SOME ASPECTS OF THE QUANTIZATION OF THEORIES WITH A GAUGE INVARIANCE

Thesis by
George Siopsis

In Partial Fulfillment of the Requirements
for the Degree of
Doctor of Philosophy

California Institute of Technology
Pasadena, California.

1987
(Submitted May 15, 1987)

ACKNOWLEDGEMENTS

First of all, I wish to express my gratitude to Prof. R. P. Feynman for providing the initial stimulation and guidance. Special thanks go to Prof. H. D. Politzer for an enjoyable, yet not so fruitful collaboration. But above all, I am indebted to Prof. J. P. Preskill for the numerous enlightening discussions we had, his constant encouragement, and his multifaceted support.

I also wish to thank all my friends, without whom my stay at Caltech would have been far less stimulating and enjoyable. I shall mention them by name, despite the fact that I am running the risk of omitting some: Alexis (my office-mate), Brian, Christos, Elias (my collaborator for part of this work), Hidenori, Milan, Rolf, Shahram, Ted, and, most important of all, Diane, whose understanding, patient enduring of my periods of depression, and her imaginative contributions to the moments of joy have been invaluable.

ABSTRACT

We discuss some problems that arise when one tries to quantize a theory that possesses gauge degrees of freedom. First, we identify the Gribov problem that is encountered when the Faddeev-Popov procedure of fixing the gauge is employed to define a perturbation expansion. We propose a modification of the procedure that takes this problem into account. We then apply this method to two-dimensional gauge theories where the exact answer is known. Second, we try to build chiral theories that are consistent in the presence of anomalies, without making use of additional degrees of freedom. We are able to solve the model exactly in two dimensions, arriving at a gauge-invariant theory. We discuss the four-dimensional case and also the application of this method to string theory. In the latter, we obtain a model that lives in arbitrary dimensions. However, we do not compute the spectrum of the model. Third, we investigate the possibility of compactifying the unwanted dimensions of superstings on a group manifold. We give a complete list of conformally invariant models. We also discuss one-loop modular invariance. We consider both type-II and heterotic superstring theories. Fourth, we discuss quantization of string field theory. We start by presenting the lagrangian approach, to demonstrate the non-uniqueness of the measure in the path-integral. It is fixed by demanding unitarity, which manifests itself in the hamiltonian formulation, studied next.

Την εργασία αυτή αφιερώνω
στους γονείς μου,
και στ' αδέλφια μου,
Αναστασία και Σταύρο.

Dedicated to my parents,
my sister Anastasia,
and my brother Stavros.

Table of Contents

1. Introduction

Ο άνθρωπος φύσει του ειδέναι ορέγεται.

Αριστοτέλης

Man, by nature, yearns to know.

Aristotle

The first quantum theory with a gauge symmetry was quantum electrodynamics (QED), based on the simplest gauge group, $U(1)$. Its beauty lies in the fact that interactions are mediated by the photon, whose introduction is necessary so that the theory possesses a local symmetry. The profound successes of QED in explaining experimental data led people to propose [1] that an extension of the idea of "gauge symmetry" to more complicated Lie groups could explain other interactions in nature.

Thus, quantum chromodynamics (QCD) was developed, based on the group $SU(3)$, that was designed to explain the strong interactions. Being a non-abelian theory, however, meant that extraction of numbers that could be compared with experimental data would be much more cumbersome. In fact, no such numbers have yet been calculated. The main obstacle is the strong coupling that does not allow the straightforward use of perturbation theory.

The gauge principle was more successful with weak interactions, where it was combined with the idea of "symmetry breaking" through Higgs particles.† However, the attempt to extend these ideas to build a theory of all known forces except gravity (grand unified theories) [3] has for all intents and purposes failed.

When QCD was at its infancy, another approach seemed promising for strong interaction physics: dual resonance models [4]. It was soon discovered that they were

† For a review and references, see ref.[2].

equivalent to postulating that elementary particles be the modes of vibrating strings [4], and therefore have a certain intrinsic extent. Although computationally much simpler than QCD, this theory had serious shortcomings: it contained a particle with imaginary mass (tachyon) and could only exist in the unworldly 26 dimensions [5]. Strings were set aside by QCD after the discovery of asymptotic freedom [6].

Another crucial development in gauge theories was the discovery of "anomalies." This term was invented to describe the breakdown of classical symmetries due to quantum effects. They were first discovered by Steinberger [7] but they started to attract attention much later [8]. It was then realized [9] that they imposed strong constraints on the particle content of gauge theories. Anomalies also resurrected string theories, which were no longer viewed as theories of strong interactions, but as theories of everything [10]. They were combined with a gauge principle and anomalies led to an almost unique choice by restricting the gauge groups [11]. However, the task of extracting any numbers became formidable, because string theories were computationally even more complex than QCD.

More recently, anomalies were investigated from a different point of view: not as a plague, but as a challenge to build a consistent theory in their presence [12]. They are also very useful in that they constitute a probe to the structure of the space of gauge potentials, which is largely unexplored. Another probe is the study of the Gribov ambiguity [13], which is also a consequence of the topology of gauge space. These two problems (the Gribov ambiguity and the construction of consistent anomalous theories) are discussed in the next two chapters. The general motivation for this work is the belief that a better understanding of gauge space is vital for the development of QCD or any other theory with a gauge freedom, that will include more interactions.

It is interesting to note that if any program of circumventing anomalies succeeds, it will strip string theories, that are meant to describe all interactions, of their beauty: their uniqueness. However, this by no means implies that the study of string theories is useless. On the contrary, they possess a rich mathematical structure, whose study will certainly lead to a better understanding of field theory. We believe that if they fail to explain everything, strings will still succeed in a more modest, yet extremely important task: to help extract numbers that will explain some experimental data in strong interactions, which was the original motivation for investigating them. However, a lot more work is needed before this goal is achieved.

Two aspects of string theories are discussed here. One is the compactification of the extra dimensions that are necessary for the consistency of the theory (Chapter 4). The other is the quantization of string field theory (Chapter 5).

In more detail, the organization of this work is as follows. In Chapter 2, we discuss the Gribov ambiguity. First, we identify the two problems that arise in the Faddeev-Popov procedure of fixing the gauge. We also exhibit their topological origin. We then proceed to solve these problems. We give a heuristic solution for the first problem. For the second problem, we derive an expression for the path integral that takes it into account. We find that this problem has an effect in two-dimensional perturbation theory. We show that the naïve expansion disagrees with the exact result, whereas correct results are obtained when the second Gribov problem is properly accounted for. It is possible to follow the same steps in four dimensions, however no surprises are expected there. It would also be interesting to find the equivalent modification of perturbation theory for lattice gauge theories, where the perturbation expansion is also in error. However, we have not succeeded in formulating the Gribov problem there.

Chapter 3 is devoted to the subject of anomalies. We introduce two regularization procedures that lead to the consistent and the covariant forms of the anomaly, respectively. It should be noted that this result is not in contradiction with the fact that the covariant anomaly is not related to the consistent one by the addition of counterterms in the effective action, because, as we show, the two models we obtain satisfy different Dyson-Schwinger equations. We also introduce a gauge-invariant regulator, which we apply to three different cases: (a) two-dimensional chiral Schwinger model, (b) four-dimensional chiral gauge theories, (c) conformal anomalies in bosonic strings. In case (a), we solve the model exactly, thus showing explicitly that it is possible to construct gauge-invariant anomalous theories. In case (b), we limit ourselves to computing the divergence of the gauge current and the commutator of generators of gauge transformations. The results indicate that the model is gauge-invariant. However, more work is needed in order to prove the unitarity of the theory. In case (c), we obtain a consistent theory in arbitrary space-time dimensions. However, we did not succeed in computing the spectrum of the model. It should be noted that in all cases we do not make use of extra degrees of freedom.

Chapter 4 deals with compactification of strings on group manifolds.‡ The advantage of group manifolds over other means of compactification (e.g., Calabi-Yau spaces) lies in the fact that the theory is exactly solvable in the former case. We study all Lie groups and identify the spectrum in each case. We consider both type-II and heterotic string models. Some models contain fermionic multiplets that realize the supersymmetry independently. When the model contains bosons that cannot be fermionized, we show that it is necessary to twist their boundary conditions, in order to obtain a realistic theory. We discuss the complications that arise concerning the proof of modular invariance in this case.

Finally, in Chapter 5, we discuss the quantization of string field theory. We first present the lagrangian formulation following the program proposed by Fradkin et al. [14]. Our starting point is the action proposed by Witten [15]. Unfortunately, we do not obtain a unique measure for the path integral. It is fixed by requiring unitarity of the quantum theory, which can be achieved in the hamiltonian formalism. We therefore define the path integral using the hamiltonian formalism and do an explicit calculation to demonstrate agreement with the lagrangian formalism at tree level. It will be interesting to calculate a loop diagram and see whether closed string poles appear in open strings.

‡ This work was done in collaboration with E. B. Kiritsis.

References

1. C. N. Yang and R. Mills, *Phys. Rev.* **96** (1954) 191.
2. E. S. Abers and B. W. Lee, *Phys. Rep.* **9** (1973) 1.
3. P. Langacker, *Phys. Rep.* **72** (1981) 85.
4. Y. Nambu, Lectures at the Copenhagen Symposium (1970);
   L. Susskind, *Nuovo Cim.* **69A** (1970) 457.
5. C. Lovelace, *Phys. Lett.* **34B** (1971) 500.
6. H. D. Politzer, *Phys. Rev. Lett.* **26** (1973) 1346;
   D. Gross and F. Wilczek, *Phys. Rev. Lett.* **26** (1973) 1343.
7. J. Steinberger, *Phys. Rev.* **76** (1949) 1180.
8. S. L. Adler, *Phys. Rev.* **177** (1969) 2426;
   W. A. Bardeen, *Phys. Rev.* **184** (1969) 1848;
   J. S. Bell and R. Jackiw, *Nuovo Cim.* **60A** (1969) 47.
9. D. J. Gross and R. Jackiw, *Phys. Rev.* **D6** (1972) 477;
   C. Bouchiat, J. Iliopoulos and Ph. Meyer, *Phys. Lett.* **38B** (1972) 519.
10. J. Scherk and J. H. Schwarz, *Nucl. Phys.* **B81** (1974) 118.
11. M. B. Green and J. H. Schwarz, *Phys. Lett.* **149B** (1984) 117.
12. R. Jackiw and R. Rajaraman, *Phys. Rev. Lett.* **54** (1985) 1219;
    L. D. Faddeev and Shatashvili, *Phys. Lett.* **167B** (1986) 225;
    A. J. Niemi and G. W. Semenoff, *Phys. Rev. Lett.* **56** (1986) 1019.
13. V. N. Gribov, *Nucl. Phys.* **B139** (1978) 1.
14. E. S. Fradkin and G. A. Vilkovisky, *Phys. Lett.* **55B** (1975) 224;
    I. A. Batalin and G. A. Vilkovisky, *Phys. Lett.* **102B** (1981) 27.
15. E. Witten, *Nucl. Phys.* **B268** (1986) 253.

2. The Gribov ambiguity

2.1 The Faddeev-Popov procedure

An indispensable tool in the quantization of gauge theories is the Faddeev-Popov procedure [1]. We shall briefly illustrate it in the case of Yang-Mills theories. For definiteness, we shall be working in the adjoint representation of $SU(N)$.

The lagrangian density for a pure Yang-Mills theory is

$$\mathcal{L} = -\frac{1}{4g^2} F^a_{\mu\nu} F^a_{\mu\nu} \quad , \tag{2.1.1}$$

where $F^a_{\mu\nu} = \partial_\mu A^a_\nu - \partial_\nu A^a_\mu + i f^{abc} A^b_\mu A^c_\nu$ is the field strength, f^{abc} are the structure constants of the gauge group $G = SU(N)$ ($[T^a, T^b] = f^{abc} T^c$, where T^a are the generators of the Lie algebra of G, normalized by $\mathrm{tr} T^a T^b = \frac{1}{2}\delta^{ab}$), and A^a_μ is the vector potential ($\mu = 0, ..., 3$). The momenta are defined by

$$P^a_\mu \equiv \frac{\delta \mathcal{L}}{\delta(\partial_0 A^a_\mu)} = F^a_{0\mu} \quad . \tag{2.1.2}$$

It is clear that $P^a_0 = 0$, so the transformation to the hamiltonian picture is singular. The hamiltonian density is

$$\mathcal{H} \equiv \frac{1}{g^2} P^a_i \partial_0 A^a_i - \mathcal{L} = \frac{1}{2g^2}(P^a_i P^a_i + B^a_i B^a_i) - \frac{1}{g^2} A^a_0 D^{ab}_i P^b_i \quad , \tag{2.1.3}$$

($i = 1, 2, 3$) where $B_i = \frac{1}{2}\epsilon_{ijk} F^a_{jk}$ is the "magnetic" field, and $D^{ab}_\mu = \partial_\mu \delta^{ab} - i f^{abc} A^c_\mu$. The time component of the vector potential, A^a_0 plays the rôle of a Lagrange multiplier enforcing the constraint (Gauss's law)

$$D^{ab}_i P^b_i = 0 \quad . \tag{2.1.4}$$

This is also the generator of time-independent gauge transformations, $\delta A^a_i = -D^{ab}_i \alpha^b$, $\delta P^a_i = -i f^{abc} P^b_i \alpha^c$ and $\delta A^a_0 = 0$, which leave the hamiltonian invariant.

A quantum theory is defined through the generating functional

$$e^{iW[J]} = \int \mathcal{D}A\mathcal{D}P \exp\left\{ i \int \frac{1}{g^2} P_i^a \partial_0 A_i^a - \mathcal{H} + J_\mu^a A_\mu^a \right\} \quad , \tag{2.1.5}$$

where J_μ^a is an external current that the system is coupled to. We do not introduce dynamical fermions, as they add nothing to our discussion.

Next, we fix the gauge using the Faddeev-Popov procedure. For simplicity, we choose a linear gauge, $MA = 0$, where M is a linear operator independent of A_μ^a and P_i^a. This class of gauges includes the Coulomb gauge ($\partial_i A_i^a = 0$) and the axial gauge ($A_3^a = 0$). Inserting

$$1 = \Delta[A_i^a] \int \mathcal{D}\alpha \delta[M_{ij}^{ab}(A_j^b + D_j^{bc}\alpha^c)] \tag{2.1.6}$$

in the path-integral, performing a gauge transformation and integrating over A_0^a, we obtain

$$e^{iW[J]} = \int \mathcal{D}A\mathcal{D}P \Delta[A] \delta[MA] \delta[D_i^{ab} P_i^b + g^2 J_0^a] e^{i\int \frac{1}{g^2} P_i^a \partial_0 A_i^a - \mathcal{H}_0} \quad , \tag{2.1.7}$$

where $\mathcal{H}_0 = \frac{1}{2g^2} P_i^a P_i^a + \frac{1}{2g^2} B_i^a B_i^a$. Δ is the Faddeev-Popov determinant which is easily computed using eq.(2.1.6):

$$\Delta[A] = \det MD \quad . \tag{2.1.8}$$

Now let $P = \tilde{P} + M\phi$ and $A = \tilde{A} + A'$, where $M\tilde{P} = M\tilde{A} = 0$. The gauge-fixing condition $MA = 0$ becomes $A' = 0$. Thus, Gauss's law becomes an equation in ϕ:

$$\partial_i \partial_i \phi^a + i f^{abc} \tilde{A}_i^b \partial_i \phi^c = \rho^a \quad , \tag{2.1.9}$$

where $\rho^a = g^2 J_0^a - i f^{abc} \tilde{A}_i^b \tilde{P}_i^c$ is the total "color" density ($g^2 J_0^a$ and $i f^{abc} \tilde{A}_i^b \tilde{P}_i^c$ can be thought of as the "quark" and "gluon" contributions, respectively). Integration over

ϕ gives a factor of $\det(MD)^{-1} = \Delta^{-1}$, which cancels the Faddeev-Popov determinant (eq.(2.1.8)). The final result is

$$e^{iW[J]} = \int \mathcal{D}\tilde{A}\mathcal{D}\tilde{P}e^{i\int \frac{1}{g^2}\tilde{P}_i^a\partial_0\tilde{A}_i^a - \mathcal{H}_0 + A_i^a J_i^a} \quad , \tag{2.1.10}$$

where $\mathcal{H}_0 = \frac{1}{2g^2}(\tilde{P}^2 + (M\phi)^2 + B^2)$, and ϕ is given by eq.(2.1.9). Thus, the generating functional has been expressed in terms of only physical degrees of freedom. This serves as the starting point for perturbative calculations.

It should be noted that the whole analysis can be carried out at the classical level. The path-integral is then defined in the standard way, once the independent degrees of freedom have been identified.

2.2 The problem(s)

The procedure that we described in the previous section only works under certain assumptions that we shall now discuss. For $g(A)$ to be a good gauge-fixing condition, it has to intersect each gauge orbit exactly once. If there are two or more points of intersection,† then we count gauge equivalent potentials twice or more times, respectively. These extra potentials are called Gribov copies, after Gribov who discovered them [2], and seem to be commonplace in Yang-Mills theories.

Singer showed that if the boundary conditions at infinity amount to working in a space that has the topology of a sphere,‡ then it is impossible to choose a gauge so as to avoid Gribov copies [3]. The argument goes as follows.

Suppose that there exists a gauge-fixing surface $\Sigma = \{A^\mu : g(A) = 0\}$ that crosses each orbit exactly once. Then, Σ, together with a fixed orbit, define a good coordinate system in the space of all vector potentials, $\mathcal{A}$. Indeed, let $A^\mu \in \mathcal{A}$ be an arbitrary vector potential. There exists a unique potential $\bar{A}^\mu \in \Sigma$ that is gauge equivalent to

---

† We do not discuss the case of zero points of intersection, because it does not arise.

‡ This is normally the case, because one considers only gauge transformations that become the identity at spatial infinity.

A^μ:

$$A^\mu = \Lambda^{-1}\bar{A}^\mu\Lambda + i\Lambda^{-1}\partial^\mu\Lambda \quad , \tag{2.2.1}$$

for some $\Lambda \in \mathcal{G}$, where $\mathcal{G}$ is the group of all three-dimensional maps from S^3 (the three-dimensional sphere) to the gauge group G. It is clear that we can associate a pair $(\bar{A}^\mu, \Lambda)$ with each potential A^μ in a unique fashion, where $\bar{A}^\mu \in \Sigma$ and $\Lambda \in \mathcal{G}$. Since each orbit is a copy of $\mathcal{G}$, it follows that the space of all potentials is the direct product of Σ and $\mathcal{G}$:

$$\mathcal{A} = \Sigma \times \mathcal{G} \quad . \tag{2.2.2}$$

Now, $\mathcal{A}$ is topologically trivial, therefore so is $\mathcal{G}$. However, this is not true, because for any Lie group G, there exists at least one homotopy group $\pi_j(\mathcal{G}) = \pi_{j+3}(G)$ that is not trivial [3]. Thus, we reached a contradiction, which shows that a continuous surface Σ crossing each orbit only once does not exist. Hence, Gribov copies are inevitable. A (heuristic) solution to this problem will be presented in the next section.

As an illustration, consider the space $\mathcal{A} = \mathbf{R}^3$, in which the orbits are the paraboloids $\Sigma_c = \{(x, y, z) : z = x^2 + y^2 + c\}$. Each Σ_c is topologically trivial, since $\pi_j(\Sigma_c)$ is trivial, $\forall j$. Therefore, we expect to be able to find a continuous line that crosses each orbit exactly once. Indeed, such a line exists; it is the z-axis.

