C8}$$

(We can set $\bar{C}$ to zero at the end, once the renormalization constants Z_1, Z_2 have been calculated; see [17] for more details.) In this representation in terms of φ, effectively, the current J^a is replaced by $Z_1\sqrt{2\pi/mc_A}\;\bar{\varphi}t^a\varphi$. The function $\mathcal{F}(Z_1\sqrt{2\pi/mc_A}\;\bar{\varphi}t^a\varphi)$ contains vertices, $\mathcal{F}^{(2)}$ with two currents (quartic in φ, $\bar{\varphi}$), $\mathcal{F}^{(3)}$ with three currents, etc. For example, we may diagrammatically represent $\mathcal{F}^{(2)}$, with two φ's and two $\bar{\varphi}$'s, as

$$\mathcal{F}^{(2)} = \tfrac{2\pi}{mc_A}\int (\bar{\varphi}t^a\varphi)_x\, f^{(2)}(x,y)(\bar{\varphi}t^a\varphi)_y \;=$$



The corrections to $\mathcal{F}^{(2)}$, which is what we are interested in, may be viewed as loop corrections to the quartic vertex in this two-dimensional field theory.

In calculating the corrections to $\mathcal{F}^{(2)}$, the vertices $\mathcal{F}^{(3)}$, $\mathcal{F}^{(4)}$, etc., can be included perturbatively since they carry powers of e. However, the lowest term in the vertex $\mathcal{F}^{(2)}$, corresponding to $f^{(2)}_{0a_1a_2}(x_1,x_2)$, has no powers of e and hence its contributions will have to be included to all orders and summed up. The result for the current-current

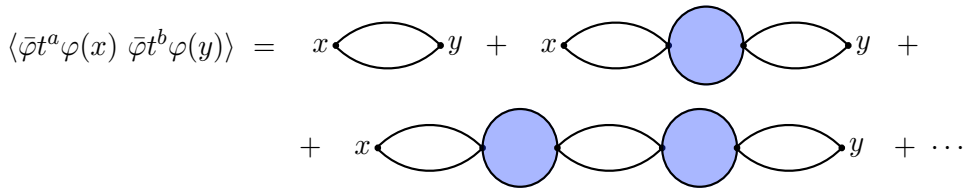


Figure 8: The current-current correlator including all contributions from $\mathcal{F}_0^{(2)}$

correlator is

$$\langle \bar{\varphi} t^a \varphi(x)\ \bar{\varphi} t^b \varphi(y) \rangle = \delta^{ab} \frac{c_A}{\pi} \int \frac{d^2k}{(2\pi)^2} e^{ik(x-y)}\ \frac{k}{\bar{k}} \left(\frac{m}{E_k} \right) \tag{C9}$$

Here $E_k = \sqrt{k^2 + m^2}$; the (m/E_k) factor arises from the summation of corrections due to $\mathcal{F}_0^{(2)}$, shown diagrammatically in Fig. 8. Any vertex can acquire a series of corrections from $\mathcal{F}_0^{(2)}$, so that we may consider an effective vertex

$$\bar{\varphi} t^a \varphi(x) \Big]_{\text{eff}} = \int \frac{d^2k}{(2\pi)^2} e^{ik(x-z)}\ \frac{m}{E_k}\ (\bar{\varphi} t^a \varphi)(z) \tag{C10}$$

This is shown diagrammatically in Fig. 9. As a result, all corrections acquire powers of (m/E_k) in the integrands and in fact, we can classify contributions in powers of (m/E_k). Since $(m/E_k) \leq 1$, the numerical values will decrease as we go down the series.

The basic strategy for calculating corrections may then be summarized as follows:

1. Construct loop diagrams generated by $\mathcal{F}^{(3)}$ (3 factors of $\bar{\varphi} t^a \varphi$) and $\mathcal{F}^{(4)}$ (4 factors of $\bar{\varphi} t^a \varphi$).