Now suppose that the orbits are coaxial cylinders: $\Sigma_c = \{(x, y, z) : x^2 + y^2 = c\}$. In this case $\pi_1(\Sigma_c)$ is not trivial and one cannot find a line that crosses each orbit only once. The same is true if the orbits are concentric spheres: $\{\Sigma_c = \{(x, y, z) : x^2 + y^2 + z^2 = c\}$, where $\pi_2(\Sigma_c)$ is not trivial.

There is also the degenerate case in which the gauge-fixing surface $g(A) = 0$ is tangent to an orbit. This is true for the Coulomb gauge,[§]

$$g(A) \equiv \partial_i A_i^a(x) = 0 \quad , \tag{2.2.3}$$

as was first shown by Gribov[♮] [2]. In this case, the Faddeev-Popov determinant (eq.(2.1.8)) vanishes. Consequently, the physical degrees of freedom are ambiguously defined, because, as can be seen from eq.(2.1.9), the operator MD has zero modes and is therefore not invertible. This would not have been an obstacle, if the vector potentials for which MD has zero modes formed a set of measure zero. However, it has been shown [4] that this is not the case.

To show that the Faddeev-Popov determinant vanishes, it suffices to show that the operator $MD \equiv \partial_i \partial_i \delta^{ab} - i f^{abc} A_i^c \partial_i$ has a zero mode. Thus, we have to find a $\phi^a(x)$ satisfying

$$\partial_i \partial_i \phi^a + i f^{abc} A_i^b \partial_i \phi^c = 0 \ . \tag{2.2.4}$$

Consider the equation

$$\partial_i \partial_i \phi^a + i f^{abc} A_i^b \partial_i \phi^c = E \phi^a \ . \tag{2.2.5}$$

This is like a Schrödinger equation with an attractive potential [3]. Therefore, for large enough A_i there exist bound states, i.e., solutions of (2.2.5) with $E < 0$. It follows that for intermediate magnitudes of A_i, there exist zero-energy solutions, i.e., solutions of eq.(2.2.4). Such solutions were first discovered by Gribov [2] and have been further investigated [5].

---

§ Another way of defining the Coulomb gauge is as the gauge minimizing the expression

$$\int d^3x A_i^a A_i^a \ .$$

♮ Singer [3] has given a proof for a generalized version of the Coulomb gauge, in which the surface $g(A)$ is orthogonal to the orbit at some fixed potential $A_\mu^{(0)}$. (For the ordinary Coulomb gauge $A^{(0)} = 0$.) The proof follows the same lines.

In section 2.4, we shall show that a careful implementation of the Faddeev-Popov procedure leads to an expression free of this second Gribov problem [6].

2.3 Solution to the first problem

We shall argue that the first problem is not a real handicap of the Faddeev-Popov procedure. For clarity, we shall make use of a simple illustrative example. Consider the integral

$$Z = \int dxdy e^{-S(x-y^{1/3})} \quad , \tag{2.3.1}$$

where the number of variables has been reduced to two! Of course, Z diverges, but if one is interested in expressions of the form

$$\langle X \rangle = \frac{1}{Z} \int dxdy X(x - y^{1/3}) e^{-S(x-y^{1/3})} \quad , \tag{2.3.2}$$

it makes sense to interpret Z as $Z = \int due^{-S(u)}$, because the (infinite) factor $\int dy$ cancels between the numerator and the denominator. The "action" S is invariant under the transformation $\delta x = \epsilon$, $\delta y = 3y^{2/3}\epsilon$. This transformation generates a flow along the lines (orbits) $x - y^{1/3}$ $=$*const*. So, fixing the gauge means choosing a line that intersects each orbit exactly once. Let us impose the gauge $y = a$. This is implemented by introducing the factor

$$1 = \Delta_1(y) \int d\epsilon \delta(y + 2y^{2/3}\epsilon - a) \quad , \tag{2.3.3}$$

inside the integral (eq.(2.3.1)). Changing the order of integration and performing the appropriate "gauge" transformation, we obtain

$$\begin{aligned} Z &\sim \int dxdy \delta(y-a) \Delta_1(y) e^{-S(x-y^{1/3})} \\ &\sim \int dx e^{-S(x-a^{1/3})} \\ &\sim \int du e^{-S(u)} \quad , \end{aligned} \tag{2.3.4}$$

where we have ignored overall normalization constants. Thus, we obtain the desired result.

Now, consider the gauge-fixing line $y = x$, which crosses some of the orbits more than once. Defining $\Delta_2(x, y)$ by

$$1 = \Delta_2(x,y) \int d\epsilon \delta(y + 2y^{2/3}\epsilon - x - \epsilon) \quad , \tag{2.3.5}$$

and arguing as before, we obtain

$$Z \sim \int_{-\infty}^{\infty} dx |2x^{2/3} - 1| e^{-S(x - x^{1/3})} \quad . \tag{2.3.6}$$

To change variables to $u = x - x^{1/3}$, we split the integral in three parts. Thus, we write

$$Z \sim \int_{-\infty}^{-\lambda} dx f(x) - \int_{-\lambda}^{+\lambda} dx f(x) + \int_{+\lambda}^{+\infty} dx f(x) \quad , \tag{2.3.7}$$

where $f(x) = (1 - 2x^{2/3})e^{-S(x-x^{1/3})}$ and $\lambda = 3^{-3/2}$. In terms of u, eq.(2.3.7) becomes

$$\begin{aligned} Z &\sim \int_{-\infty}^{+\sigma} du \; e^{-S(u)} - \int_{-\sigma}^{+\sigma} du \; e^{-S(u)} + \int_{-\sigma}^{+\infty} du \; e^{-S(u)} \\ &= \int_{-\infty}^{+\infty} du \; e^{-S(u)} \quad , \end{aligned} \tag{2.3.8}$$

where $\sigma = 3^{-1/2} - 3^{-3/2}$. Therefore, we obtain the right result. The reason is that the "determinant" Δ_2 is the absolute value of an expression ($\Delta_2 = |2x^{2/3} - 1|$), so the integrals between points of intersection cancel. Were $\Delta_2 = 2x^{2/3} - 1$, instead, the middle integral in eq.(2.3.8) would have had a plus sign leading to an incorrect expression.

Thus, the procedure works even if there are more than one points where the gauge-fixing line intersects the gauge orbits. Yang-Mills theories work in a similar way.

From the example we just discussed, we can also see how the second problem may arise. Indeed, suppose we choose the gauge-fixing line $y = 0$, which is tangent to all orbits. Then the resulting expression (eq.(2.3.4) with $a = 0$) is meaningless, owing to the fact that $\Delta_1(0) = 0$. Thus, the second problem invalidates the Faddeev-Popov procedure, and therefore requires more care.

2.4 Solution to the second problem

The second Gribov problem arises whenever the operator MD (with indices suppressed) has zero modes. Let us call $\mathcal{K}$ the space of all the zero modes of this operator. Whenever α^a is in $\mathcal{K}$, $MD\alpha = 0$ and so in the expression for the Faddeev-Popov determinant,

$$\Delta^{-1} = \int \mathcal{D}\alpha \delta[MD\alpha] \quad , \tag{2.4.1}$$

which is valid for the vector potentials lying on the gauge-fixing surface $MA = 0$, we get a $\delta(0)$ factor. It is these factors that produce the unwanted infinity. To eliminate them, we replace Δ^{-1} by [6]

$$\tilde{\Delta}^{-1} = \int \mathcal{D}\alpha \int_{\Re} \mathcal{D}\lambda \exp\left\{ i \int d^4x \lambda^a [M^{ab}_{ij} A^b_j + (MD)^{ab} \alpha^b] \right\} \quad , \tag{2.4.2}$$

where the integration over λ^a is restricted in the space $\Re$ of the functions orthogonal to the zero modes of MD. Inserting the factor

$$1 = \tilde{\Delta} \int \mathcal{D}\alpha \int_{\Re} \mathcal{D}\lambda \exp\left\{ i \int d^4x \lambda^a [M^{ab}_{ij} A^b_j + (MD)^{ab} \alpha^b] \right\} \quad , \tag{2.4.3}$$

in expression (2.1.5) for the path-integral, we interchange the order of integration and make an infinitesimal gauge transformation to obtain

$$e^{iW[J]} = \int \mathcal{D}P_i \mathcal{D}A_i \mathcal{D}A_0 \tilde{\Delta} \int_{\Re} \mathcal{D}\lambda e^{i \int \lambda^a M^{ab}_{ij} A^b_j} e^{iS} \quad , \tag{2.4.4}$$

where the (infinite) constant $\int \mathcal{D}\alpha$ has been erased.

Integration over λ produces a δ-function that says that the gauge-fixing condition $MA = 0$ has to be imposed, as the range of integration for λ has been restricted appropriately so that this information has not been lost. To clarify this statement, we cite a simple N-dimensional example. Let M be the $N \times N$ matrix

$$M = \mathrm{diag}(m_1, \ldots, m_k, 0, \ldots, 0) \quad , \tag{2.4.5}$$

with $N - k$ zeros along the diagonal. Consider the expression

$$D = \int d^N x e^{i\mathbf{x}^T(M\mathbf{y}-\mathbf{h})} \quad , \tag{2.4.6}$$

where $\mathbf{x}$, $\mathbf{y}$, $\mathbf{h}$ are N-vectors and $\mathbf{h}$ is orthogonal to the zero modes of M. If the integration is over all directions, then we get a string of $N - k$ $\delta(0)$'s under the integral sign. To avoid them, we restrict the integration over those $\mathbf{x}$'s that are of the form

$$\mathbf{x} = (x_1, \ldots, x_k, 0, \ldots, 0)^T \quad , \tag{2.4.7}$$

($N - k$ zeros), which form the range of M. Thus, we replace D by

$$\tilde{D} = \int dx_1 \cdots dx_k e^{i\mathbf{x}^T(M\mathbf{y}-\mathbf{h})} \quad , \tag{2.4.8}$$

with $\mathbf{x}$ given by eq.(2.4.7). A simple manipulation yields

$$\tilde{D} = \delta(m_1 y_1 - h_1) \cdots \delta(m_k y_k - h_k) \quad . \tag{2.4.9}$$

Hence, by doing this, we have retained all the "meaningful" δ's, having only dropped the $\delta(0)$'s. One can also see from eq.(2.4.9) how the determinant $\det' M$, where the zero modes are excluded, will arise upon integration over the $\mathbf{y}$'s, since

$$\det' M = m_1 \cdots m_k \quad . \tag{2.4.10}$$

Coming back to our discussion of the path-integral (eq.(2.4.4)), the next step according to our discussion in section 2.1 is to split the vector potential and the fields into $A = \tilde{A} + A'$ and $P = \tilde{P} + P'$, respectively. Since $M\tilde{P} = 0$, and P' is orthogonal to $\tilde{P}$, we have $P' = M\phi$, for some ϕ. Since D has no zero modes, in order for ϕ to have the same number of degrees of freedom as P', it has to be orthogonal to the zero modes of MD.*

Proceeding as in section 2.1, we arrive at

$$e^{iW[J]} = \int \mathcal{D}\tilde{A}\mathcal{D}\tilde{P}\mathcal{D}\phi\mathcal{D}A_0\tilde{\Delta}[\tilde{A}]e^{iS} \ . \tag{2.4.11}$$

The modified Faddeev-Popov determinant is

$$\tilde{\Delta} = \frac{1}{V_{\mathcal{K}}}\mathrm{det}'MD \ , \tag{2.4.12}$$

where $V_{\mathcal{K}}$ is the (infinite) volume of the space of zero modes $\mathcal{K}$. However, integration over A_0^a produces a factor $V_{\mathcal{K}}$, because the integrand does not depend on the projection of A_0^a onto the space $\mathcal{K}$. These two factors cancel and the final result is a finite expression, that is the same as eq.(2.1.10), with no ambiguity in the choice of ϕ, because out of all possible solutions of eq.(2.1.9) we have to choose the one that is orthogonal to the zero modes of the operator MD.

To illustrate the above results, we shall now consider the example of the Coulomb gauge [7],

$$\partial_i A_i^a = 0 \ . \tag{2.4.13}$$

The operator M is thus $M = \nabla$. The vector potential is split into its transverse and longitudinal components: $A = A_T + A_L$. Eq.(2.4.13) implies $A_L = 0$. The momentum is split similarly: $P = P_T + P_L$, and we write $P_L = \nabla\phi$ (which is possible, because $\nabla \times P_L = 0$), where ϕ is the solution of Gauss's law: $\nabla^2\phi^a + if^{abc}A_T^b \cdot \nabla\phi^c = \rho^a$, which is orthogonal to the zero modes of the operator $\nabla^2\delta^{ac} + if^{abc}A_T^b \cdot \nabla$. The hamiltonian is $\mathcal{H}_0 = \frac{1}{2g^2}(P_T^2 + (\nabla\phi)^2 + B^2)$. We shall use these results in section 2.6,

---

⋆ Notice the error in ref.[6], where ϕ was not restricted. We are indebted to A. P. Polychronakos for pointing this out to us.

where the vacuum expectation value of a Wilson loop in two dimensions is computed explicitly.

2.5 Lattice gauge theories

While the Faddeev-Popov procedure is necessary for perturbation theory, vacuum expectation values of gauge-invariant operators $X[A,P]$:

$$\langle X \rangle \equiv \frac{\int \mathcal{D}A\mathcal{D}PX[A,P]e^{-S}}{\int \mathcal{D}A\mathcal{D}Pe^{-S}} \ , \tag{2.5.1}$$

can be defined by placing the system on a lattice. In this case, no gauge-fixing is needed and direct numerical calculations are sufficient to extract numbers. However, since no powerful enough computers have yet been developed, analytical techniques are still useful. As we now have two definitions of gauge theories (perturbation theory *à la* Faddeev-Popov and lattices), it is important to check that they are identical. This is easily done in two dimensions, because two-dimensional pure Yang-Mills theories are exactly solvable models [8]. As we shall see, the two definitions do not agree with each other. In this section, we derive the exact solution on a lattice and compare it to perturbation theory on a lattice. In the next section, we shall do the same calculations using continuum perturbation theory. In that case, we will be able to trace the cause of the discrepancy in the second Gribov problem.

The partition function for a pure Yang-Mills theory on a lattice is

$$Z_N(g) = \int \mathcal{D}U \exp\left\{ \frac{1}{g^2} \sum_{plaq.} \mathrm{tr}(U_{12}U_{23}U_{34}U_{41} + h.c.) \right\} \ , \tag{2.5.2}$$

where N is the number of plaquettes and U_{ab} is the Wilson factor along the link joining the points a and b: $U_{ab} = \mathcal{P}e^{i\int_a^b A_\mu dx^\mu}$. The points $1,2,3,4$ in the exponent (eq.(2.5.2)) lie on a single plaquette, and the sum is over all plaquettes. We start by integrating out one edge link. The relevant piece is

$$I(g) = \int dU \exp\left\{\frac{1}{g^2}\mathrm{tr}(UV^\dagger + VU^\dagger)\right\} \ , \tag{2.5.3}$$

where U is the Wilson factor along the edge link and V traces the other three sides of the plaquette. I is independent of V, as can be seen by the change of variables $U \to UV$, because $dU \to dU$. It follows that

$$Z_N(g) = I(g)Z_{N-1}(g) \ , \tag{2.5.4}$$

which is a recursion relation whose solution is

$$Z_N(g) = [I(g)]^N \ . \tag{2.5.5}$$

We shall now calculate the vacuum expectation value of a Wilson loop, $X = \prod_{loop} U$. Since

$$\langle X \rangle = \frac{1}{Z_N(g)} \int \mathcal{D}U \ X \ \exp\left\{\frac{1}{g^2}\sum_{plaq.} \mathrm{tr}(U_{12}U_{23}U_{34}U_{41} + h.c.)\right\} \ , \tag{2.5.6}$$

by integrating over links that are outside the loop, we get factors of $I(g)$ in both the numerator and the denominator that cancel. Therefore, if there are A plaquettes inside the loop, eq.(2.5.6) becomes

$$\langle X \rangle = \frac{1}{Z_A(g)} \int \mathcal{D}U \ X \ \exp\left\{\frac{1}{g^2}\sum_{plaq.} \mathrm{tr}(U_{12}U_{23}U_{34}U_{41} + h.c.)\right\} \ , \tag{2.5.7}$$

where we only integrate over the plaquettes inside the loop. To integrate over an edge link, we observe that the relevant piece is

$$J(g) = \int dU\ U\ \exp\left\{\frac{1}{g^2}\mathrm{tr}(UV^\dagger + VU^\dagger)\right\} \ . \tag{2.5.8}$$

Evidently,

$$J(g) = -Vg^4\frac{d}{dg^2}I(g) \ . \tag{2.5.9}$$

Repeating the integration until we exhaust all plaquettes, we obtain

$$\langle X\rangle = \left(-g^4\frac{d\ln I}{dg^2}\right)^A \ . \tag{2.5.10}$$

Using eq.(2.5.3), we find [9] (setting $V = I$)

$$-g^4\frac{d\ln I}{dg^2} = 1 - \frac{N^2-1}{4N^2}(Ng^2) + \frac{N^2-1}{32N^4}(Ng^2)^2 + \cdots \ , \tag{2.5.11}$$

for an $SU(N)$ gauge theory. Notice that the series terminates in the large-N limit [10].

Let us now use perturbation theory. For simplicity, take two plaquettes. After we integrate over an edge link, we find

$$Z(g) = \int dU_1 dU_2 \exp\left\{\frac{1}{g^2}\mathrm{tr}(U_1U_2^\dagger + U_2U_1^\dagger)\right\} \ . \tag{2.5.12}$$

To do perturbation theory, we expand U_1 and U_2 around the identity:

$$U = e^{iu} = 1 + iu - \frac{1}{2}u^2 + \cdots \ , \ u = u^\dagger \ , \ u = u_aT^a \ . \tag{2.5.13}$$

The measure is

$$dU = \prod du_a \det\frac{\sin\frac{1}{2}u_b\Omega^b}{\frac{1}{2}u_b\Omega^b} \ , \tag{2.5.14}$$

where the matrices Ω^b are in the adjoint representation of $SU(N)$. We obtain [11]

$$-g^4 \frac{d \ln I}{dg^2} = 1 - \frac{N^2-1}{4N^2}(Ng^2) + \frac{(N^2-1)(N^2-2)}{64N^4}(Ng^2)^2 + \cdots \ , \tag{2.5.15}$$

which disagrees with eq.(2.5.11) to second order in Ng^2. The discrepancy is

$$\frac{(N^2-1)(N^2-4)}{64N^4}(Ng^2)^2 \ , \tag{2.5.16}$$

and it does not vanish in the large-N limit.

This result does not change if more plaquettes are considered, as a study of the $SU(2)$ case indicates [12]. The source of the error in perturbation theory can be traced to its severe infrared divergences. Remarkably enough, the introduction of an infrared regulator leads to finite expressions for all gauge-invariant Green functions, when the regulator is removed. Unfortunately, as we have just discussed, those expressions are incorrect [11,12,13]. We do not have a satisfactory solution for this problem. In the next section, we shall see that the same problem arises in the continuum theory. The problem can be solved in that case. As the source of the error is the same (infrared divergences), it should be possible to find a solution along the same lines in the case of a lattice.

2.6 Two dimensions

In this section, we show that there is an error in continuum perturbation theory in two dimensions, and that the error can be traced to the fact that the second Gribov problem [2] has not been taken into account [14].

In general, the Gribov problem does not affect perturbation theory, because it is usually possible to expand around the zero configuration ($A_\mu = 0$). We shall see that in our two-dimensional model this is not the case. Thus, the naïve perturbation expansion is incorrect, although still possible, because there is no ambiguity right at the zero potential.

We shall modify the naïve procedure in the manner described in section 2.4 [6]. This amounts to correctly treating the zero modes of the Faddeev-Popov determinant (eq.(2.1.8)). We work in the Coulomb gauge employing the hamiltonian formalism. As an infrared regulator, we enclose the system in a box of size $2L$ $(-L \leq x_0, x_1 \leq L)$, imposing vanishing boundary conditions. Ultraviolet divergences are regulated by a one-dimensional Pauli-Villars regulator, as we shall explain. Specifically, we shall calculate the expectation value of a square Wilson loop of size $1/M$, where M is the ultraviolet cutoff:

$$W \equiv \frac{1}{N}\langle \mathrm{tr}\mathcal{P}e^{i\int A_\mu dx^\mu}\rangle = \frac{1}{N}\langle e^{iM^{-2}F}\rangle + o(1/M) \quad , \tag{2.6.1}$$

where $F = \partial_0 A_1 - \partial_1 A_0 + [A_1, A_0]$ is the only non-vanishing component of the field strength. As was explained in the previous section, W can be calculated exactly, if we put the system on a two-dimensional lattice. The exact answer is given by eq.(2.5.11). We can also find the β-function of the theory by considering the Callan-Symanzik equation for the string tension $\sigma = -\frac{1}{a^2}\ln W$. We obtain

$$\beta(\alpha) \equiv a\frac{d\alpha}{da} = 2\alpha + \frac{N^2-2}{4N^2}\alpha^2 + o(\alpha^3) \quad , \tag{2.6.2}$$

where $\alpha = Ng^2$.

We now proceed to derive an expression for the Wilson loop in continuum perturbation theory. The generators of the gauge group $SU(N)$ obey the algebra

$$[T^a, T^b] = f^{abc}T^c \quad , \tag{2.6.3}$$

and are normalized by $\mathrm{tr}T^aT^b = \frac{1}{2}\delta^{ab}$. The generating functional is

$$Z[J^\mu] = \int \mathcal{D}P_1\mathcal{D}A_1\mathcal{D}A_0 e^{iS} \quad , \tag{2.6.4}$$

where $S = \int d^2x(\frac{1}{g^2}P_1^a\partial_0 A_1^a - h)$ and h is the hamiltonian density

$$h = \frac{1}{2g^2}P_1^aP_1^a + A_1^aJ_1^a - \frac{1}{g^2}A_0^a(D_1^{ac}P_1^c - g^2J_0^a) \quad . \tag{2.6.5}$$

$P_1 = P_1^aT^a = F$ is the momentum conjugate to A_1, and $D_\mu^{ac} = \partial_\mu\delta^{ac} + if^{abc}A_\mu^b$ is the covariant derivative.[♮]

We fix the gauge by imposing the Coulomb gauge (cf. eq.(2.4.13))

$$\partial_1 A_1^a = 0 \quad . \tag{2.6.6}$$

Notice that in two-dimensions this condition implies that A_1 is independent of the spatial coordinate x_1. The Faddeev-Popov determinant is (cf. eq.(2.1.8))

$$\Delta = \det \partial_1 D_1 \quad . \tag{2.6.7}$$

Integrating over A_0^a in the generating functional we obtain Gauss's law as a constraint,

$$D_1^{ac} P_1^c - g^2 J_0^a = 0 \quad . \tag{2.6.8}$$

Splitting P_1^a in its transverse and longitudinal components, $P_1^a = P_T^a + P_L^a$, where $\partial_1 P_T^a = 0$ and $P_L^a = \partial_1 \phi^a$ for some ϕ^a , Gauss's law becomes

$$\partial_1^2 \phi^a + i f^{abc} A_1^b \partial_1 \phi^c = \rho^a \quad , \tag{2.6.9}$$

where $\rho^a = g^2 J_0^a - i f^{abc} A_1^b P_T^c$ is the total color charge density. In the measure we have

$$\mathcal{D}P_1 = \mathcal{D}P_T \mathcal{D}\phi \quad , \tag{2.6.10}$$

where we have omitted the jacobian $\det \partial_1$, which is a constant and can therefore be absorbed into the overall normalization of the path integral. The hamiltonian density becomes

$$h_0 = \frac{1}{2g^2} P_T^2 + \frac{1}{2g^2} (\partial_1 \phi)^2 + A_1^a J_1^a \quad . \tag{2.6.11}$$

Integration over ϕ^a produces a factor Δ^{-1} , because of the constraint (eq.(2.6.9)) in the path integral, which exactly cancels the Faddeev-Popov determinant. Thus,

---

♮ The coupling constant now has the dimensions of mass. A dimensionless coupling constant can be defined by $\alpha = Ng^2/M^2$, where M is a large momentum cutoff.

the Gribov problem consists in choosing the correct ϕ^a out of all the solutions of eq.(2.6.9). This ambiguity only exists when the operator $\partial_1 D_1$ has zero modes, i.e., when $\Delta = 0$.

Disregarding the problem for the moment, we can formally solve eq.(2.6.9) for ϕ^a. We obtain

$$\phi^a = \frac{1}{\partial_1^2}\rho^a - i\frac{1}{\partial_1^2}\partial_1\frac{1}{\partial_1^2}f^{abc}A_1^b\rho^c + \frac{1}{\partial_1^2}\partial_1\frac{1}{\partial_1^2}\partial_1\frac{1}{\partial_1^2}f^{abc}A_1^b f^{cde}A_1^d\rho^e + \cdots \quad . \tag{2.6.12}$$

Since we have imposed vanishing boundary conditions, the Green function for the operator ∂_1^2 is

$$G_0(x,y) = \frac{1}{2}L - \frac{1}{2}|x-y| - \frac{1}{2L}xy \quad , \tag{2.6.13}$$

satisfying $\partial_1^2 G_0(x,y) = -\delta(x-y)$. We regulate $G_0(x,y)$ by introducing a Pauli-Villars field of mass M. The explicit form of the regulated propagator is

$$G_0^{\text{reg}}(x,y) = G_0(x,y) - \frac{1}{4M\sin ML}[\cos M(L-|x-y|) - \cos M(L+x+y)] \quad . \tag{2.6.14}$$

Henceforth, we shall omit the superscript 'reg' for simplicity. Eq.(2.6.12) reads

$$\begin{aligned}\phi^a(x) = &\int_{-L}^{L} dy G_0(x,y)\rho^a(y) \\ &- \int_{-L}^{L} dy G_1(x,y) i f^{abc}A_1^b\rho^c(y) + \int_{-L}^{L} dy G_2(x,y) f^{abc}A_1^b f^{cde}A_1^d\rho^e(y) + \cdots \quad ,\end{aligned} \tag{2.6.15}$$

where

$$G_1(x,y) = \int_{-L}^{L} dz G_0(x,z)\partial_1 G_0(z,y) \quad , \tag{2.6.16}$$

$$G_2(x,y) = \int_{-L}^{L} dz G_0(x,z)\partial_1 G_1(z,y) \quad . \tag{2.6.17}$$

Therefore, the contribution to the hamiltonian coming from the longitudinal part of the field strength is

$$
\begin{aligned}
h' \equiv & -\frac{1}{2}\int_{-L}^{L} dx_1 \phi^a \partial_1^2 \phi^a \\
= & \frac{1}{2}\int dxdy G_0(x,y)\rho^a(x)\rho^a(y) - \int dxdy G_1(x,y) f^{abc} A_1^a \rho^b(x)\rho^c(y) \\
& + \frac{3}{2}\int dxdy G_2(x,y) f^{abc} f^{ade} A_1^b \rho^c(x) A_1^d \rho^e(y) + \cdots \quad ,
\end{aligned}
\tag{2.6.18}
$$

where we have rescaled fields by g . This is the part of the hamiltonian describing the interactions. The rest contains a kinematical quadratic in P_T piece. To find the propagator for A_1 , we diagonalize the quadratic part of the lagrangian by completing the square. By shifting $P_T \rightarrow P_T + \partial_0 A_1$, we obtain the gaussian piece

$$
h_1 = L(P_T^a P_T^a + \partial_0 A_1^a \partial_0 A_1^a) \quad . \tag{2.6.19}
$$

Having eliminated all the unphysical degrees of freedom, we are now in a position to calculate correlation functions perturbatively. We shall calculate a square Wilson loop, of side $1/M$, which, as an expansion in the dimensionless coupling constant $\alpha = Ng^2/M^2$, is

$$
W = 1 + \alpha W_1 + \alpha^2 W_2 + o(\alpha^3) \quad . \tag{2.6.20}
$$

We are interested in the limit of an infinite box. Thus, we shall drop all terms that vanish in the limit $L \rightarrow \infty$. To first order in the coupling constant we have

$$
\begin{aligned}
W_1 &= -\frac{1}{2N^2}\frac{1}{M^2}\langle(\partial_0 A_1 - \partial_1 A_0)^2\rangle_0 + o\left(\frac{1}{ML}\right) \\
&= -\frac{1}{4N^2}\frac{1}{M^2}\left(\partial_1 \frac{\delta}{\delta J_0^a}\right)^2 Z[J_\mu]\Big|_{J_\mu=0} + o\left(\frac{1}{ML}\right) \\
&= -\frac{N^2-1}{4N^2}\frac{1}{M^2}M^2 + o\left(\frac{1}{ML}\right) \quad .
\end{aligned}
\tag{2.6.21}
$$

Taking the limit $ML \to \infty$, we obtain

$$W_1 = -\frac{N^2-1}{4N^2} \quad . \tag{2.6.22}$$

To second order in α , there are two contributions. One comes from the α^2 piece of $\langle \mathrm{tr}F^4 \rangle$, the other one comes from $\langle \mathrm{tr}F^2 \rangle$ (cf eq.(2.6.1)). There is a potential contribution from $\langle \mathrm{tr}F^3 \rangle$, but it is seen to vanish in the infinite-ML limit. So

$$W_2 = W_2^{(1)} + W_2^{(2)} \quad , \tag{2.6.23}$$

where

$$\begin{aligned} W_2^{(1)} &= \frac{1}{24N^3}\frac{1}{M^4}\langle F^4\rangle_2 = \frac{1}{24N^3}\frac{1}{M^4}\langle(\partial_0 A_1 - \partial_1 A_0)^4\rangle_0 \\ &= \frac{1}{24N^3}\frac{1}{M^4}(T_1^a \cdots T_4^a)\partial_1 \frac{\delta}{\delta J_0^{a_1}} \cdots \partial_1 \frac{\delta}{\delta J_0^{a_4}} Z[J_\mu]\Big|_{J_\mu=0} + o\left(\frac{1}{ML}\right) \quad , \end{aligned} \tag{2.6.24}$$

and

$$W_2^{(2)} = -\frac{1}{2N^2}\frac{1}{M^2}\langle F^2\rangle_2 = -\frac{1}{2N^2}\frac{1}{M^2}\langle(\partial_0 A_1 - \partial_1 A_0 + [A_0, A_1])^2\rangle_2 \quad . \tag{2.6.25}$$

A short computation shows that

$$W_2^{(1)} = -\frac{1}{48}\frac{N^2-1}{N^2} + \frac{1}{32}\frac{N^2-1}{N^4} \quad . \tag{2.6.26}$$

We also find that all the infinities cancel between the various terms in $W_2^{(2)}$ and the result is

$$W_2^{(2)} = \frac{1}{96}\frac{N^2-1}{N^2} \quad . \tag{2.6.27}$$

Therefore,

$$W_2 = -\frac{1}{96}\frac{N^2-1}{N^2} + \frac{1}{32}\frac{N^2-1}{N^4} \quad . \tag{2.6.28}$$

Because of eqs.(2.6.22) and (2.6.28), eq.(2.6.20) becomes

$$W = 1 - \frac{N^2-1}{4N^2}\alpha + \left(-\frac{1}{48}\frac{(N^2-1)}{N^2} + \frac{1}{32}\frac{N^2-1}{N^4}\right)\alpha^2 + o(\alpha^3) \quad , \tag{2.6.29}$$

which disagrees with the correct expression (eq.(2.5.11)) for the Wilson loop in second order in α . We also obtain a different β-function. Following the same procedure as before, we obtain

$$\beta(\alpha) = 2\alpha + \left(\frac{N^2-2}{4N^2} - \frac{1}{12}\right)\alpha^2 + o(\alpha^3) \quad . \tag{2.6.30}$$

Notice that the extra term does not vanish in the large-N limit. As we saw in the previous section, one encounters the same problem on the lattice, even if it consists of only two plaquettes [11]. In that case it can be seen that the problem is due to the possibility of performing a global unitary transformation on the Wilson loops on all plaquettes. In our formulation, we shall see that the problem is due to the ambiguity in choosing the right ϕ^a in eq.(2.6.9).

According to the modification of the Faddeev-Popov procedure discussed in section 2.4, in order to account for the Gribov problem, we have to choose the solution of eq.(2.6.9) that is orthogonal to the zero modes. A word of caution is in order here. $\partial_1\phi$ represents the longitudinal component of P_1 , which is by construction orthogonal to the transverse component, P_T , in the sense that $\int_{-L}^{L} dx \mathrm{tr}\partial_1\phi P_T = 0$. Therefore, strictly speaking, the longitudinal component is $-\partial_1^\dagger\phi$, where $\partial_1^\dagger$ is the adjoint of ∂_1. We did not have to be careful above, because with our choice of ϕ (eq.(2.6.15)), $\partial_1^\dagger\phi = -\partial_1\phi$, but in general, $\partial_1^\dagger \neq -\partial_1$. Also the Faddeev-Popov determinant should be $\det(-\partial_1^\dagger D_1)$ instead of $\det \partial_1 D_1$ (eq.(2.6.7)). The zero modes of the operator $\partial_1^\dagger D_1$ are the solutions of the equation

$$\partial_1^\dagger\partial_1 u^a + i f^{abc} A_1^b \partial_1^\dagger u^c = 0 \quad . \tag{2.6.31}$$

This equation has N^2-1 linearly independent solutions, which we shall call $u_{(k)}(k = 1,...,N^2-1)$. Suppose that the $\partial_1^\dagger u_{(k)}$ form an orthonormal set, the inner product being $\langle w_1, w_2\rangle = \int_{-L}^{L} dx \mathrm{tr} w_1(x) w_2(x)$, for the vectors w_1 and w_2 . Then, in the hamiltonian, $\partial_1\phi$ has to be replaced by $\partial_1\omega$, where

$$\partial_1\omega = \partial_1\phi - \sum_{k=1}^{N^2-1} \partial_1^\dagger u_{(k)} \int_{-L}^{L} dx \partial_1^\dagger u_{(k)} \partial_1\phi \tag{2.6.32}$$

is the projection of $\partial_1\phi$ orthogonal to the zero modes. This implies that the interaction hamiltonian h' (eq.(2.6.18)) acquires an extra piece $\tilde{h}$, where

$$\tilde{h} = \sum_{k=1}^{N^2-1} \left(\int_{-L}^{L} dx u_{(k)} \partial_1^2 \phi \right)^2 , \tag{2.6.33}$$

which describes new interactions. It represents a non-trivial correction to ordinary perturbation theory. To write down the explicit form of $\tilde{h}$, we have to solve eq.(2.6.31). To this end, it is convenient to define a matrix v by $u = \partial_1 v$, where v is such that $\partial_1^\dagger v = -\partial_1 v$. Then, eq.(2.6.31) reads

$$\partial_1^3 v^a + i f^{abc} A_1^b \partial_1^2 v^c = 0 \ . \tag{2.6.34}$$

We can solve eq.(2.6.34) perturbatively. We find $N^2 - 1$ solutions,

$$v_{(k)}(x) = Q_0(x)T^k - Q_1(x)i[A_1, T^k] - Q_2(x)[A_1, [A_1, T^k]] + \cdots \ , \tag{2.6.35}$$

where $Q_0(x) = \frac{1}{\sqrt{2L}}(x^2 - L^2)$, and $Q_n(x) = \int_{-L}^{L} dy G_0(x,y) \partial_1 Q_{n-1}$, for $n \geq 1$. We can check that $\partial_1^\dagger v_{(k)} = -\partial_1 v_{(k)}$, which justifies our replacing of $\partial_1^\dagger$ by $-\partial_1$. We can also see that $\{\partial_1^2 v_{(k)}\}$ is an orthonormal set: $\int_{-L}^{L} dx \mathrm{tr} \partial_1^2 v_{(k)} \partial_1^2 v_{(l)} = \delta_{kl}$. Therefore, eq.(2.6.33) becomes

$$\tilde{h} = \sum_{k=1}^{N^2-1} \left(\int_{-L}^{L} dx \partial_1 v_{(k)} \partial_1^2 \phi \right)^2 \ . \tag{2.6.36}$$

Using eqs.(2.6.15) and (2.6.35), we obtain the perturbation expansion of $\tilde{h}$,

$$\begin{aligned} \tilde{h} = &\int dx dy \tilde{G}_0(x,y) \rho^a(x) \rho^a(y) + \int dx dy \tilde{G}_1(x,y) f^{abc} A_1^a \rho^b(x) \rho^c(y) \\ &+ \int dx dy \tilde{G}_2(x,y) f^{abc} f^{ade} A_1^b \rho^c(x) A_1^d \rho^e(y) + \cdots \ , \end{aligned} \tag{2.6.37}$$

where

$$\tilde{G}_0(x,y) = \partial_1 Q_0(x) \partial_1 Q_0(y) \ , \tag{2.6.38a}$$

$$\tilde{G}_1(x,y) = \partial_1 Q_0(x)[\partial_1 Q_1(y) + Q_0(y)] - (x \leftrightarrow y) \ , \qquad (2.6.38b)$$

$$\tilde{G}_2(x,y) = \partial_1 Q_0(x) \left\{ \partial_1 Q_2(y) + Q_1(y) - aQ_0(y) \int_{-L}^{L} dz \partial_1 G_0(y,z) \right\} + (x \leftrightarrow y)$$
$$+ [\partial_1 Q_1(x) + Q_0(x)][\partial_1 Q_1(y) + Q_0(y)] \ . \qquad (2.6.38c)$$

Coming back to the calculation of the Wilson loop, we easily see that W_1 does not change, nor does $W_2^{(1)}$. However, $W_2^{(2)}$ gets an additional contribution $\tilde{W}_2^{(2)}$ from the new interactions. Again, all infinities cancel and the result is

$$\tilde{W}_2^{(2)} = \frac{1}{96} \frac{N^2 - 1}{N^2} \ . \qquad (2.6.39)$$

Therefore, to second order in the coupling constant, we obtain (using eqs.(2.6.26), (2.6.27) and (2.6.39))

$$W_2 = \frac{1}{32} \frac{N^2 - 1}{N^4} \ , \qquad (2.6.40)$$

in agreement with the exact result (eq.(2.5.11)).

The above analysis can be generalized to four dimensions. Clearly, a lot more work is needed to see if the Gribov ambiguity has perturbative effects there, as calculations in four dimensions are a lot more involved. It would also be interesting to investigate the possibility of carrying out a similar analysis on the lattice using the gauge-invariant Wilson action. Unfortunately, we have not been able to find a procedure that solves the Gribov problem in this case.