2. They can have arbitrary insertions of $\mathcal{F}_0^{(2)}$'s, leading to a factor of (m/E_k), as in Fig. 10.

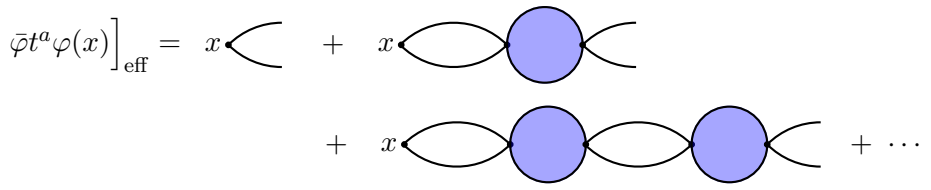


Figure 9: The effective current vertex

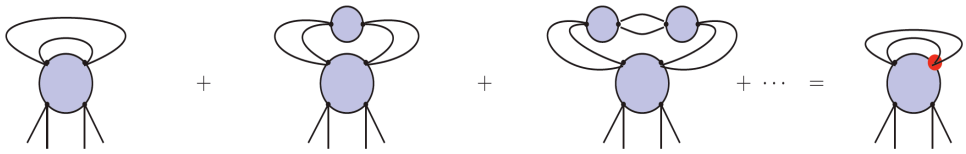


Figure 10: Corrections from $F_0^{(2)}$ summed up as a factor of m/E_k (shaded circle at vertex) and sample renormalization diagrams

3. Sum up $\mathcal{F}_0^{(2)}$ insertions in all diagrams (of order e^2) generated by $\mathcal{F}^{(3)}$ and $\mathcal{F}^{(4)}$.

4. Classify and group these by the number of factors of (m/E_k).

There will be corrections generated to the terms $\bar{\varphi}\bar{\partial}\varphi$ and $\bar{\varphi}\bar{C}\varphi$ in the action; these are renormalization effects due to $\mathcal{F}_0^{(2)}$. These have to be cancelled by Z_1 Z_2 factors. They are discussed in more detail in [17].

We have calculated corrections to order e^2 and up to 4 powers of (m/E_k). Denoting the factors of (m/E_k) by shaded circles at the vertices, the corrections to the low momentum limit of $f^{(2)}$ may summarized as in Fig. 11. We show the coefficients of $\bar{q}^2/2m$, for small $q,\ \bar{q}$ for each diagram. Let C_n denote the partial sum of corrections up to terms with $(m/E_k)^n$, starting with $C_0 = 1.1308$ from the recursive procedure (C1). We then

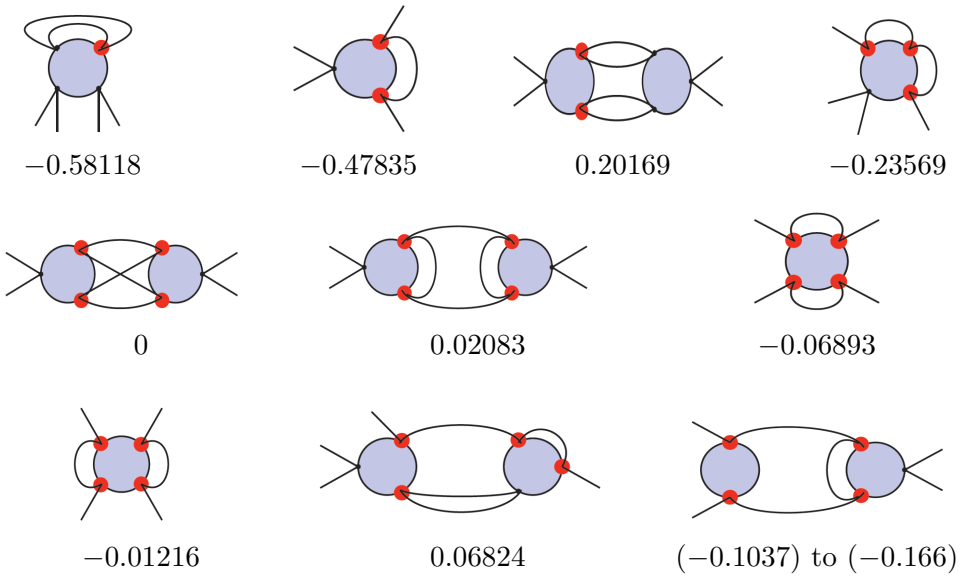


Figure 11: Corrections to the low momentum limit of the $\mathcal{F}^{(2)}$ vertex

find

$$\begin{aligned} C_1 &= 0.5496 \\ C_2 &= 0.2730 \\ C_3 &= 0.0373 \\ C_4 &= -0.05843 \text{ to } -0.00583 \end{aligned} \tag{C11}$$

Many of the integrals have to be evaluated numerically. There is a small ambiguity in one of the integrals for the last diagram in Fig. 11 [17], so that C_4 is in the range indicated.

Notice that the partial sums are systematically decreasing in value, showing that the ordering of diagrams by powers of m/E_k does constitute a viable expansion. The cumulative value of the corrections to the order we have calculated is indeed small. For the string tension, we then find

$$\sqrt{\sigma_R} = e^2 \sqrt{\frac{c_A c_R}{4\pi}} \begin{cases} (1 - 0.02799 + \cdots) \\ (1 - 0.0029 + \cdots) \end{cases} \tag{C12}$$

This correction, of the order of -2.8% to -0.03%, is entirely consistent with lattice calculations. Terms of order $(m/E_k)^5$ are expected to contribute at the level of a fraction of 1%.

The corrections seem to play out in an intriguing way here. The first term C_0 is rather large, but is then mostly cancelled out by the "loop corrections". This suggests that there must be a different way to organize the corrections so that many of the cancellations are packaged together.

There are also several types of corrections we have not calculated. First of all, there is the issue of corrections to Ψ_0 due to the next set of terms (of order e^4) in the solution of

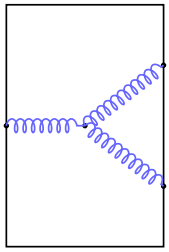


Figure 12: Example of a correction from $\mathcal{F}^{(3)}$ term to the Wilson loop expectation value. Each"gluon" stands for a $\langle J^a J^b \rangle$ propagator.

the Schrödinger equation, as in section 9. Secondly, even to the order we are calculating here, there are diagrams with two or more current loops in the effective 2d theory of the φ-fields. Finally, there could also be corrections which do not appear as loop corrections to the quadratic terms in Ψ_0 but have to be included in the computation of the expectation value of the Wilson loop operator. An example of such a diagram is shown in Fig. 12. These types of corrections can also be representation-dependent in general and they can be important for string-breaking effects as well.

This work was supported in part by the U.S. National Science Foundation Grant No. PHY-2112729.

References

[1] K. Symanzik, Nucl. Phys. **B 190** [FS3], 1 (1981). Some of the earlier references can be traced from here.

[2] M. Lüscher, Nucl. Phys. **B 254**, 52 (1985).

[3] R.P. Feynman, Nucl. Phys. **B 188**, 479 (1981).

[4] I.M. Singer, Physica Scripta **T24** (1981) 817. For a general formulation of the configuration space for gauge theories, see I.M. Singer, Commun. Math. Phys. **60**, 7(1978).

[5] V.P. Nair, *Liquid Helium and QCD(2+1): Feynman's last problem, 35 years later*, available at `https://www.researchgate.net/publication/312056227` DOI: 10.13140/RG.2.2.14570.72640

[6] D. Gross, R. Pisarski and L. Yaffe, Rev. Mod. Phys. **53**, 43 (1981) and references therein.

[7] R. Jackiw in *Field Theory and Particle Physics*, V Jorge Andre Swieca Summer School 1989, O.J.P. Eboli, M. Gomes and A. Santoro (eds.), World Scientific, 1990.

[8] M.F. Atiyah and R. Bott, Phil. Trans. R. Soc. Lond. **A 308**, 523 (1982).

[9] M. Bos and V. P. Nair, Int. J. Mod. Phys. A **5**, 959 (1990).

[10] A. Agarwal and V. P. Nair, Nucl. Phys. B **816**, 117 (2009) [arXiv:0807.2131 [hep-th]].

[11] K. Gawedzki and A. Kupiainen, Phys. Lett. B **215**, 119 (1988); Nucl. Phys. B **320**, 625 (1989).

[12] D. Karabali and V. P. Nair, Nucl. Phys. B **464**, 135 (1996) [arXiv:hep-th/9510157]; Phys. Lett. B **379**, 141 (1996) [hep-th/9602155].

[13] D. Karabali, C. j. Kim and V. P. Nair, Nucl. Phys. B **524**, 661 (1998) [arXiv:hep-th/9705087].

[14] A.M. Polyakov and P.B. Wiegmann, Phys. Lett. **141 B** (1984) 223; see also D.Gonzales and A.N.Redlich, Ann. Phys.(N.Y.) **169**, 104 (1986); B.M. Zupnik, Phys. Lett. **B183**, 175 (1987).

[15] V.N. Gribov, Nucl. Phys. **B 139**, 1 (1978).

[16] D. Karabali, C. j. Kim and V. P. Nair, Phys. Lett. B **434**, 103 (1998) [arXiv:hep-th/9804132].