References

1. L. D. Faddeev and V. N. Popov, *Phys. Lett.* **25B** (1967) 30.
2. V. N. Gribov, *Nucl. Phys.* **B139** (1978) 1.
3. I. M. Singer, *Comm. Math. Phys.* **60** (1978) 7.
4. T. Maskawa and H. Nakajima, *Prog. Theor. Phys.* **60** (1978) 526;
   V. Moncrief, *J. Math. Phys.* **20** (1979) 579;
   E. Malec, *J. Math. Phys.* **23** (1982) 21.
5. R. Jackiw, I. Muzinich and C. Rebbi, *Phys. Rev.* **D17** (1978) 1576;
   C. Bender, T. Eguchi and H. Pagels, *Phys. Rev.* **D17** (1978) 1086.
6. G. Siopsis, *Phys. Lett.* **136B** (1984) 175.
7. R. D. Peccei, *Phys. Rev.* **D17** (1978) 1097.
8. M. W. Roth, *Phys. Rev.* **D15** (1976) 1084.
9. P. Cvitanović, J. Greensite and B. Lautrup, *Phys. Lett.* **105B** (1981) 197.
10. D. J. Gross and E. Witten, *Phys. Rev.* **D21** (1980) 446.
11. Y. Brihaye and P. Rossi, *Nucl. Phys.* **B235** (1984) 226.
12. V. F. Müller and W. Rühl, *Ann. Phys. (N.Y.)* **133** (1981) 240.
13. A. Patrascioiu, *Phys. Rev. Lett.* **54** (1985) 1102.
14. G. Siopsis, *Phys. Lett.* **187B** (1987) 351.

3. Chiral and conformal anomalies

3.1 The geometrical connection

It was first observed by Steinberger, then rediscovered by Schwinger and again later by Adler, Bell and Jackiw [1] that the divergence of the axial current in a gauge theory does not vanish, due to quantum effects. These effects have come to be known as "anomalies." In chiral gauge theories, they seem to be unwanted, because the effective action ceases to be gauge-invariant in their presence [2].

To study the effect, consider a Yang-Mills theory with gauge group G coupled to a doublet of left-handed fermions. The partition function is

$$e^{iW} = \int \mathcal{D}A_\mu e^{i\int \frac{1}{4}F^a_{\mu\nu}F^a_{\mu\nu}} e^{iW_f[A_\mu]} \quad , \tag{3.1.1}$$

where

$$e^{iW_f[A_\mu]} = \int \mathcal{D}\psi_L \mathcal{D}\bar{\psi}_L e^{i\int \bar{\psi} i \not{D} P_- \psi} \quad , \tag{3.1.2}$$

$iD_\mu = i\partial_\mu + eA_\mu$, and $P_\pm = \frac{1}{2}(1 \pm \gamma_5)$ are projection operators onto the right(left)-handed fermions. In general, under a gauge transformation, we have

$$W_f[A_\mu + D_\mu\alpha] = W_f[A_\mu] + \int \alpha^a Q^a[A_\mu] \ . \tag{3.1.3}$$

If Q does not vanish, gauge-invariance is broken. Thus, Q is the anomaly. In terms of the generators of gauge transformations,

$$X^a = -\partial_\mu \frac{\delta}{\delta A^a_\mu} - f^{abc} A^b_\mu \frac{\delta}{\delta A^c_\mu} \quad , \tag{3.1.4}$$

eq.(3.1.3) reads $X^a W_f = Q^a$. Since the X^a's satisfy

$$[X^a(x), X^b(y)] = f^{abc} X^c(x) \delta(x-y) \quad , \tag{3.1.5}$$

the anomaly has to satisfy

$$X^a(x) Q^b(y) - X^b(y) Q^a(x) = f^{abc} Q^c(x) \delta(x-y) \ . \tag{3.1.6}$$

These are the celebrated Wess-Zumino consistency conditions [3]. To put them in a more geometrical form, define an operation δ by[†]

$$\delta A_\mu = -D_\mu \alpha \quad , \quad \delta \alpha^a = -\frac{1}{2} f^{abc} \alpha^b \alpha^c \ , \tag{3.1.7}$$

where α is considered to be an anti-commuting variable. Then eq.(3.1.6) can be written as

$$\delta \int \alpha^a Q^a = 0 \ . \tag{3.1.8}$$

Since $\delta^2 = 0$, this equation has an obvious solution: $\int \alpha^a Q^a = \delta X$, where X is any functional of A^μ. However, these solutions are trivial in the sense that they can be canceled by the addition of counterterms in the effective action. Our problem is therefore to see whether the anomaly is a non-trivial solution of eq.(3.1.8).

We can now see the connection with topology. The operator δ defines a cohomology. The anomaly belongs to a non-trivial cohomology class of δ. We are therefore naturally led to a study of the structure of orbits of gauge transformations. Each orbit is a replica of the group $\mathcal{G}$ of all maps from the four-dimensional sphere S^4 to the gauge group G.[‡] To see this more explicitly, consider a closed path in $\mathcal{G}$: $\{\Lambda(\theta) : \theta \in [0,1]\}$, $\Lambda(0) = \Lambda(1) =$*identity*. Let $A_\mu^\Lambda = \Lambda^{-1} A_\mu \Lambda + \Lambda^{-1}\partial_\mu \Lambda$. The set $\{A^{\Lambda(\theta)} : \theta \in [0,1]\}$ is a path in the orbit that goes through A_μ. Now imagine carrying $e^{iW[A]}$ along this path, from $\theta = 0$ to $\theta = 1$. As it goes from point θ to $\theta + d\theta$, it picks up a factor $e^{id\phi}$ where the phase is $d\phi = \int_{S^4} \mathrm{tr}\alpha Q$, and $\alpha = \Lambda^{-1}\frac{\partial \Lambda}{\partial \theta} d\theta$ is the infinitesimal gauge transformation along the path at θ. The total phase is $\int_{\theta=0}^{\theta=1} d\phi$ and has to be a multiple of 2π. If we can continuously contract this path to a point, the total phase must vanish. In this case, there is no anomaly. Therefore, the anomaly is associated with the existence of a non-contractable path in $\mathcal{G}$. Such a path exists whenever the first homotopy group $\pi_1(\mathcal{G}) = \pi_5(G)$ is non-trivial.[§] It follows that an

† δ is generally known as the BRS operator [4].

‡ Note the difference with section 2.3, where we considered time-independent gauge transformations, and therefore the maps were from S^3 to G.

§ Comparing with the results of section 2.3, we see that anomalies have the same geometrical origin as the Gribov problem. This connection is yet to be explored [5].

anomaly exists only if the group $\mathcal{G}$ is topologically non-trivial. However, this makes it hard to obtain any results on differential geometry by working in an orbit, because orbits are replicas of $\mathcal{G}$. It is therefore imperative that we consider the space $\mathcal{A}$ of all vector potentials, which is topologically trivial (cf. section 2.3).

The space $\mathcal{A}$ is endowed with a connection

$$w = D^{-1}\delta A \ . \tag{3.1.9}$$

On the other hand, A_μ itself is a connection in the fiber bundle $\mathcal{B}$ over S^4 with group G. Therefore, the connection in the space $\mathcal{B} \times \mathcal{A}$ is $A + w$. Using the language of forms,♮ we can define the curvature in the space $\mathcal{B} \times \mathcal{A}$ by

$$\mathcal{F} = (d+\delta)(A+w) + (A+w)^2 \ , \tag{3.1.10}$$

where d is the ordinary derivative in S^4 and δ is the derivative in $\mathcal{A}$.

Atiyah and Singer have derived a powerful result (the Family Index Theorem, [6]) that relates the curvature $\mathcal{F}$ to differential geometry, via the Dirac operator, $D_- = \not{D} P_-$. They define a space $\mathcal{I}$ (the "Index" of D_-) endowed with a connection

$$w' = D_-^{-1}\delta D_- \ . \tag{3.1.11}$$

The Family Index Theorem establishes a connection between w and w'. For our purposes, it is sufficient to consider the result*

$$\mathrm{ch}_1(\mathcal{I}) = \int_{S^4} \mathrm{ch}_5(\mathcal{B} \times \mathcal{A}) \ , \tag{3.1.12}$$

where the Chern-Simons form is defined by $\mathrm{ch}_j(\mathcal{B} \times \mathcal{A}) = \frac{1}{(2\pi)^j}\mathrm{tr}\mathcal{F}^j$, and similarly for $\mathcal{I}$.

♮ A one-form is $A = A_\mu dx^\mu$, where the dx^μ anti-commute. The field strength is the two-form $F \equiv F_{\mu\nu} dx^\mu dx^\nu = dA + A^2$.

⋆ The Family Index Theorem is a series of equations relating Chern-Simons forms to all orders. Eq.(3.1.12) is one of these relations. As we discussed above, it is related to the non-triviality of $\pi_5(G)$. The other equations are related to higher homotopy groups ($\pi_{2j+1}(G)$, $j > 2$) [6]. However, their physical implications are yet to be seen.

By defining $\mathcal{F}_t = t\mathcal{F} + (t - t^2)w^2$, it is easy to see that

$$\mathrm{tr}\mathcal{F}^j = (d+\delta)j \int_0^1 dt \mathrm{tr}(w\mathcal{F}_t^{j-1}) \ . \tag{3.1.13}$$

Thus, the right-hand side of eq.(3.1.12) becomes $\int_{S^4} \mathrm{ch}_5(\mathcal{B} \times \mathcal{A}) = \delta Q$ where

$$Q = \int_0^1 dt \int_{S^4} \frac{5}{(2\pi)^5} \mathrm{tr}(w\mathcal{F}_t^4) \ . \tag{3.1.14}$$

Also, the left-hand side of eq.(3.1.12) is

$$ch_1(\mathcal{I}) = \delta \frac{1}{2\pi} \mathrm{tr} w' \ . \tag{3.1.15}$$

Therefore,

$$\mathrm{tr} w' = 2\pi Q + \delta h \ . \tag{3.1.16}$$

We were able to obtain eq.(3.1.16), because the space $\mathcal{A}$ is topologically trivial ($\delta v = 0$ implies $v = \delta h$, for some form h). We can now restrict ourselves to a specific orbit. The vectors tangent to an orbit are of the form $\delta A = -D\alpha$. Therefore, the connection is $w = D^{-1}\delta A = -\alpha$. Also, $\delta\alpha = -\alpha^2$. It follows that, when restricted to an orbit, δ coincides with the BRS operator that we discussed above.

A straightforward calculation shows that in an orbit,

$$\mathrm{tr} w' = \frac{1}{24\pi^2} \int_{S^4} \mathrm{tr}\alpha d(AdA + \frac{1}{3}A^3) + \delta h \ . \tag{3.1.17}$$

Also, $\mathrm{tr} w' = \mathrm{tr} D_-^{-1} \delta D_- = \delta \mathrm{tr} \ln D_- = \delta \det D_-$ and $\det D_- = e^{iW}$. Therefore, the right-hand side of eq.(3.1.17) is the anomaly. Comparing with eq.(3.1.3), we obtain

$$\int \mathrm{tr}\alpha Q = \frac{1}{24\pi^2} \int_{S^4} \mathrm{tr}\alpha d(AdA + \frac{1}{3}A^3) \ . \tag{3.1.18}$$

Having established that the anomaly exists for a wide class of theories, it is natural to ask whether it is still possible to find viable theories when anomalies are present. The remainder of this chapter is devoted to this task. We shall first discuss a proposal by Faddeev and Shatashvili [7] to circumvent the problem by adding a degree of freedom. In the other sections of this chapter, we investigate approaches that do not make use of additional fields.

Faddeev and Shatashvili's suggestion consists in changing the definition of the partition function (eq.(3.1.1)) to

$$e^{iW} = \int \mathcal{D}A_\mu \mathcal{D}\Lambda e^{i\int \frac{1}{4}F^a_{\mu\nu}F^a_{\mu\nu}} e^{iW_f[A^\Lambda_\mu]} \quad , \tag{3.1.19}$$

where $A^\Lambda_\mu = \Lambda^{-1}A_\mu\Lambda + \Lambda^{-1}\partial_\mu\Lambda$. This theory possesses a gauge invariance:

$$A \to U^{-1}AU + U^{-1}dU \quad , \quad \Lambda \to U^{-1}\Lambda \ . \tag{3.1.20}$$

In a $U(1)$ theory, the effect of the new field Λ is to enforce the additional constraint

$$\varepsilon^{\mu\nu\rho\sigma}F_{\mu\nu}F_{\rho\sigma} = 0 \ , \tag{3.1.21}$$

i.e., that the anomaly vanish. Unfortunately, it has not yet been possible to define a perturbation theory.

The procedure is also applicable to conformal anomalies. As an example, consider strings. The classical action is

$$S = \frac{1}{2}\int d^2\sigma \sqrt{-\det g^{ab}}\, g^{ab}\partial_a x^\mu \partial_b x^\mu \quad , \tag{3.1.22}$$

where $g^{ab}(\sigma)$ is the metric on the world-sheet and $x^\mu(\sigma)$ are coordinates in space-time. It is a geometrical fact that every metric on a two-dimensional surface can be brought to the form $g^{ab} = e^\phi \delta^{ab}$. (We assume that the surface has euclidean signature.) Hence, (3.1.22) becomes

$$S = \frac{1}{2}\int d^2\sigma \partial_a x^\mu \partial_a x_\mu \ , \tag{3.1.23}$$

and therefore the action is independent of the conformal factor, $\phi(\sigma)$. However, in the quantum theory, this factor does not decouple, due to the conformal anomaly

[8]. Instead, one has to integrate over it. Thus, ϕ plays the rôle of Λ in eq.(3.1.19). It follows that Faddeev and Shatashvili's ansatz [7] coincides with Polyakov's [8] in this case. As with chiral theories, however, it is hard to develop perturbation theory, because the action for ϕ is hard to work with, being a Liouville action.

3.2 New chiral theories

In this section, we describe regulators that lead to inequivalent chiral gauge theories. There are two ways of regulating the jacobian of gauge transformations coming from the measure $\mathcal{D}\psi\mathcal{D}\bar{\psi}$ in the path-integral. This has led to a controversy [9], because the two methods lead to different forms of the anomaly. One of them satisfies the Wess-Zumino conditions [3] and is therefore named the consistent anomaly (eq.(3.1.18)). The other one does not satisfy these conditions, and, therefore, it cannot be related to the consistent anomaly by the addition of a counterterm in the effective action. It can be expressed in terms of the field strength $F^{\mu\nu}$, hence it is called the covariant anomaly. Inasmuch as the Wess-Zumino conditions are fundamental, the latter anomaly is rejected as not representing the variation of the path integral. We shall show explicitly that the covariant anomaly can be derived using the path-integral formalism. This peculiar fact cannot be attributed to the inconsistency of a theory with anomalies, because the two methods produce conflicting results even in anomaly-free theories. In particular, they disagree on the divergence of the $U(1)$ current [3]. We shall argue that the two forms of the anomaly belong to two distinct theories satisfying different Dyson-Schwinger equations.

For pedagogical reasons, we start by discussing the axial anomaly in a gauge theory with only vector couplings [10]. We obtain Ward identities by studying the transformation properties of the measure

$$\mathcal{D}\mu = \frac{1}{Z}\mathcal{D}\psi\mathcal{D}\bar{\psi}e^{iS} \ , \tag{3.2.1}$$

where $Z = \int e^{iS}$, $S = \int \bar{\psi}i\not{D}\psi$ is the action, and $iD_\mu = i\partial_\mu + eA_\mu$, under the action of a chiral transformation,

$$\psi(x) \to \psi(x) + ie\gamma_5\alpha(x)\psi(x) \ , \ \bar{\psi}(x) \to \bar{\psi}(x) + ie\bar{\psi}(x)\gamma_5\alpha(x) \ . \tag{3.2.2}$$

Our regularization procedure consists of two steps. In the first step we define the regularized lagrangian as the limit $\lim_{\epsilon\to 0}\mathcal{L}_\epsilon$, where

$$\mathcal{L}_\epsilon = \bar{\psi} i\not{D} g(-\epsilon D_\mu D^\mu)\psi \ , \tag{3.2.3}$$

$g(x)f(x) = 1$, and f is a function satisfying $f(0) = 1$ and $f(x) \to 0$ sufficiently fast, as $x \to \infty$. In the second step we define perturbation theory by requiring that all operators be expanded in powers of ϵ , before evaluating diagrams. In essence, this means that, at each vertex, the factor $g(\epsilon p^2)$, where p^μ is the momentum, will have to be expanded, before any loop integrations are performed. Notice that the inverse propagator in the regularized lagrangian $\mathcal{L}_\epsilon$ contains a factor of $g(\epsilon p^2)$, rendering all diagrams finite.

For convenience, we introduce the basis $\{\phi_n\}$, which consists of the eigenfunctions of the Dirac operator,

$$i\not{D}\phi_n = \lambda_n\phi_n \ . \tag{3.2.4}$$

Expressing ψ and $\bar{\psi}$ in terms of the basis $\{\phi_n\}$,

$$\psi = \sum_n a_n\phi_n \ , \ \bar{\psi} = \sum_n b_n\phi_n^\dagger \ , \tag{3.2.5}$$

we obtain the current corresponding to a chiral transformation of $\mathcal{L}_\epsilon$,

$$\langle j^\mu_{5\epsilon}\rangle = \sum_n \frac{1}{\lambda_n}\phi_n^\dagger\gamma^\mu\gamma_5 g(\epsilon\lambda_n^2)\phi_n \ . \tag{3.2.6}$$

According to our prescription, this operator is defined by its formal expansion in powers of ϵ . The divergence of the current is

$$\langle \partial_\mu j^\mu_{5\epsilon} \rangle = \sum_{\ell=0}^{\infty} \frac{g^{(\ell)}(0)}{\ell!} \sum_n \phi_n^\dagger \gamma_5 (\epsilon \lambda_n^2)^\ell f(\epsilon \lambda_n^2) \phi_n \ . \tag{3.2.7}$$

By switching to a plane-wave basis, we can write this as

$$\langle \partial_\mu j^\mu_{5\epsilon} \rangle = \sum_{\ell=0}^{\infty} \frac{g^{(\ell)}(0)}{\ell!} \int \frac{d^4k}{(2\pi)^4} e^{-ikx} \mathrm{tr} \gamma_5 h_\ell (\epsilon (D_\mu D^\mu + \frac{i}{2}[\gamma_\mu, \gamma_\nu] F^{\mu\nu})) e^{ikx} \ , \tag{3.2.8}$$

where $h_\ell(x) = x^\ell f(x)$. It is a consequence of the properties of γ-matrices that the divergent part of the right-hand side of eq.(2.8) vanishes. The finite part is obtained by expanding h_ℓ about $\epsilon D_\mu D^\mu$. It is straightforward to see that it is proportional to

$$\frac{\ell(\ell-1)}{2} \int_0^\infty t^{\ell-1} f(t) dt + \ell \int_0^\infty t^\ell f'(t) dt + \frac{1}{2} \int_0^\infty t^{\ell+1} f''(t) dt \ . \tag{3.2.9}$$

Integrating by parts as many times as needed, we see that this is zero for $\ell > 0$. It follows that only the $\ell = 0$ term contributes to the series in eq.(2.7). Therefore, only the lowest order contribution to the current is significant.† It is now straightforward [11] to show that

$$\partial_\mu j^\mu_5 = \frac{e^2}{16\pi^2} \varepsilon_{\mu\nu\rho\sigma} \mathrm{tr} F^{\mu\nu} F^{\rho\sigma} \ . \tag{3.2.10}$$

It is interesting to note that we could have obtained the same result if we had expressed the fermionic fields in terms of a plane-wave basis from the beginning. This would have meant that we were using perturbation theory with a regulated propagator $\not{p}^{-1} f(\epsilon p^2)$. This fact demonstrates the equivalence of the two bases.