[17] D. Karabali, V. P. Nair and A. Yelnikov, Nucl. Phys. B **824**, 387 (2010) [arXiv:0906.0783 [hep-th]].

[18] M. Teper, Phys. Rev. **D59**, 014512 (1999); B. Lucini and M. Teper, Phys. Rev. **D66**, 097502 (2002);

[19] H.B. Meyer and M.J. Teper, Nucl. Phys. **B668**, 111 (2003).

[20] B. Bringoltz and M. Teper, Phys. Lett. B **645**, 383 (2007) [arXiv:hep-th/0611286].

[21] N.D. Hari Dass and P. Majumdar, Phys. Lett. **B 658**, 273 (2008) [arXiv:hep-lat/0702019].

[22] J. Kiskis and R. Narayanan, JHEP 0809:080 (2008) [arXiv:0807.1315[hep-th]].

[23] B.H. Wellegehausen, A. Wipf and C. Wozar, Phys. Rev.**D 83**, 016001 (2011) [arXiv:1006.2305[hep-lat]].

[24] C. Bonati, M. Caselle and S. Morlacchi, Phys. Rev. **D 104**, 054501 (2012) [arXiv:2106.08784[hep-lat]].

[25] M.R. Douglas and S. Shenker, Nucl. Phys. **B 447**, 271 (1995) [arXiv:hep-th/9503163].

[26] N. Seiberg and E. Witten, Nucl. Phys. **B 426**, 19 (1994) [arXiv:hep-th/9407087]; Nucl. Phys. **B 431**, 484 (1994) [arXiv:hep-th/9408099].

[27] A. Hanany, M.J. Strassler and A. Zaffaroni, Nucl. Phys. **B 513**, 87 (1998) [arXiv:hep-th/9707244].

[28] C.P. Herzog and I.R. Klebanov, Phys. Lett. **B 526**, 388 (2002) [arXiv:hep-th/0111078].

[29] A. Armoni and M. Shifman, Nucl. Phys. **B 664**, 233 (2003) [arXiv:hep-th/0304127]; Nucl. Phys. **B 671**, 67 (2003) [arXiv:hep-th/0307020].

[30] M. Shifman, Acta. Phys. Polon. **B 36**, 3805 (2005) [arXiv:hep-th/0510098].

[31] P.N. Meisinger and M.C. Ogilvie, Phys. Rev. **D 81**, 025012 (2010) [arXiv:0905.3577[hep-lat]].

[32] C.P. Herzog, Phys. Rev. **D 66**, 065009 (2002) [arXiv:hep-th/0205064].

[33] L.A. Pando Zayas, V.G.J. Rodgers and K. Stiffler, JHEP 0812:036,2008 [arXiv:0809.4119[hep-th]].

[34] C. Doran, L.A. Pando Zayas, V.G.J. Rodgers and K. Stiffler, JHEP 0911:064,2009 [arXiv:0907.1331[hep-th]]

[35] J. Greensite, B. Lucini and A, Patella, Phys. Rev. **D 83** 125019 (2011) [arXiv:1101.5344[hep-th]].

[36] For the numerical lattice-based evaluation of the Casimir energy for nonabelian gauge theories, see M. N. Chernodub, V. A. Goy, A. V. Molochkov, Ha Huu Nguyen, Phys. Rev. Lett. **121**, 191601 (2018). For the analytical work on the Casimir energy and the analytic calculation of the propagator mass, see D. Karabali and V.P. Nair, Phys. Rev. **D 98**, 105009 (2018).

[37] R. Pasechnik and M. Sumbera, Universe **3**, 7 (2017) [arXiv:1611.01533[hep-ph]].

[38] V.P. Nair, Phys. Lett. **B 352**, 117 (1995); G. Alexanian and V.P. Nair, Phys. Lett. **B 352**, 435 (1995).

[39] R. Jackiw and S.Y. Pi, Phys. Lett. **B368**, 131 (1996); *ibid* **B403**, 297 (1997).

[40] W. Buchmüller and O. Philipsen, Nucl.Phys. B443 (1995) 47; W. Buchmüller and O. Philipsen, Phys. Lett. **B397**, 112 (1997).

[41] V.P. Nair, Rev. in Math. Phys. **33**, 214002 (2021) [arXiv:1910.06051[hep-th]].

[42] G.M. von Hippel and R.R. Horgan, Phys. Rev. Lett. **90**, 132001 (2003) [arXiv:hep-ph/0207203].

[43] F. Eberlein, Phys. Lett. **B439**, 130 (1998); Nucl. Phys. **B550**303 (1999); see also the talk by O. Philipsen at the *Workshop on QCD Green's Functions, Confinement and Phenomenology, QCD-TNT09*, September 7-11, 2009, Trento, Italy; published in the *Proceedings of Science*, PoS (QCD-TNT09) 052, `http://pos.sissa.it//archive/conferences/087/030/QCD-TNT09_052.pdf`

[44] D. Bieletzki, K. Lessmeier, O. Philipsen and Y. Schroder, JHEP 2012, 58 (2012) [arXiv:1203.6538[hep-ph]].

[45] A.C. Aguilar, D. Binosi and J. Papavassiliou, Phys. Rev.**D 81**, 125025 (2010) and references therein.

[46] The literature on this is enormous, but a good status report with references to earlier work is J.M. Cornwall, J. Papavassiliou and D. Binosi, *The Pinch Technique and its Application to Nonabelian Gauge Theories*, Cambridge University Press (2011). A complementary review is N. Vandersickel and D. Zwanziger, [arXiv:1202.1491[hep-th]].

[47] A more recent review article which also updates the status of this approach is A.C. Aguilar, D. Binosi and J. Papavassiliou, Front. Phys. **11**(2), 111203 (2016).

[48] The possibility of calculating the magnetic mass by summing Feynman diagrams was addressed in J.M. Cornwall, Phys. Rev. **D26**, 1453 (1982). The paper J.M. Cornwall, W-S. Hou, J.E. King, Phys. Lett. **B153**, 173 (1985) gave the bound $m/e^2 \geq 0.58$; the later paper J.M. Cornwall, Phys. Rev. **D57**, 3694 (1998) gave a modified procedure with the estimate $0.248 \approx 0.25$, as quoted. See also J.M. Cornwall, Phys. Rev. **D10**, 500 (1974); *ibid.* **D57**, 3694 (1998); J. M. Cornwall and B. Yan, Phys. Rev. **D 53**, 4638 (1996); J.M. Cornwall, Phys. Rev. **D57**, 3694 (1998); J.M. Cornwall, Phys. Rev. **D76**, 025012 (2007).

[49] F. Karsch *et al*, Nucl. Phys. **B474**, 217 (1996); F. Karsch, M. Oevers and P. Petreczky, Phys. Lett. **B442**, 291 (1998).

[50] U. M. Heller, F. Karsch and J. Rank, Phys. Rev. **D 57**, 1438 (1998) [arXiv:hep-lat/9710033].

[51] O. Philipsen, Phys. Lett. **B521**, 273 (2001); talk at *Lattice 2001*, [arXiv:hep-lat/0110114].

[52] O. Philipsen in *Strong and Electroweak Matter 2002*, Proceedings of the SEWM2002 Meeting, Heidelberg, Germany, October 2002, M.G. Schmidt (ed), https://doi.org/10.1142/5243 — June 2003

[53] A. Nakamura, I. Pushkina, T. Saito and S. Sakai, Phys. Lett. **B 549**, 133 (2002) [arXiv:0208075[hep-lat]].

[54] A. Nakamura, T. Saito and S. Sakai, Phys. Rev. **D 69**, 014506 (2004) [arXiv:0311024[hep-lat]].

[55] A. Agarwal, D. Karabali and V. P. Nair, Nucl. Phys. **B 790**, 216 (2008) [arXiv:0705.0394 [hep-th]].

[56] O. Philipsen and H. Wittig, Phys. Lett. **B 451**, 146 (1999) [arXiv:hep-lat/9902003].

[57] S. Kratochvila and P. de Forcrand, Nucl. Phys. **B 671**, 103 (2003) [arXiv:hep-lat/0306011].

[58] F. Knechtli and R. Sommer, Nucl. Phys. **B 590**, 309 (2000) [arXiv:hep-lat/0005021].

[59] M. Pepe and U.-J. Wiese, Phys. Rev. Lett. **102**, 191601 (2009) [arXiv:0901.2510 [hep-lat]].

[60] H.D. Trottier and K. Y. Wong, Phys. Rev. **D 72**, 054505 (2005) [arXiv:0408028[hep-lat]].

[61] C. Bonati and S. Morlacchi, Phys. Rev. **D 101**, 094506 (2020) [arXiv:2003.07244[hep-lat]].