Since the terms of higher-order in ϵ in the expansion of the current $j^\mu_{5\epsilon}$ vanish, we could have defined the axial current by its zeroth-order contribution $j^\mu_5 = \bar{\psi} \gamma^\mu \gamma_5 \psi$, without altering the results. It will be seen later that keeping higher-order terms in the expression for the current alters chiral theories.

† This justifies Fujikawa's method [11] of only regulating the jacobian of the chiral transformation, coming from the measure $\mathcal{D}\psi \mathcal{D}\bar{\psi}$.

We shall now apply the same regularization procedure to chiral gauge theories. For simplicity, we will concentrate on a theory of left-handed fermions only, defined by the measure

$$\mathcal{D}\mu = \frac{1}{Z}\mathcal{D}\psi\mathcal{D}\bar{\psi}e^{iS} \ , \tag{3.2.11}$$

where $S = \int \bar{\psi}i\hat{D}\psi$, $i\hat{D} = i\partial\!\!\!/ + eA\!\!\!/P_-$ and $P_\pm = \frac{1}{2}(1 \pm \gamma_5)$. We shall describe two methods that lead to two theories that satisfy different Dyson-Schwinger equations. Method A yields a theory whose current has an anomalous divergence that obeys the Wess-Zumino consistency conditions. Method B leads to a different form of the anomaly that is expressed solely in terms of the field strength $F^{\mu\nu}$, called the covariant anomaly, [12].

METHOD A.

We define the lagrangian as the limit $\lim_{\epsilon\to 0}\mathcal{L}_\epsilon$, where

$$\mathcal{L}_\epsilon = \bar{\psi}i\hat{D}g(-\epsilon\hat{D}^2)\psi \ . \tag{3.2.12}$$

The operator $i\hat{D}$ is not hermitian. Its eigenvalue problem is

$$i\hat{D}\phi_n = \lambda_n\phi_n \ , \ i\hat{D}^\dagger\chi_n = \lambda_n^*\chi_n \ . \tag{3.2.13}$$

We expand ψ and $\bar{\psi}$ in terms of ϕ_n and χ_n respectively,

$$\psi = \sum_n a_n\phi_n \ , \ \bar{\psi} = \sum_n b_n\chi_n^\dagger \ . \tag{3.2.14}$$

Since ϕ_n is orthogonal to χ_m , $\mathcal{D}\psi\mathcal{D}\bar{\psi} = \prod_n da_n db_n$ [13], and the action can be written as $S_\epsilon = \sum_n \lambda_n g(\epsilon\lambda_n^2)a_n b_n$. Under a gauge transformation, $\delta A_\mu = D_\mu\alpha$, $i\hat{D}$ transforms as

$$i\hat{D} \to (1 + ie\alpha P_+)i\hat{D}(1 - ie\alpha P_-) \ . \tag{3.2.15}$$

This can be compensated for by a change in the fermionic fields, $\delta\psi = ie\alpha P_-\psi$ and

$\delta\bar{\psi} = -ie\bar{\psi}\alpha P_+$. One then obtains

$$d\mu \to d\mu \left\{1 - \frac{ie^3}{96\pi^2}\int \mathrm{tr}\alpha\varepsilon_{\mu\nu\rho\sigma}\partial^\mu \left(A^\nu\partial^\rho A^\sigma + \frac{1}{2}A^\nu A^\rho A^\sigma\right)\right\} \ , \tag{3.2.16}$$

which is the consistent form of the anomaly (cf. eq.(3.1.18)). Notice that this vanishes in a theory that obeys the anomaly-free condition, $\mathrm{tr}\{T^a, \{T^b, T^c\}\} = 0$. However, even in that case, the theory is not necessarily gauge-invariant. The reason is that the regulated action is not invariant under the combined transformation of the gauge potential and the fermionic fields. To see this, we transform

$$F = g'(0)\sum_n \epsilon\lambda_n^3 a_n b_n \ . \tag{3.2.17}$$

This is the first-order in ϵ contribution to the regulated action S_ϵ . We find

$$\delta F = g'(0)\sum_{m,n} a_m b_n \epsilon\lambda_m\lambda_n(\lambda_n - \lambda_m)\int \chi_n^\dagger\gamma_5\alpha\phi_m \ . \tag{3.2.18}$$

If we are only interested in $\det i\hat{D} = \int e^{iS}$, we need $\langle\delta F\rangle$, which is zero, because integration over a_n and b_n contracts m and n . The same argument applies to higher order contributions to the regulated action S_ϵ . Thus we see that $\det i\hat{D}$ is gauge-invariant under infinitesimal gauge transformations in an anomaly-free theory. This is also true for all Green functions [14], but it may fail in the case of composite operators. We conjecture that this non-invariance of the action has two consequences. One is the existence of higher-order anomalies found by Atiyah and Singer [6]. The other one is the non-perturbative $SU(2)$ anomaly [15]. Notice that theses anomalies are present even when the anomaly cancellation condition $\mathrm{tr}\{T^a, \{T^b, T^c\}\} = 0$ is satisfied.

METHOD B.

This time let us define the lagrangian as the limit $\lim_{\epsilon\to 0}\mathcal{L}_\epsilon$, where

$$\mathcal{L}_\epsilon = \bar{\psi}i\hat{D}g(-\epsilon\hat{D}^\dagger\hat{D})\psi \ . \tag{3.2.19}$$

The operators $-\hat{D}^\dagger\hat{D}$ and $-\hat{D}\hat{D}^\dagger$ are hermitian. Let $-\hat{D}^\dagger\hat{D}\phi_n = \lambda_n^2\phi_n$, and define $\chi_n = \frac{1}{\lambda_n}i\hat{D}\phi_n$. Then $-\hat{D}\hat{D}^\dagger\chi_n = \lambda_n^2\chi_n$. Expanding ψ and $\bar{\psi}$ in terms of ϕ_n and χ_n

respectively,

$$\psi = \sum_n a_n \phi_n \quad , \quad \bar{\psi} = \sum_n b_n \chi_n^\dagger \ , \tag{3.2.20}$$

the action becomes $S_\epsilon = \sum_n \lambda_n g(\epsilon\lambda_n^2) a_n b_n$, and [9]

$$\mathcal{D}\psi\mathcal{D}\bar{\psi} = \det \chi_n^\dagger(x) \prod_n da_n db_n \det \phi_n(x) \ , \tag{3.2.21}$$

where $\phi_n(x)$ denotes a matrix whose columns and rows are labeled by n and x , and similarly for $\chi_n^\dagger(x)$. The determinants arise because we make a transformation from the basis formed by the eigenfunctions of position to the one formed by the eigenfunctions of $-\hat{D}^\dagger\hat{D}$ and $-\hat{D}\hat{D}^\dagger$. Thus, we originally define the fermionic measure by $\mathcal{D}\psi\mathcal{D}\bar{\psi} = \prod_x d\psi(x) d\bar{\psi}(x)$.

Therefore, the fermionic determinant is

$$\det i\hat{D} = \det \chi_n^\dagger(x) \prod_n \lambda_n g(\epsilon\lambda_n^2) \det \phi_n(x) \ . \tag{3.2.22}$$

Under a gauge transformation, the eigenvalues of $i\hat{D}$ do not change and the eigenfunctions change as follows

$$\phi_n \to (1 + ie\alpha P_-)\phi_n \quad , \quad \chi_n \to (1 - ie\alpha P_-)\chi_n \ . \tag{3.2.23}$$

Therefore,

$$\det i\hat{D} \to \det \left(\lambda_n g(\epsilon\lambda_n^2)\delta_{mn} + \lambda_n g(\epsilon\lambda_n^2) \int \bar{\chi}_n ie\alpha\gamma_5\phi_m \right) \ . \tag{3.2.24}$$

According to the second step of our prescription, we have to expand the change in the determinant in powers of ϵ . Thus, eq.(3.2.24) becomes

$$\det i\hat{D} \rightarrow \det \left(\lambda_n g(\epsilon\lambda_n^2)\delta_{mn} + \sum_{\ell=0}^{\infty} \frac{g^{(\ell)}(0)}{\ell!}(\epsilon\lambda_n^2)^\ell \lambda_n \int \bar{\chi}_m i e \gamma_5 \phi_n \right) \quad ,$$

or

$$\det i\hat{D} \rightarrow \det i\hat{D} \left(1 + \sum_{\ell=0}^{\infty} \frac{g^{(\ell)}(0)}{\ell!}(\epsilon\lambda_n^2)^\ell f(\epsilon\lambda_n^2) \int \bar{\chi}_n i e \alpha \gamma_5 \phi_n \right) \quad . \tag{3.2.25}$$

An explicit calculation shows that [9]

$$\det i\hat{D} \rightarrow \det i\hat{D} \left(1 - \frac{ie^3}{32\pi^2} \int \varepsilon_{\mu\nu\rho\sigma} \mathrm{tr} \alpha F^{\mu\nu} F^{\rho\sigma} \right) \quad , \tag{3.2.26}$$

which is the covariant form of the anomaly. This anomaly does not satisfy the Wess-Zumino conditions, which means that $\det i\hat{D}$ is not a functional of A^μ. One normally concludes that the theory is inconsistent. However, the fermionic determinant is only a sum of vacuum graphs and, therefore, strictly speaking, not a physical quantity. To find the physical content of the theory, one has to compute the Green functions. Since method B is a prescription that regulates all Green functions, the resulting theory is consistent, even though the fermionic determinant is not defined as a functional of A^μ.

If the jacobian of the gauge transformation vanishes, this regularization produces a gauge-invariant theory, which contains no higher-order or non-perturbative anomalies. This can only be true if the theory is different from the one we obtained using method A. This conclusion is supported by the fact that the two methods are not related by counterterms, as the covariant anomaly does not satisfy the Wess-Zumino conditions. Indeed, we shall now show that method B obeys different Dyson-Schwinger equations [10].

Using method B, one can obtain the divergence of the $U(1)$ current $J^\mu = \bar{\psi}\gamma^\mu P_- \psi$,

[9]

$$\partial_\mu J^\mu = \frac{e^2}{32\pi^2}\varepsilon_{\mu\nu\rho\sigma}\mathrm{tr}F^{\mu\nu}F^{\rho\sigma} \ . \tag{3.2.27}$$

The argument is similar to that leading to eq.(3.2.10). If the Dyson-Schwinger equations are obeyed, eq.(3.2.27) to second order in A_μ reads

$$\langle \partial_\mu J_\mu(x) j^a_\nu(y) j^b_\rho(z)\rangle = -\frac{e^2}{8\pi^2}\varepsilon_{\mu\nu\rho\sigma}\delta^{ab}\partial^\mu\delta^4(x-y)\partial^\sigma\delta^4(x-z) \ , \tag{3.2.28}$$

where $j^a_\mu = \frac{\delta S}{\delta A^a_\mu}$. We obtain eq.(3.2.28) by functionally differentiating both sides of eq.(3.2.27) twice with respect to A^μ. Taking fourier transforms of both sides, we obtain

$$G^{ab}_{\nu\rho}(k_1,k_2,k_3) = -\frac{e^2}{8\pi^2}\varepsilon_{\mu\nu\rho\sigma}\delta^{ab}k_1^\mu k_2^\sigma \ , \tag{3.2.29}$$

where $\delta^4(k_1+k_2+k_3)G^{ab}_{\nu\rho}(k_1,k_2,k_3)$ is the fourier transform of $\langle\partial^\mu J_\mu(x)j^a_\nu(y)j^b_\rho(z)\rangle$.

An explicit computation of the three-point function $\langle\partial^\mu J_\mu(x)j^a_\nu(y)j^b_\rho(z)\rangle$, making use of the regulated propagator $\not{p}^{-1}f(\epsilon p^2)$, which is obtained from the regularized lagrangian (eq.(3.2.19)), gives

$$G^{ab}_{\nu\rho}(k_1,k_2,k_3) = -\frac{e^2}{24\pi^2}\varepsilon_{\mu\nu\rho\sigma}\delta^{ab}k_1^\mu k_2^\sigma \ . \tag{3.2.30}$$

This disagrees with eq.(3.2.27) by a factor of $\frac{1}{3}$. It follows that the resulting theory does not satisfy the Dyson-Schwinger equations and is therefore not equivalent to the one obtained using method A. The remarkable fact is that this is true even in the case where all anomalies cancel, e.g., in an $SU(2)$ gauge theory [9]. Nevertheless, it may still be interesting to investigate the physical content of this theory.

3.3 A chiral Schwinger model

We shall now discuss a procedure that leads to gauge-invariant chiral theories [10]. We start by applying the procedure to a two-dimensional abelian theory (chiral Schwinger model), where we can compute the exact form of the generating functional [16]. In the following sections (3.4 and 3.5) we discuss the four-dimensional case and bosonic strings.

Our regularization procedure is as follows. First, we integrate over the fermionic degrees of freedom, and then over the gauge potential. Thus initially, we have to work with a theory of fermions in a background gauge field. We regularize the action by splitting points in such a way as to preserve gauge invariance. In a theory with vector couplings, chiral symmetry is also preserved. For definiteness, we choose to work with a gauge field coupled only to left-handed fermions. The regularized action for the fermions is

$$S_\epsilon = \int d^2x \bar{\psi}_L(x + \frac{1}{2}\epsilon)\mathcal{P}e^{ie\int_{x+\epsilon/2}^{x} A\cdot dl} i\not{D}\mathcal{P}e^{ie\int_{x}^{x-\epsilon/2} A\cdot dl}\psi_L(x - \frac{1}{2}\epsilon) \quad , \qquad (3.3.1)$$

where $iD_\mu = i\partial_\mu + eA_\mu$. This is a symmetric point-splitting, in the sense that, in the end, we have to average over directions of ϵ^μ. The effect of the regulator is to add new vertices that are formally of order ϵ. Moreover, both the propagator and the vertices contain a factor of $e^{i\epsilon\cdot p}$. These factors cancel in a diagram and so diagrams are still not finite. To remedy this, we supplement our procedure with an additional step. Before we perform any loop integrations, we expand all factors that depend on the external momenta in powers of ϵ^μ , and then truncate the series after the first-order term ($e^{i\epsilon\cdot p}$ is replaced by $1 + i\epsilon \cdot p$, where p^μ is an external momentum). The factors $e^{i\epsilon\cdot q}$, where q^μ is the loop momentum, are not expanded, thus regulating the diagram. To make sense, this truncation has to be done consistently. To this end, we write the integrand as a sum of terms, each of which is a function of q^μ with only one pole. Then in each term we shift the integration variable q^μ so as to move the pole to zero. Finally, we expand the factors that involve the external momenta and not q^μ , in the manner described above, before performing any loop integrations. It should be noted that this procedure leads to a redefinition of Feynman diagrams and is not merely the addition of counterterms in the action. We therefore obtain a new theory, satisfying different Dyson-Schwinger equations, as we shall demonstrate. We shall show that this new theory is consistent.

The interaction lagrangian is

$$\mathcal{L}_{int} = eA_\mu \left\{ \bar{\psi}\gamma^\mu P_-(1 + ie\epsilon_\nu A^\nu)\psi - \frac{1}{2}\bar{\psi}(\epsilon^\mu\gamma^\nu + \epsilon^\nu\gamma^\mu)\overleftrightarrow{\partial}_\nu P_-\psi \right\} + o(\epsilon^2) \quad . \qquad (3.3.2)$$

The higher-order terms in ϵ^μ are irrelevant because we truncate all operators after the first-order term.

We now restrict ourselves to two dimensions, where we can compute all the Green functions exactly. We are interested in Green functions that contain insertions of the gauge-invariant operator $j^\mu = \bar{\psi}\gamma^\mu P_-\psi$. It should be pointed out that the gauge potential couples to the current

$$j^\mu_\epsilon = \bar{\psi}\gamma^\mu P_-\psi - \frac{1}{2}\bar{\psi}(\epsilon^\mu\gamma^\nu + \epsilon^\nu\gamma^\mu)\overleftrightarrow{D}_\nu P_-\psi \quad . \tag{3.3.3}$$

However, we are not considering the operator j^μ_ϵ , because Green functions containing insertions of this operator are trivial. As we shall show, $\langle j^\mu_\epsilon\rangle = 0$, in the presence of an external gauge field A_μ . Thus, we introduce sources B_μ and J^μ coupled to j^μ and A_μ , respectively. We wish to calculate the generating functional

$$e^{iW_L[J,B]} = \int \mathcal{D}A_\mu e^{i\int d^2x(\frac{1}{2}F^2+eJ^\mu A_\mu)} \int \mathcal{D}\psi_L\mathcal{D}\bar{\psi}_L e^{iS_\epsilon + ie\int j^\mu B_\mu} \quad . \tag{3.3.4}$$

$W_L[J,B]$ describes a theory that does not obey the naïve Dyson-Schwinger equations, if j^μ is defined to be the electromagnetic current. For convenience, we introduce light-cone coordinates $x_\pm = \frac{1}{\sqrt{2}}(x_0 \pm x_1)$. It is easy to see that $(\gamma_\pm)^2 = 0$ and $P_- = \frac{1}{2}\gamma_+\gamma_-$. Notice that j^μ is actually a one-component object ($j_+ = 0$, because $\gamma_+P_- = 0$). Therefore, the external source B_μ also has only one component, namely B_+ .

We first integrate over the fermions. Let Z_f be the fermionic contribution,

$$Z_f[A_\mu, B_\mu] = \int \mathcal{D}\psi\mathcal{D}\bar{\psi} e^{iS_f + ie\int j^\mu B_\mu} \quad , \tag{3.3.5}$$

where $S_f = \int[\bar{\psi}(x+\frac{1}{2}\epsilon)i\not{\partial}\psi(x-\frac{1}{2}\epsilon)+\mathcal{L}_{int}]$ is the fermionic action. We have multiplied the generating functional by a constant, $\int \mathcal{D}\psi_R\mathcal{D}\bar{\psi}_R e^{i\int\bar{\psi}(x+\frac{1}{2}\epsilon)i\not{\partial}P_+\psi(x-\frac{1}{2}\epsilon)}$, to define a propagator for the fermions. The propagator is $\Delta(x) = \int\frac{d^2p}{(2\pi)^2}e^{ipx}\frac{e^{ip\epsilon}}{\not{p}}$. Differentiating

Z_f with respect to the coupling constant e , we obtain

$$\begin{aligned}\frac{1}{Z_f}\frac{\partial Z_f}{\partial e} &= i\int d^2x\{A_\mu(x)\langle j_\epsilon^\mu(x)\rangle + B_\mu(x)\langle j^\mu(x)\rangle\} \\ &= i\int d^2x\Big\{[A_\mu(x)+B_\mu(x)]\langle j^\mu(x)\rangle + 2ie\epsilon^\nu A_\mu(x)A_\nu(x)\langle j^\mu(x)\rangle \\ &- \frac{1}{2}A_\mu(x)\langle\bar{\psi}(x)(\epsilon^\mu\gamma^\nu+\epsilon^\nu\gamma^\mu)\overleftrightarrow{\partial}_\nu P_-\psi(x)\rangle\Big\} \quad .\end{aligned} \tag{3.3.6}$$

Using perturbation theory, we see that only the bilinear terms in A_μ and B_μ survive. Thus, eq.(3.3.6) reads

$$\begin{aligned}\frac{1}{Z_f}\frac{\partial Z_f}{\partial e} &= -e\int d^2x d^2y\Big\{(A_\mu(x)+B_\mu(x))(A_\nu(y)+B_\nu(y))\langle j^\mu(x)j^\nu(y)\rangle_0 \\ &+ 2\epsilon^\nu A_\mu(x)A_\nu(y)\langle j^\mu(x)\rangle_0\delta^2(x-y) \\ &- A_\mu(x)(A_\nu(y)+B_\nu(y))\langle\bar{\psi}(x)(\epsilon^\mu\gamma^\lambda+\epsilon^\lambda\gamma^\mu)\overleftrightarrow{\partial}_\lambda P_-\psi(x)j^\nu(y)\rangle_0\Big\} \quad ,\end{aligned} \tag{3.3.7}$$

where $\langle\Theta\rangle_0$ denotes the expectation value of the operator Θ with vanishing background gauge fields ($A_\mu = B_\mu = 0$) . To calculate $\partial \ln Z_f/\partial e$, we have to compute the following Green functions,

$$G_{\mu\nu}^{(1)}(p) = \int d^2x e^{-ipx}\langle j_\mu(x)j_\nu(0)\rangle_0 \quad , \tag{3.3.8a}$$

$$G_{\mu\nu}^{(2)}(p) = \int d^2x e^{-ipx}2\epsilon_\nu\langle j_\mu(x)\rangle_0\delta^2(x) \quad , \tag{3.3.8b}$$

$$G_{\mu\nu}^{(3)}(p) = \int d^2x e^{-ipx}\epsilon_\mu\langle\bar{\psi}(x)\gamma^\rho\overleftrightarrow{\partial}_\rho P_-\psi(x)j_\nu(0)\rangle_0 \quad , \tag{3.3.8c}$$

$$G_{\mu\nu}^{(4)}(p) = \int d^2x e^{-ipx}\epsilon^\lambda\langle\bar{\psi}(x)\gamma_\mu\overleftrightarrow{\partial}_\lambda P_-\psi(x)j_\nu(0)\rangle_0 \quad . \tag{3.3.8d}$$

With these definitions, eq.(3.3.7) can be written as

$$\begin{aligned}\frac{1}{Z_f}\frac{\partial Z_f}{\partial e} = -e\int\frac{d^2p}{(2\pi)^2}&\Big\{(\tilde{A}^\mu(p)+\tilde{B}^\mu(p))(\tilde{A}^\nu(-p)+\tilde{B}^\nu(-p))G_{\mu\nu}^{(1)}(p) \\ &+ \tilde{A}^\mu(p)\tilde{A}^\nu(-p)G_{\mu\nu}^{(2)}(p) \\ &-\tilde{A}^\mu(p)(\tilde{A}^\nu(-p)+\tilde{B}^\nu(-p))(G_{\mu\nu}^{(3)}+G_{\mu\nu}^{(4)}(p))\Big\} \quad ,\end{aligned} \tag{3.3.9}$$

where $\tilde{A}_\mu(\tilde{B}_\mu)$ is the fourier transform of $A_\mu(B_\mu)$. $G_{\mu\nu}^{(1)}$ is the ordinary two-point function and we easily find $G^{(1)}_{++} = G^{(1)}_{+-} = G^{(1)}_{-+} = 0$ and

$$G^{(1)}_{--} = -\frac{i}{4\pi^2}\frac{p_-}{p_+} \quad , \tag{3.3.10}$$

which shows that the current j^μ has an anomalous divergence $(p_+ G^{(1)}_{--} = -\frac{i}{4\pi^2}p_-)$. However, this does not imply the breakdown of gauge invariance, because $\partial_\mu j^\mu$ does not represent the variation of the fermionic part of the path integral.

It is easy to show that $G^{(2)}_{++} = G^{(2)}_{--} = G^{(2)}_{-+} = 0$ and

$$G^{(2)}_{+-} = -\frac{i}{4\pi^2} \quad . \tag{3.3.11}$$

The calculation of $G^{(3)}_{\mu\nu}$ is also straightforward. We find $G^{(3)}_{++} = G^{(3)}_{-+} = 0$. Also,

$$G^{(3)}_{+-} = \epsilon_+ \int \frac{d^2q}{(2\pi)^2} e^{iq\epsilon} q_- \frac{(p+q)_-}{(p+q)^2 + i\eta} \frac{(p-q)_-}{(p-q)^2 + i\eta} \quad , \tag{3.3.12}$$

where the limit $\eta \longrightarrow 0$ is implied. We write

$$G^{(3)}_{+-} = -\epsilon_+ \int \frac{d^2q}{(2\pi)^2} e^{iq\epsilon} \frac{1}{2} \left(\frac{(p+q)_-}{(p+q)^2 + i\eta} - \frac{(p-q)_-}{(p-q)^2 + i\eta} \right) \quad . \tag{3.3.13}$$

Shifting the integration variable to move the poles to zero, we obtain

$$G^{(3)}_{+-} = \frac{1}{2}\epsilon_+ \int \frac{d^2q}{(2\pi)^2} e^{iq\epsilon} (e^{-ip\epsilon} + e^{ip\epsilon}) \frac{q_-}{q^2 + i\eta} \quad . \tag{3.3.14}$$

According to our procedure, the next step is to drop the factor $\frac{1}{2}(e^{-ip\epsilon} + e^{ip\epsilon})$, which depends on the external momentum p^μ . It should be noted, however, that this does not modify the definition of the Green function, because the integrand has only a simple pole in q^μ . The final answer is

$$G^{(3)}_{+-}(p) = \frac{i}{4\pi^2} \quad . \tag{3.3.15}$$

Similarly, we obtain $G^{(3)}_{--} = 0$.

Finally, we have to compute $G^{(4)}_{\mu\nu}$. Plainly, $G^{(4)}_{++} = G^{(4)}_{-+} = G^{(4)}_{+-} = 0$. Now,

$$G^{(4)}_{--} = \epsilon^{\lambda} \int \frac{d^2q}{(2\pi)^2} e^{iq\epsilon} q_{\lambda} \frac{(p+q)_-}{(p+q)^2 + i\eta} \frac{(p-q)_-}{(p-q)^2 + i\eta} \quad . \tag{3.3.16}$$

As before, we separate the poles, obtaining

$$G^{(4)}_{--} = -\epsilon^{\lambda} \int \frac{d^2q}{(2\pi)^2} e^{iq\epsilon} q_{\lambda} \frac{1}{2p_-} \left\{ \frac{(p+q)_-}{(p+q)^2 + i\eta} - \frac{(p-q)_-}{(p-q)^2 + i\eta} \right\} \quad . \tag{3.3.17}$$

The right-hand side consists of two terms. The first term is

$$I_1 = -\epsilon^{\lambda} \int \frac{d^2q}{(2\pi)^2} e^{iq\epsilon} q_{\lambda} \frac{1}{2p_-} \frac{(p+q)_-}{(p+q)^2 + i\eta} \quad , \tag{3.3.18a}$$

or

$$I_1 = -\epsilon^{\lambda} \frac{1}{2p_-} \int \frac{d^2q}{(2\pi)^2} e^{iq\epsilon} e^{-ip\epsilon} \frac{q_-}{q^2 + i\eta} (q_{\lambda} - p_{\lambda}) \quad , \tag{3.3.18b}$$

where we shifted the integration variable to derive eq.(3.3.18b). We now drop the factor $e^{-ip\epsilon}$. Unlike before, however, this modification changes the Green function, because the integrand has a second-order pole. After discarding $e^{-ip\epsilon}$, it can be seen that only a simple pole remains and so gauge invariance is retained, as we shall see shortly. Integration over q^{μ} gives

$$I_1 = \frac{i}{8\pi^2} \frac{p_-}{p_+} + R \quad , \tag{3.3.19}$$

where R vanishes after averaging over directions of ϵ^{μ} . The second term in eq.(3.3.17) is computed similarly and the final result is

$$G^{(4)}_{--} = \frac{i}{4\pi^2} \frac{p_-}{p_+} \quad . \tag{3.3.20}$$

Because of eqs. (3.3.10,11,15,20), eq. (3.3.7) becomes

$$\frac{1}{Z_f} \frac{\partial Z_f}{\partial e} = ie \int \frac{d^2p}{(2\pi)^2} \left\{ \frac{1}{\pi} \frac{p_-}{p_+} \tilde{B}_+(p) \tilde{B}_+(-p) + \frac{1}{\pi} \tilde{B}_+(p) (\tilde{A}_-(-p) - \frac{p_-}{p_+} \tilde{A}_+(p)) \right\} \quad . \tag{3.3.21}$$

Integrating, we obtain

$$Z_f = e^{\frac{1}{2}ie^2 \int \frac{d^2p}{(2\pi)^2} \left\{ \frac{1}{\pi}\frac{p_-}{p_+}\tilde{B}_+(p)\tilde{B}_+(-p) + \frac{1}{\pi}\tilde{B}_+(p)(\tilde{A}_-(-p) - \frac{p_-}{p_+}\tilde{A}_+(p)) \right\}} \quad . \tag{3.3.22}$$

which is manifestly gauge-invariant (under the transformation $\delta\tilde{A}_\mu = p_\mu\alpha, \delta\tilde{B}_\mu = 0$). It is interesting to note that when $B_\mu = 0$, $Z_f = 1$, which shows that the fermion determinant has no dependence on the gauge potential A_μ . In particular, there is no quadratic piece in the effective action that could give a mass to A_μ . It also follows that $\langle j_\epsilon^\mu \rangle = 0$, because $\langle j_\epsilon^\mu \rangle = \frac{\delta \ln Z_f}{\delta A_\mu}\Big|_{B_\mu=0}$.

Now only integrating over the gauge field remains. To do so, we have to fix the gauge. A convenient choice is the axial gauge $A_+ = 0$. The Faddeev-Popov determinant does not depend on A_μ , so we can integrate over A_- by completing the square in the exponential in eq.(3.3.4). Our final result is

$$\begin{aligned} W_L[J,B] = \frac{1}{2}\int \frac{d^2p}{(2\pi)^2} \Bigg\{ & \frac{m^2}{p_+^2}\tilde{J}_+(p)\tilde{J}_+(-p) \\ & + (m^2 - p^2)\frac{m^2}{p_+^2}\tilde{B}_+(p)\tilde{B}_+(-p) + \frac{2m^3}{p_+^2}\tilde{J}_+(p)\tilde{B}_+(-p) \Bigg\} \quad , \end{aligned} \tag{3.3.23}$$

where $m = e/\sqrt{\pi}$. This is an uninteresting theory except for the fact that it exists. The only pole with non-vanishing residue appears in the term that is quadratic in B_+. We therefore conclude that the model described by W_L contains just a massless mode that does not couple to the gauge field, but only couples to the electromagnetic current.

To extend our results to the vector case, we have to add an interaction between the vector potential and the right-handed current. This just adds the right-handed generating functional W_R, which is computed similarly, to W_L. which is computed similarly. The resulting generating functional for the vector theory is

$$\begin{aligned} W = W_L + W_R = \frac{1}{2}\int \frac{d^2p}{(2\pi)^2} \Bigg\{ & \frac{m^2}{p^2}\tilde{J}_\mu\pi^{\mu\nu}\tilde{J}_\nu \\ & + (m^2 - p^2)\frac{m^2}{p^2}\tilde{B}_\mu\pi^{\mu\nu}\tilde{B}_\nu + \frac{2m^3}{p^2}\tilde{J}_\mu\pi^{\mu\nu}\tilde{B}_\nu \Bigg\} \quad , \end{aligned} \tag{3.3.24}$$

where

$$\pi_{\mu\nu}(p) = g_{\mu\nu} - \frac{p_\mu p_\nu}{p^2} \tag{3.3.25}$$

is a projection operator. The current J^μ is conserved: $p_\mu \tilde{J}^\mu = 0$. The theory described by W (eq.(3.3.24)) does not contain a massive state, in disagreement with Schwinger's result [1]. We can obtain the Schwinger model by performing a non-local transformation whose effect is to multiply J_μ and B_μ by a factor that depends on momentum. Thus, we see that the two models are related by non-local counter terms, yet they are both unitary. This comes about in a trivial way in two dimensions, because there is no physical state coupled to the gauge field. However, in higher dimensions this method may provide a way of canceling the anomaly by adding non-local counter terms, yet retaining unitarity and locality. The factor needed to multiply the sources J_μ and B_μ is uniquely determined, if we require that the naïve Dyson-Schwinger equations be satisfied. This implies that the factor is $\alpha^{\mu\nu}(p) = g^{\mu\nu}\frac{|\mathrm{p}|}{\sqrt{p^2-m^2}}$, as we shall now demonstrate.

The Dyson-Schwinger equations are generalizations of the equations of motion, $\partial_\mu F^{\mu\nu} = ej^\nu + eJ^\nu$. A straightforward calculation shows that in momentum space,

$$\frac{1}{e} p^2 \pi^{\mu\nu} \frac{\delta W}{\delta \tilde{J}_\nu(p)} + \frac{\delta W}{\delta \tilde{B}_\mu(p)} + e\tilde{J}_\mu(p) = 0 \ . \tag{3.3.26}$$

Using eq.(3.3.24) with $\tilde{J}_\mu$ and $\tilde{B}_\mu$ replaced by $\alpha_{\mu\nu}\tilde{J}^\nu$ and $\alpha_{\mu\nu}\tilde{B}^\nu$, respectively, we can evaluate $\alpha^{\mu\nu}$. An easy way to obtain $\alpha^{\mu\nu}$ is by first differentiating eq.(3.3.20) with respect to $\tilde{B}_\rho$. A direct computation yields

$$\alpha^{\mu\nu}(p) = g^{\mu\nu} \frac{|\mathrm{p}|}{\sqrt{p^2 - e^2}} \quad , \tag{3.3.27}$$

as promised. Therefore, W becomes

$$\tilde{W} = \frac{1}{2} \int \frac{d^2p}{(2\pi)^2} \left\{ \frac{m^2}{p^2 - m^2} \tilde{J}_\mu \pi^{\mu\nu} \tilde{J}_\nu + m^2 \tilde{B}_\mu \pi^{\mu\nu} \tilde{B}_\nu + \frac{2m^3}{p^2 - m^2} \tilde{J}_\mu \pi^{\mu\nu} \tilde{B}_\nu \right\} \quad , \tag{3.3.28}$$

which shows the existence of a massive state of mass $m = e/\sqrt{\pi}$, in agreement with Schwinger's result. Notice that we cannot satisfy the Dyson-Schwinger equations when we only have left-handed fermions, because the current j_μ is not conserved. (Whereas J_μ is conserved and $\partial_\mu \partial_\nu F^{\mu\nu} = 0$ identically.) This is the price we have to pay if we insist on maintaining gauge invariance.

The above method is readily applicable to non-abelian groups. However, in this case, we cannot solve the model exactly. The model fails to satisfy the naïve Dyson-Schwinger equations, as can be seen by making use of the Maxwell equations to express the current in terms of the gauge field and then compute, e.g., the expectation value of the product of two currents,

$$G^{ab}_{\mu\nu}(x,y) = \langle j^a_\mu(x) j^b_\nu(y) \rangle \quad . \tag{3.3.29}$$

The procedure is the same as in the abelian case. By computing $G^{ab}_{\mu\nu}$ perturbatively, we can deduce the existence of a massless mode that does not couple to the gauge field. However, non-perturbative effects may give rise to a massive state. More work is also needed to establish the unitarity of the theory to all orders in the coupling constant.

At first glance, our results seem to be in conflict with the recent solution of the chiral Schwinger model [17] that gave rise to massive states. In that model the mass depends on a parameter a that is the coefficient of a term in the action that explicitly breaks gauge invariance. Effectively, if $a < 1$, an extra field is added of opposite statistics to the fermions. In particular, when $a = 0$ (i.e., when the gauge field is massless), negative norm states are obtained [17]. However, in our case, although we obtain a massless physical state, the theory is unitary, owing to the fact that we have not introduced additional degrees of freedom. This leads to a consistent theory, because there is no physical state coupled to the gauge field.

Finally, it should be noted that the method is applicable to a higher number of dimensions. This is the subject of the next section.

3.4 Four dimensions

We now turn to a discussion of a four-dimensional chiral gauge theory [10]. Unfortunately, explicit results are hard to obtain in this case. The reason is that separating the poles in a four-dimensional integral is not a straightforward process. We shall only be able to check for the absence of an anomalous divergence of the current and of an anomalous term in the commutator of the generators of gauge transformations. A lot more work seems to be needed for the study of the physical spectrum. Also, questions regarding unitarity will remain unanswered. Again for simplicity, we assume that the theory only contains left-handed fermions. We define $\mathcal{L}$ as the limit $\lim_{\epsilon\to 0}\mathcal{L}_\epsilon$, where

$$\mathcal{L}_\epsilon = \bar{\psi}(x+\frac{1}{2}\epsilon)i\not{\partial}\psi(x-\frac{1}{2}\epsilon)+eA_\mu^a(x)j_\epsilon^{\mu a}(x) \ , \tag{3.4.1}$$

and

$$j_\epsilon^{\mu a} = \bar{\psi}(x)\left(T^a\gamma^\mu - \frac{1}{2}T^a(\epsilon^\nu\gamma^\mu+\epsilon^\mu\gamma^\nu)\overleftrightarrow{\partial}_\nu + \frac{1}{2}ie\epsilon^\nu\{T^a,A_\nu(x)\}\gamma^\mu\right)P_-\psi(x) \ . \tag{3.4.2}$$

In analogy with the two-dimensional case, the current is expected to have vanishing divergence:

$$D_\mu j_\epsilon^{\mu a} = 0 \ . \tag{3.4.3}$$

To see how this comes about, let us define the current $j^{\mu a}$ as the zeroth order term in $j_\epsilon^{\mu a}$. By making repeated use of the Dirac equation, we deduce that

$$D_\mu j_\epsilon^{\mu a} = D_\mu j^{\mu a} - ie\epsilon_\nu\bar{\psi}\gamma_\mu P_- F^{\mu\nu}\psi \ . \tag{3.4.4}$$

It is easy to show that

$$\langle j_\mu^a\rangle \equiv \langle\bar{\psi}\gamma_\mu T^a P_-\psi\rangle = \frac{e^2}{8\pi^2}\epsilon_\lambda\varepsilon^{\mu\lambda\rho\sigma}F_{\rho\sigma} \ , \tag{3.4.5}$$

where we neglect higher-order terms in ϵ. Therefore, eq.(3.4.3) holds, provided

$$D_\mu j^{\mu a} = \frac{e^2}{32\pi^2}\varepsilon_{\mu\nu\rho\sigma}\mathrm{tr}T^a F^{\mu\nu}F^{\rho\sigma} \ . \tag{3.4.6}$$

Notice that this result agrees with the second method of Section 3.2. It disagrees with the result

$$D_\mu j^{\mu a} = \frac{e^2}{96\pi^2}\varepsilon_{\mu\nu\rho\sigma}\mathrm{tr}T^a\partial^\mu(A^\nu\partial^\rho A^\sigma + \frac{1}{2}A^\nu A^\rho A^\sigma) \ , \tag{3.4.7}$$

which is obtained in conventional perturbation theory, or by using the first method of Section 3.2. Although this way we do not obtain the consistent anomaly, the Wess-Zumino conditions are still satisfied by the Noether current of the theory, $j^a_{\mu\epsilon} = \frac{\delta}{\delta A^a_\mu}S_\epsilon$, which has no anomaly, i.e., $D_\mu j^{\mu a}_\epsilon = 0$. Eq.(3.4.6) can be proven by a straightforward computation of diagrams. This shows that the theory is gauge-invariant.