[62] R. G. Leigh, D. Minic and A. Yelnikov, Phys. Rev. Lett. **96**, 222001 (2006); Phys. Rev. **D76**, 065018 (2007).

[63] J. Greensite, Nucl. Phys. **B 158**, 469 (1979).

[64] S. Samuel, Phys. Rev. **D 55**, 4189 (1997) [arXiv:hep-ph/9604405]

[65] J. Greensite and S. Olejnik, Phys. Rev. **D 77**, 065003 (2008) [arXiv:0707.2860[hep-lat]].

[66] J. Greensite *et al*, Phys. Rev. **D 83**, 114509 (2011) [arXiv:1102.3941[hep-lat]].

[67] R. Jackiw and S. Templeton, Phys. Rev. **D 23**, 2291 (1981); S. Deser, R. Jackiw and S. Templeton, Phys. Rev. Lett. **48**, 975 (1982), S. Deser, R. Jackiw and S. Templeton, Annals Phys. **140**, 372 (1982)

[68] J. Schonfeld, Nucl. Phys. .**B 185**, 157 (1981).

[69] M. Asorey and P.K. Mitter, Phys. Lett. **B 153**, 147 (1985).

[70] D. Karabali, C. -j. Kim and V. P. Nair, Nucl. Phys. B **566**, 331 (2000) [hep-th/9907078]. In this connection, see also the paper G. Grignani, G. Semenoff, P. Sodano and O. Tirkkonen, Nucl. Phys. **B489** (1997) 360.

[71] M. Asorey, J. Geom. Phys. **11**, 63 (1993).

[72] A. Agarwal and V.P. Nair, Phys. Rev **D85**, 085011 (2012) [arXiv:1201.6609[hep-th]].

[73] A. Agarwal and V.P. Nair, Jour. Phys. A: Math and Theor. **48**, 465401 (2015) [arXiv:1504.07201[hep-th]].

[74] E. Witten, Commun. Math. Phys. **117**, 353 (1988).

[75] V.G. Knizhnik and A.B. Zamolodchikov, Nucl. Phys. **B 247**, 83 (1984).

[76] R. D. Pisarski and S. Rao, Phys. Rev. D **32**, 2081 (1985).

[77] H. -C. Kao, K. -M. Lee and T. Lee, Phys. Lett. B **373**, 94 (1996) [hep-th/9506170].

[78] N. Seiberg and E. Witten, In *Saclay 1996, The mathematical beauty of physics* 333-366 [hep-th/9607163].

[79] J. W. Elliott and G. D. Moore, JHEP **0711**, 067 (2007) [arXiv:0708.3214 [hep-lat]].

[80] E. Witten, In *Shifman, M.A. (ed.): The many faces of the superworld* 156-184 [hep-th/9903005].

[81] I. Affleck, J. A. Harvey and E. Witten, Nucl. Phys. B **206**, 413 (1982), J. de Boer, K. Hori and Y. Oz, Nucl. Phys. B **500**, 163 (1997) [hep-th/9703100], O. Aharony, A. Hanany, K. A. Intriligator, N. Seiberg and M. J. Strassler, Nucl. Phys. B **499**, 67 (1997) [hep-th/9703110].

[82] J. Gomis and J. G. Russo, JHEP **0110**, 028 (2001) [hep-th/0109177].

[83] M. Unsal, Phys. Rev. D **80**, 065001 (2009) [arXiv:0709.3269 [hep-th]].

[84] D. N. Kabat, Nucl. Phys. B **453**, 281 (1995) [hep-th/9503016].

[85] D. Burghelea, L. Friedlander and T. Kappeler, J. of Funct. Anal. **107**, 34 (1992).

[86] A. Agarwal, D. Karabali and V.P. Nair, Phys. Rev. **D 96**, 125008 (2017) [arXiv:1701.00014[hep-th]]

[87] D. Karabali and V.P. Nair, Phys. Rev. **D 77**, 025014 (2008) [arXiv:0705.2898[hep-th]].

[88] P. Mansfield, JHEP 04 (2004) 059 [arXiv:hep-th/0406237].

[89] S. Krug and A. Pineda, Phys. Rev. **D 88**, 125001 (2013) [arXiv:1301.6992[hep-th]]; Nucl. Phys. **B 878**, 82 (2014) [arXiv:1308.2663[hep-th]];
S. Krug, arXiv:1404.7005[hep-th].