It is interesting to compare this calculation to the conventional one which parallels Bardeen [18]. Let the S-matrix be $S = Te^{i\int d^4x e A^a_\mu J^{\mu a}}$, where

$$J^{\mu a}(x) = e\bar{\psi}(x+\frac{1}{2}\epsilon)\left(T^a\gamma^\mu - \frac{1}{2}\epsilon^\nu\overleftrightarrow{\partial} + \frac{1}{2}ie\epsilon^\nu\{A_\nu, T^a\}\gamma^\mu\right)\psi(x-\frac{1}{2}\epsilon) \tag{3.4.8}$$

is the Noether current of the theory. We have introduced extra terms in the lagrangian that are formally of order ϵ , to avoid having to introduce counterterms at the end. The difference with the previous method is that in calculating Green functions we do not expand the fermionic fields in powers of ϵ . Also the propagator is now $\not{p}$ instead of $\not{p}e^{i\epsilon p}$. Following Bardeen, we can show that $J^{\mu a}$ satisfies eq.(3.4.7). Comparing the two methods, it is obvious that the discrepancy arises from the difference in the expansion in ϵ , before taking the limit $\epsilon \to 0$.

We can also see how the above procedure modifies the Maxwell equations, so as not to reproduce any inconsistencies. The argument showing the inconsistency of theories with anomalies is as follows [2]. The equations of motion for a dynamical gauge field are $D_\mu F^{\mu\nu} = ej^\nu$. Now $D_\mu D_\nu F^{\mu\nu} = 0$, identically and so $D_\mu j^\mu = 0$. But $D_\mu j^\mu \neq 0$, by eq.(3.4.7). Hence the inconsistency.

However, after regularizing, using our gauge invariant procedure, the equations of motion become

$$D_\mu F^{\mu\nu} = ej^\nu_\epsilon \ , \tag{3.4.9}$$

where j^μ_ϵ is the Noether current of the theory, given by eq.(3.4.2). Since $D_\mu j^\mu_\epsilon = 0$ (cf eq.(3.4.3)), there is no inconsistency.

As our procedure gives the covariant anomaly like method B of Section 3.2, the resulting theory satisfies different Dyson-Schwinger equations. Unlike that method, however, the anomaly we find using our gauge invariant procedure does not break gauge invariance, because it is not associated with the Noether current of the theory. Therefore, this procedure produces a theory that is not equivalent to the theory that resulted from method B of Section 3.2.

Next we ask the question whether Gauss's law can be implemented as a first-order constraint. Using cohomology, Faddeev [19] showed that this is not the case for theories with anomalies. Specifically, he showed that there is an anomalous term contributing to the commutator of two generators of the gauge group. Subsequently, various methods were employed for the explicit evaluation of the anomalous term, with conflicting results [20]. Since our regularization is gauge invariant, there are no such terms. In fact the generators vanish by virtue of the equations of motion. Yet, it might be of interest to investigate the possibility that the time component of $j^a_\mu = \bar{\psi}(x)\gamma_\mu P_- T^a \psi(x)$ is a generator of gauge transformations. We shall see that this is true, although in conventional perturbation theory there is an anomalous term in the commutator of two currents.

We therefore wish to compute the equal-time commutator of the time components of two currents $j^a_0(x)$ and $j^b_0(0)$, where $j^a_\mu = \bar{\psi}\gamma_\mu P_- T^a \psi$. We define a quantity $C^{ab}(\mathrm{x})$ by

$$C^{ab}(\mathrm{x}) = [j^a_0(\mathrm{x},0), j^b_0(0)] - if^{abc} j^c_0(0)\delta^3(\mathrm{x}) \ . \tag{3.4.10}$$

If one uses the method of Section 3.2 which produces the consistent anomaly, or cohomology [19], one finds that

$$C^{ab}(\mathrm{x}) = \frac{1}{24\pi^2} \mathrm{tr} T^a \{T^b, T^c\} \varepsilon^{ijk} \partial_i A^c_j(\mathrm{x}) \partial_k \delta^3(\mathrm{x}) \ . \tag{3.4.11}$$

Therefore j^a_0 cannot be a generator of infinitesimal gauge transformations.

We shall use gauge-invariant point-splitting regularization to show that $C = 0$, as naïvely expected. The Fourier transform of C is

$$\tilde{C}^{ab}(\mathbf{p}) = \int d^3x e^{i\mathbf{p}\cdot\mathbf{x}}[j_0^a(\mathbf{x},0), j_0^b(0)] - if^{abc}j_0^c(0) \ . \tag{3.4.12}$$

The BJL theorem [21] gives

$$\int d^3x e^{i\mathbf{p}\cdot\mathbf{x}}[j_0^a(\mathbf{x},0), j_0^b(0)] = \lim_{p^0\to\infty} p^0 \int d^4x e^{ipx} Tj_0^a(x)j_0^b(0) \ . \tag{3.4.13}$$

We shall calculate

$$\sigma^{ab}(\mathbf{p}) = \langle 0|\tilde{C}^{ab}(\mathbf{p})|0\rangle \tag{3.4.14}$$

to first order in A_μ . Only connected diagrams contribute and, to use the BJL theorem, we have to drop all polynomials in p^0 . Because of (3.4.12) and (3.4.13) , eq.(3.4.14) becomes

$$\sigma^{ab}(\mathbf{p}) = \lim_{p^0\to\infty} p^0 \int d^4x e^{ipx}\langle Tj_0^a(x)j_0^b(0)\rangle - if^{abc}\langle j_0^c(0)\rangle \ . \tag{3.4.15}$$

We observe that, on the right-hand side of eq.(3.4.15), the second term serves to remove the part of the first term that does not depend on p . Let F_1, F_2 be the p-dependent parts of the diagrams that contribute to $\langle Tj_0j_0\rangle$ and contain the vertices that come from the first and second terms contributing to the current $j_\epsilon^{\mu a}$ (eq.(3.4.2)), respectively. Then

$$\sigma = F_1 + F_2 \ . \tag{3.4.16}$$

An explicit calculation shows that

$$F_1 = \frac{1}{12\pi^2}\int \frac{d^4q}{(2\pi)^4}\mathrm{tr}T^a\{T^b, T^c\}\varepsilon^{ijk}p_iq_j\tilde{A}_k^c(q) \ , \tag{3.4.17}$$

where $\tilde{A}_\mu$ is the Fourier transform of A_μ . It is easy to see that $F_2 = -F_1$. Therefore, eq.(3.4.16) becomes

$$\sigma = 0 \ , \tag{3.4.18}$$

as advertised.

It is interesting to compare this calculation with the one that makes use of the current defined in eq.(3.4.8). We can follow the same steps as above to compute $\sigma^{ab}(\mathrm{p})$. We see that the second vertex does not contain the second term that is proportional to $\epsilon^\mu\gamma^\nu$. Thus, we find that in this case $F_2 = -\frac{1}{2}F_1$. It follows that $\sigma = \frac{1}{2}F_1$, or

$$\sigma^{ab}(\mathrm{p}) = \frac{1}{24\pi^2}\int\frac{d^4q}{(2\pi)^4}\mathrm{tr}T^a\{T^b,T^c\}\varepsilon^{ijk}p_iq_j\tilde{A}_k^c(q) \ , \tag{3.4.19}$$

proving the existence of an anomalous term in the commutator of two currents.

It is interesting to observe that we obtained two conflicting results by considering essentially the same diagrams. The only difference between the two regularization procedures was in the expansion in powers of ϵ .

3.5 Bosonic strings

We shall now apply the same method to bosonic strings. We shall find that it is possible to obtain a unitary theory in an arbitrary number of space-time dimensions. However, we shall not be able to construct vertex operators explicitly. To this end, we may have to consider the ghost coordinates as well.

We start with the action‡

$$S = \frac{1}{2\alpha'}\int\sqrt{-\det g^{ab}}g^{ab}\partial_a x^\mu\partial_b x_\mu d^2\sigma \quad , \tag{3.5.1}$$

where g_{ab} and x^μ are independent degrees of freedom, (g^{ab} is the world-sheet metric and x^μ are the coordinates ($\mu = 1,\dots,d$, d being the dimension of space-time)) and α' is an arbitrary constant that can be identified with the string tension. Define a regulated action S_ϵ by symmetrically splitting points [22]:

$$S_\epsilon = \frac{1}{2\alpha'}\int\sqrt{g}g^{ab}\partial_a x^{\mu(+)}\partial_b x_\mu^{(-)}d^2\sigma \quad , \tag{3.5.2}$$

where ${x_\mu}^{(\pm)} = x_\mu(\sigma \pm \frac{1}{2}\epsilon)$ and ϵ^a is a fixed two-vector. This is a symmetric point-splitting in the sense that, before taking the limit $\epsilon \to 0$, we have to average over the

‡ For a comprehensive introduction to string theory and references on the subject, see ref.[24].

directions of ϵ^a.

To work with perturbation theory, we set

$$g_{ab} = \delta_{ab} + \alpha' h_{ab} \quad , \tag{3.5.3}$$

where h_{ab} is small. The linearized action is

$$S_\epsilon = \frac{1}{2} \int \left(\partial_a x^{\mu(+)} \partial_a x_\mu^{(-)} + \alpha' h_{ab} T_{ab} \right) d^2\sigma \quad , \tag{3.5.4}$$

where $T_{ab} = \partial_a x^{\mu(+)} \partial_b x_\mu^{(-)} - \frac{1}{2}\delta_{ab}\partial_c x^{\mu(+)}\partial_c x_\mu^{(-)}$ and we have rescaled $x^\mu \to \sqrt{\alpha'} x^\mu$. The zeroth-order action is

$$S_0 = \frac{1}{2} \int d^2\sigma \partial_a x^{\mu(+)} \partial_a x_\mu^{(-)} \quad , \tag{3.5.5}$$

giving a propagator $\delta(\sigma) = \int \frac{d^2p}{(2\pi)^2} e^{ip\sigma} e^{ip\epsilon} \frac{1}{p^2}$, whose short-distance behavior is $\delta(\sigma) \sim -\ln \sigma^2$.

We can now describe how we will be calculating correlation functions. If the correlator involves an operator that depends on the regularization parameter ϵ^a, we expand the operator in ϵ^a, keeping only the terms in the series that are of order at most 2 in ϵ^a. Thus, e.g., the stress tensor becomes

$$T_{ab} = \partial_a x^\mu \partial_b x_\mu + \frac{1}{4} \epsilon_c \epsilon_d \partial_a x^\mu \partial_b \partial_c \overleftrightarrow{\partial}_d x_\mu - trace \quad . \tag{3.5.6}$$

It should be noted that this step does not modify the usual regularization procedure, because the contribution to the stress tensor that is formally of order 2 in ϵ^a is just a local counterterm in the action. What is not equivalent to a local counterterm and therefore modifies the standard perturbation expansion is an additional step which we shall now describe. Before we perform any loop integrations, we write the integrand as a sum of terms, each of which has only one pole in the external momenta. Then in each term we shift the loop momenta, so as to move the pole to zero. Finally, we expand the factors that involve the external momenta in powers of ϵ^a keeping only the terms of order at most 2 in ϵ^a. The shifting of the integration variables typically gives rise to factors of the form $e^{ip\epsilon}$, where p^a is an external momentum. As we just described, we have to replace $e^{ip\epsilon}$ by $1 + ip\cdot\epsilon - \frac{1}{2}(p\cdot\epsilon)^2$. This expansion of factors is not the same for all the terms in which we split the Feynman diagram. This is the reason why the procedure we just presented results in a theory that is not related to the standard theory via local counterterms.

We now proceed to compute the partition function:

$$e^{-W} = \int \mathcal{D}x^\mu e^{-S} \quad , \tag{3.5.7}$$

as a functional of the metric h_{ab}. Differentiating with respect to α', we find

$$\frac{\partial W}{\partial \alpha} = \frac{1}{2}\int d^2\sigma h_{ab}\langle T_{ab}\rangle \quad . \tag{3.5.8}$$

Now,

$$\langle T_{ab}(\sigma)\rangle = \frac{1}{4}\alpha' \int d^2\sigma' h_{a'b'}(\sigma')\langle T_{ab}(\sigma)T_{a'b'}(\sigma')\rangle_0 \quad , \tag{3.5.9}$$

so

$$\frac{\partial W}{\partial \alpha} = \frac{1}{4}\alpha' \int \frac{d^2p}{(2\pi)^2} h_{ab}(p)h_{a'b'}(-p)G_{aba'b'}(p) \quad , \tag{3.5.10}$$

where $G_{aba'b'}(p) = F.T.\langle T_{ab}(\sigma)T_{a'b'}(0)\rangle_0$ is the two-point function. It gets contributions from the various vertices (eq.(3.5.6)). The first term contributing is

$$G^{(1)}_{aba'b'}(p) = d\int \frac{d^2k}{(2\pi)^2} e^{i(2k+p)\epsilon} k_a k_{b'}(p+k)_b(p+k)_{a'}\frac{1}{k^2}\frac{1}{(p+k)^2} \quad . \tag{3.5.11}$$

Working with light-cone coordinates, $\sigma^\pm = \frac{1}{\sqrt{2}}(\sigma^0 \pm i\sigma^1)$, it is easy to show that all indices have to be equal. Thus, only $G^{(1)}_{++++}$ and $G^{(1)}_{----}$ survive. We have

$$\begin{aligned} G^{(1)}_{++++} &= d\int \frac{d^2k}{(2\pi)^2} e^{i(2k+p)\epsilon} k_+(p+k)_+ \frac{1}{p_-}\cdot\left(\frac{k_+}{k^2} - \frac{(p+k)_+}{(p+k)^2}\right) \\ &= \frac{d}{p_-}\int \frac{d^2k}{(2\pi)^2} e^{2ik\epsilon}\left[\left(1 + ip\cdot\epsilon - \frac{1}{2}(p\cdot\epsilon)^2\right)k_+^2(p+k)_+\frac{1}{k^2} - (p^a \to -p^a)\right] \end{aligned}$$

where we expanded the factor $e^{ip\epsilon}$ and kept the terms up to order two in ϵ^a . We easily find

$$G^{(1)}_{++++}(p) = \frac{d}{4\pi}\frac{1}{p_-}\left[\left(1 + ip\cdot\epsilon - \frac{1}{2}(p\cdot\epsilon)^2\right)\left(-\frac{1}{4}\right)\frac{\partial^2}{\partial\epsilon_-^2}\left(-\frac{i}{2}\frac{\partial}{\partial\epsilon_-} + p_+\right)\ln\epsilon^2 \right.$$
$$\left. -(p^a \rightarrow -p^a) \right]$$
$$= \frac{d}{16\pi}\frac{p_+^3}{p_-} + R \quad , \tag{3.5.12}$$

where R vanishes after averaging over directions of ϵ^a. Similarly, we obtain

$$G^{(1)}_{----}(p) = \frac{d}{16\pi}\frac{p_-^3}{p_+} \quad . \tag{3.5.13}$$

The second term contributing to G is

$$G^{(2)}_{aba'b'}(p) = -\frac{1}{4}\epsilon_c\epsilon_d d\int\frac{d^2k}{(2pi)^2}e^{i(2k+p)\epsilon}k_a k_{b'}(p+k)_b(p+k)_{a'}k_c(2k+p)_d\frac{1}{k^2}\frac{1}{(p+k)^2} \quad . \tag{3.5.14}$$

We find

$$G^{(2)}_{++++}(p) = -\frac{d}{16\pi}\epsilon_c\epsilon_d N\frac{1}{p_-}\left[\frac{i}{8}\frac{\partial^2}{\partial\epsilon_-^2}\left(-\frac{i}{2}\frac{\partial}{\partial\epsilon_-} + p_+\right)\frac{\partial}{\partial\epsilon_c}\left(-i\frac{\partial}{\partial\epsilon_d} + p_d\right)\ln\epsilon^2\right.$$
$$\left. -\frac{1}{4}\frac{\partial^2}{\partial\epsilon_-^2}\left(-\frac{i}{2}\frac{\partial}{\partial\epsilon_-} + p_+\right)\left(-\frac{i}{2}\frac{\partial}{\partial\epsilon_c} - p_c\right)\left(-i\frac{\partial}{\partial\epsilon_d} - p_d\right)\ln\epsilon^2\right] \quad , \tag{3.5.15}$$

where we replaced $e^{ip\epsilon}$ by 1, because the Green function is already formally of order 2 in ϵ^a . After some algebra, we find that

$$G^{(2)}_{++++}(p) = -\frac{d}{16\pi}\frac{p_+}{p_-}p_c p_d\epsilon_c\epsilon_d\frac{1}{\epsilon_-^2} \quad . \tag{3.5.16}$$

Averaging gives

$$G^{(2)}_{----}(p) = -\frac{d}{16\pi}\frac{p_-^3}{p_+} \quad . \tag{3.5.17}$$

Similarly, we obtain

$$G^{(2)}_{++++}(p) = -\frac{d}{16\pi}\frac{p_+^3}{p_-} \quad . \tag{3.5.18}$$

Therefore, $G^{(2)}$ cancels $G^{(1)}$, and so

$$G_{++++}(p) = G_{----}(p) = 0 \quad . \tag{3.5.19}$$

It follows from eq.(3.5.10) that the partition function does not depend on the string tension and is therefore also independent of the metric. Using similar arguments, one can show that the Faddeev-Popov determinant does not depend on the metric either. The reason is that it involves a conformally covariant operator, and so the above procedure is readily reproducible. It follows that the ghost degrees of freedom completely decouple and we shall therefore ignore them. Using the BJL theorem [21],

$$\int d\sigma_1 e^{ip^1\sigma_1}[T_{++}(0,\sigma_1), T_{++}(0,0)] = \lim_{p^0\to\infty} p^0 \int d^2\sigma e^{ip\sigma}\mathrm{T}(T_{++}(\sigma)T_{++}(0)) \quad , \tag{3.5.20}$$

where the polynomial terms in p^0 have to be dropped on the right-hand side, it is straightforward to show that eq.(3.5.19) implies

$$\langle 0|[T_{++}(0,\sigma_1), T_{++}(0,0)]|0\rangle = 0 \quad , \tag{21}$$

to be compared with the result of the standard theory,

$$\langle 0|[T_{++}(0,\sigma_1), T_{++}(0,0)]|0\rangle = \frac{d}{48\pi}\delta'''(\sigma_1) \quad . \tag{3.5.22}$$

Equivalently, we can write the following operator-product expansion (OPE):

$$T_{++}(\sigma)T_{++}(\sigma') \sim \frac{2}{(\sigma-\sigma')_+^2}T_{++}(\sigma) + \frac{1}{(\sigma-\sigma')_+}\partial_+T_{++}(\sigma) + \cdots \quad , \tag{3.5.23}$$

and similarly for $T_{--}(\sigma)$. This shows that the central extension term vanishes and therefore $T_{ab}(\sigma)$ contains the effects of the ghosts in the standard theory.

Physical states are annihilated by L_n $(n > 0)$ and $L_0 - c$, where $c = \langle L_0 \rangle$ is the intercept. To compute the intercept, we use the definition of L_n:

$$L_n = \frac{1}{2\pi} \int_0^\pi [e^{-2in\sigma_1} T_{--} + e^{2in\sigma_1} T_{++}] d\sigma_1 \quad . \tag{3.5.24}$$

One has to be a little careful and take into account the fact that the range of σ_1 is finite $(0 \le \sigma_1 \le \pi)$. The calculation is straightforward and we find $c = \frac{d-2}{24}$. It is useful to express L_0 in terms of creation and annihilation operators. These are defined as the fourier components of the position operator,

$$x^\mu = \frac{1}{\sqrt{2}} q^\mu + \sqrt{2} p^\mu \sigma_0 + \frac{i}{2} \sum_{n \neq 0} \frac{1}{n} (a_n^\mu e^{-2in\sigma_-} + \bar{a}_n^\mu e^{2in\sigma_+}) \quad , \tag{3.5.25}$$

where $(a_n^\mu)^\dagger = a_{-n}^\mu$, $(\bar{a}_n^\mu)^\dagger = \bar{a}_{-n}^\mu$, satisfying commutation relations $[a_m^\mu, a_n^\nu] = m\eta^{\mu\nu}\delta_{m+n}$, $[\bar{a}_m^\mu, \bar{a}_n^\nu] = m\eta^{\mu\nu}\delta_{m+n}$, $[p^\mu, q^\nu] = -i\eta^{\mu\nu}$. Thus,

$$L_0 = p^2 + \frac{1}{2} \sum_{n=1}^\infty a_{-n}^\mu a_{n\mu} e^{-2in\epsilon_-} + \frac{1}{2} \sum_{n=1}^\infty \bar{a}_{-n}^\mu \bar{a}_{n\mu} e^{2in\epsilon_+} \quad . \tag{3.5.26}$$

Next, we turn to a study of loops. A one-loop diagram with no external lines is related to the partition function

$$Z(\tau) = |\mathrm{Tr} e^{2\pi i \tau (L_0{}^{(L)} - \frac{c}{2})}|^2 \quad , \tag{3.5.27}$$

where $L_0{}^{(L)} = \frac{1}{2}p^2 + \frac{1}{2} \sum a_{-n}^\mu a_{n\mu}$ is the hamiltonian for the left-moving sector. The partition function has to be modular invariant, which is equivalent to $Z(\tau + 1) = Z(-\frac{1}{\tau}) = Z(\tau)$. A straightforward computation that makes use of coherent states shows that the ϵ-dependent terms do not contribute and the answer is

$$Z(\tau) = |\tau^{-(d-2)/2} [f(\tau)]^{d-2}|^2 \quad , \tag{3.5.28}$$

where $f(\tau) = \prod_{n=1}^\infty (1 - e^{2\pi i \tau n})^{-1}$, which is modular invariant, due to the transformation properties of the function f [23].

To compute the spectrum of the theory, one has to find vertex operators that create the physical states. The standard choice for the lowest-lying (tachyonic) state is the operator $v_k(\sigma_0) = \int_0^\pi d\sigma_1 e^{ikx(\sigma)}$. However, this will not work in our case, because

$$T_{++}(\sigma)e^{ikx(\sigma')} \sim -\frac{k^2}{2(\sigma-\sigma')_+^2} + \frac{k^2}{(\sigma-\sigma')_+^2}e^{ikx(\sigma)} + \frac{1}{(\sigma-\sigma')_+}\partial_+ e^{ikx(\sigma)} + \cdots \quad . \tag{3.5.29}$$

Therefore, we have to modify $\tilde{v}_k$ to eliminate the central term. We have not been able to find the appropriate modification of $v_k(\sigma_0)$. However, the fact that the intercept is $c = \frac{d-2}{24}$ indicates that the mass of a tachyon is $m^2 = -\frac{d-2}{24}$.†

In the case of open strings, one can also argue that a one-loop diagram with N external lines possesses no cuts. Thus, a source of violation of unitarity is absent in our formalism. To see how this works, consider

$$\begin{aligned} A_{\text{n.p.}}(k_1,\cdots,k_N) = \text{Tr}\Big\{ & \frac{1}{\tilde{L}_0 - c} v_{k_1}(0)\cdots v_{k_i}(0)\Omega \\ & \times \frac{1}{\tilde{L}_0 - c} v_{k_i+1}(0) \cdots \frac{1}{\tilde{L}_0 - c} v_{k_N}(0)\Omega \Big\} \quad , \end{aligned} \tag{3.5.30}$$

where Ω is the twist operator ($\Omega^{-1}x^\mu(0,\sigma_1)\Omega = x^\mu(\pi,\sigma_1)$), and $\tilde{L}_0$ is the hamiltonian for open strings. In the standard theory, $c = 1$, so using the identity $\frac{1}{\tilde{L}_0 - c} = \int_0^\infty d\tau e^{-\tau(\tilde{L}_0 - c)}$ we can rewrite eq.(3.5.30) as follows:

$$\begin{aligned} A_{\text{n.p.}}(k_1,\cdots,k_N) = \prod_{j=1}^{N} \int_0^\infty d\tau_j \text{Tr}\Big\{ & v_{k_1}(\tau_1)\cdots v_{k_i}(\tau_1+\cdots+\tau_i)\Omega v_{k_i+1}(\tau_1+\cdots+\tau_{i+1}) \\ & \times\cdots\times v_{k_N}(\tau_1+\cdots+\tau_N)\Omega e^{-(\tau_1+\cdots+\tau_N)(\tilde{L}_0-1)} \Big\} \quad , \end{aligned} \tag{3.5.31}$$

where we used the fact that $e^{-\tau\tilde{L}_0}v_k(0)e^{\tau\tilde{L}_0} = v_k(\tau)$ is the vertex operator at imagi-

† Notice that if $d = 26$, then $m^2 = -1$, in agreement with the standard theory.

nary time $\sigma_0 = i\tau$. Making use of coherent states, we obtain [23]

$$A_{\text{n.p.}}(k_1, \cdots, k_N) = \int_0^1 \frac{dq}{q} (\ln q)^{(d-26)/24} g(q) \quad . \tag{3.5.32}$$

The behavior of the function g for small q is $g(q) \sim q^{-s-2}$, where $s = (k_1 + \cdots + k_i)^2$ is a Mandelstam variable. Therefore, the poles are the same as for the Koba-Nielsen tree amplitude (eq.(3.5.33)). Also the diagram has a cut if $(\ln q)^{(d-26)/24}$ does not equal 1. This is the case for all $d \neq 26$. The factor $(\ln q)^{(d-26)/24}$ contains a contribution $(\ln q)^1$ from the factor involving the integration variables $e^{c(\tau_1 + \cdots + \tau_N)}$ (cf. eq. (3.5.31)), where $c = 1$. Using our method, eq.(3.5.31) is modified to

$$A_{\text{n.p.}}(k_1, \cdots, k_N) = \prod_{j=1}^{N} \int_0^\infty d\tau_j \text{Tr}[v_{k_1}(\tau_1) \cdots v_{k_i}(\tau_1 + \cdots + \tau_i) \Omega v_{k_i+1}(\tau_1 + \cdots + \tau_{i+1}) \times \\ \times \cdots \times v_{k_N}(\tau_1 + \cdots + \tau_N) \Omega e^{-(\tau_1 + \cdots + \tau_N)(L_0 - \frac{d-2}{24})}] \quad . \tag{3.5.33}$$

Therefore, instead of $e^{\tau_1 + \cdots + \tau_N}$, we obtain a factor of $e^{(\tau_1 + \cdots + \tau_N)(d-2)/24}$, which leads to a factor $(\ln q)^0 = 1$, for all dimensions d. Thus, no cuts exist in any number of dimensions and the theory appears to be unitary.

Finally, let us mention that space-time Lorentz invariance can also be demonstrated by computing vacuum expectation values of the commutators of the generators of the Lorentz group. One can follow the steps leading to eq.(3.5.21) to show that they all vanish, if our method is employed.

References

1. J. Steinberger, *Phys. Rev.* **76** (1949) 1180;

   J. Schwinger, *Phys. Rev.* **128** (1962) 2425;

   S. L. Adler, *Phys. Rev.* **177** (1969) 2426;

   J. S. Bell and R. Jackiw, *Nuovo Cim.* **60A** (1969) 47.

2. D. J. Gross and R. Jackiw, *Phys. Rev.* **D6** (1972) 477;

   C. Bouchiat, J. Iliopoulos and Ph. Meyer, *Phys. Lett.* **38B** (1972) 519.

3. J. Wess and B. Zumino, *Phys. Lett.* **37B** (1971) 95.

4. C. Becchi, A. Rouet and R. Stora, *Phys. Lett.* **52B** (1974) 344.

5. T. Sumitani, *J. Phys. A: Math. Gen.* **17** (1984) L811.

6. M. F. Atiyah and I. M. Singer, *Proc. Nat. Acad. Sci. USA* **81** (1984) 2597.

7. L. D. Faddeev and S. L. Shatashvili, *Phys. Lett.* **167B** (1986) 225.

8. A. M. Polyakov, *Phys. Lett.* **103B** (1981) 207.

9. K. Fujikawa, *Phys. Rev.* **D29** (1984) 285;

   D. W. McKay and B.-L. Young, *Phys. Rev.* **D28** (1983) 1039;

   M. B. Einhorn and D. R. T. Jones, *Phys. Rev.* **D29** (1984) 331.

10. G. Siopsis, Caltech preprint, CALT-68-1335.

11. K. Fujikawa, *Phys. Rev.* **D21** (1980) 2848.

12. L. Alvarez-Gaumé and P. Ginsparg, *Ann. Phys.* **161** (1985)423.

13. L. Alvarez-Gaumé and P. Ginsparg, *Nucl. Phys.* **B243** (1984) 449.

14. G. Moore and Ph. Nelson, *Commun. Math. Phys.* **100** (1985) 83.

15. E. Witten, *Phys. Lett.* **117B** (1982) 324.

16. G. Siopsis, *Phys. Lett.* **183B** (1987) 187.

17. R. Jackiw and R. Rajaraman, *Phys. Rev. Lett.* **54** (1985) 1219;

J. Lott and R. Rajaraman, *Phys. Lett.* **165B** (1985) 321;

I. G. Halliday, E. Rabinovici, A. Schwimmer and M. Chanowitz, *Nucl. Phys.* **B268** (1986) 413.

18. W. A. Bardeen, *Phys. Rev.* **184** (1969) 1848.

19. L. D. Faddeev, *Phys. Lett.* **145B** (1984) 81.

20. M. Kobayashi and A. Sugamoto, *Phys. Lett.* **159B** (1985) 315;

S. Jo, *Phys. Lett.* **163B** (1985) 353;

H. Sonoda, *Nucl. Phys.* **B266** (1986) 410.

21. J. Bjorken, *Phys. Rev.* **148** (1966) 1467.

22. G. Siopsis, Caltech preprint (1987), CALT-68-1396.

23. J. H. Schwarz, *Phys. Rep.* **89** (1982) 223.

24. M. B. Green, J. H. Schwarz and E. Witten, *Superstring Theory* (Cambridge University Press, 1987).

4. Strings on group manifolds

After the discovery of anomaly cancellation in string theories by Green and Schwarz [1], strings have recently emerged as the prime candidate for a theory that would explain all fundamental interactions. A remarkable property of string theories is that they are consistent only in certain critical dimensions, which are 26 for the bosonic string and 10 for the superstring. Since these are apparently unworldly dimensions, if we want strings to describe our perceived four-dimensional world, we must compactify the redundant dimensions to an internal space $\mathcal{K}$.♮ Unfortunately, most choices of $\mathcal{K}$ lead to models from which it is hard to extract any predictions. However, if $\mathcal{K}$ is a group manifold, the model is exactly solvable, as has been discussed by Knizhnik and Zamolodchikov [3] and Gepner and Witten [4]. In this Chapter, we shall first review their construction of consistent models in the bosonic case, and then implement similar ideas on superstrings.

4.1 Bosonic strings

We shall quantize a two-dimensional sigma model with group G, of dimension D, whose action is of the Wess-Zumino form [5]:

$$S = \alpha \int d^2\sigma \mathrm{tr}(\partial_a g \partial_a g^{-1}) + \frac{k}{24\pi} \int d^3\sigma \varepsilon^{ijk} \mathrm{tr}(g^{-1}\partial_i g)(g^{-1}\partial_j g)(g^{-1}\partial_k g) \ , \quad (4.1.1)$$

where $g \in G$. The second integral is over a three-dimensional surface whose boundary is the two-dimensional integration region of the first integral. This term is defined modulo 2π [6], which requires k to be an integer, so that the factor e^{iS} in the path integral be well-defined and independent of the three-dimensional surface chosen. We will be interested in the propagation of closed strings, and therefore we impose the boundary conditions $g(\sigma_1 = 0) = g(\sigma_1 = 2\pi)$. For convenience, we define new coordinates z and $\bar{z}$ defined as $z = e^{\sigma_0 + i\sigma_1}$ and $\bar{z} = e^{\sigma_0 - i\sigma_1}$. Henceforth, we shall be working with z and $\bar{z}$, instead of σ_0 and σ_1.

♮ For a review and references on the issue of compactification to four dimensions see ref.[2].

At the critical point, $\alpha = \frac{k}{16\pi}$, the theory is invariant under the transformations generated by the currents:

$$J(z) \equiv J^a T^a = -\frac{k}{2}(\partial_z g)g^{-1} \ , \tag{4.1.2a}$$

$$\bar{J}(\bar{z}) \equiv \bar{J}^a T^a = -\frac{k}{2}(\partial_{\bar{z}} g^{-1})g \ , \tag{4.1.2b}$$

where the T^a's are the generators of the Lie algebra of G ($[T^a, T^b] = f^{abc}T^c$ and $\mathrm{tr}T^aT^b = c\delta^{ab}$, c being the second Casimir of a certain representation of G). The components of $J(z)$ in a Laurent expansion, $J(z) = \sum J_n z^{-n-1}$, generate a Kac-Moody algebra,†

$$[J_n^a, J_m^b] = if^{abc}J_{n+m}^c + kn\delta^{ab}\delta_{n+m,0} \ , \tag{4.1.3}$$

and similarly for $\bar{J}$. From now on, we concentrate on the left-moving sector. The discussion of the right-moving sector follows the same lines.

The energy-momentum tensor takes the Sugawara form:

$$T(z) = \frac{1}{2k + c_A} : J^a(z)J^a(z) : \ . \tag{4.1.4}$$

Its coefficients in the Laurent expansion, $T(z) = \sum L_n z^{-n-2}$ obey the Virasoro algebra,

$$[L_n, L_m] = (n-m)L_{n+m} + \frac{1}{12}c(n^3 - n)\delta_{n+m,0} \ , \tag{4.1.5}$$

where the central charge is given by

$$c = \frac{2kD}{2k + c_A} \ , \tag{4.1.6}$$

c_A being the second Casimir of the adjoint representation of G. The physical states are annihilated by L_n and $\bar{L}_n$ ($n > 0$). We define highest-weight vectors (primary

† This is to be compared with the algebra of the coefficients a_n^μ of the coordinates x^μ (eq.(3.5.25)), which is obtained from eq.(4.1.3) for $G = U(1)^d$.

states) as eigenstates of the zero modes L_0 and J_0^a:

$$L_0|R,i\rangle = \Delta|R,i\rangle \quad , \quad J_0^a|R,i\rangle = (T_R^a)_{ij}|R,j\rangle \quad , \tag{4.1.7}$$

where R denotes an irreducible representation of the group G. These states are physical states provided

$$\Delta = \frac{c_R}{2k + c_A} \quad , \tag{4.1.8}$$

where c_R is the second Casimir of the representation R ($\mathrm{tr} T_R^a T_R^a = c_R \delta^{ab}$).

By acting with the raising operators J_{-n}^a ($n > 0$) on the primary states, we obtain all physical states. Each primary state is created from the in-vacuum by a field $\phi_i^R(z)$ (primary field):

$$|R,i\rangle = \phi_i^R(0)|0\rangle \ . \tag{4.1.9}$$

Since all physical states can be created from the primary fields, the correlators of primary fields contain enough information to determine all Green functions. To calculate them, we shall make use of the existence of null vectors, in analogy with the solution of critical systems by Belavin et al.[7].

A representation of the semi-direct product of a Kac-Moody and a conformal algebra is said to be degenerate if there is a secondary state in it (i.e., a state generated from the highest-weight vector by the action of lowering operators), which has the properties of a highest-weight vector, i.e., it is annihilated by all the raising operators of the algebra. This state is said to be null, and the corresponding primary field is said to be degenerate. It is easy to check that the following state is null:

$$|\chi\rangle = \left\{ \left(k + \frac{c_A}{2} \right) L_{-1} + T^a J_{-1}^a \right\} |R,i\rangle \ . \tag{4.1.10}$$

Its inner product with any other state in the representation generated by the initial primary field, $\phi(z)$, vanishes identically. Thus, the sub-representation generated by the null vector, $\chi(z)$, can be consistently set to zero. In particular, its correlation functions with all the other primary fields vanish:

$$\langle \phi^{R_1}(z_1) \cdots \phi^{R_{i-1}}(z_{i-1}) \chi^{R_i}(z_i) \phi^{R_{i+1}}(z_{i+1}) \cdots \phi^{R_n}(z_n) \rangle = 0 \ . \qquad (4.1.11)$$

Since $[L_{-1}, \phi] = \partial_z \phi$ and $[J^a_{-1}, \phi] = z^{-1} T^a \phi$, eq.(4.1.11) can be written as:

$$\left[\left(k + \frac{c_A}{2} \right) \frac{\partial}{\partial z_i} - \sum_{j \neq i} \frac{T^a_{R_i} T^a_{R_j}}{z_i - z_j} \right] \langle \phi^{R_1}(z_1) \cdots \phi^{R_n}(z_n) \rangle = 0 \ . \qquad (4.1.12)$$

This is a set of differential equations that can in principle determine all correlators.

One can obtain an additional null vector by considering the Kac-Moody algebra of the theory (eq.(4.1.3)). It is thus possible to extract extra algebraic matrix equations that are satisfied by correlators of primary fields. These equations imply that primary fields corresponding to non-integrable representations of G vanish identically.‡ On the other hand, each integrable representation appears exactly once.

We now couple the Wess-Zumino model (eq.(4.1.1)) to a sigma model defined on a Minkowski space of dimension d (eq.(3.5.1)). Let a^μ_n be the fourier components of the coordinates x^μ ($\mu = 1, \ldots, d$). The mass formula for this system is

$$m^2 = -p^2 = -2 + \frac{c_R + c_{\bar{R}}}{2k + c_A} + N + \bar{N} + M + \bar{M} \ , \qquad (4.1.13)$$

where $N = \frac{1}{2k+c_A} \sum_{m \neq 0} : J^a_{-m} J^a_m :$, $M = \frac{1}{2} \sum_{m \neq 0} : a^\mu_{-m} a_{\mu m} :$ are the number operators for the left-moving sector in the group and Minkowski space, respectively. $\bar{N}$ and $\bar{M}$ are the corresponding operators in the right-moving sector. It is clear that the spectrum includes a tachyon ($m^2 = -2$) that transforms as a singlet under some representation R. Since there is no preferred point on a closed string, the physical

‡ An integrable representation is defined as the one satisfying $k \geq 2(\lambda, \theta)$, where λ is the highest weight of the representation and θ is the highest root of the Lie algebra of the group G [8]. For $SU(N)$, e.g., $2(\lambda, \theta)$ is the length of the first row of the Young tableau.

states must be invariant under translations in the σ_1 direction. These are generated by $L_0 - \bar{L}_0$. It follows that

$$N + M + \frac{c_R}{2k + c_A} = \bar{N} + \bar{M} + \frac{c_{\bar{R}}}{2k + c_A} \quad . \tag{4.1.14}$$

The theory is also seen to be modular-invariant at the critical dimension: $d + c = 26$, i.e.,

$$d + \frac{2kD}{2k + c_A} = 26 \quad . \tag{4.1.15}$$

4.2 The N=1 super Wess-Zumino model

The main drawbacks of the model we have just described are the absence of space-time fermions and the existence of a tachyonic mode (cf. eq.(4.1.13)). One expects this mode to disappear if supersymmetry is incorporated into the model, in a manner similar to the case of flat space. We therefore proceed to study the N=1 supersymmetric extension of this model.

Recently, there has been a lot of interest in Kac-Moody algebras from both the mathematical and the physical points of view [6]. In particular, Kac-Moody algebras are seen to play an important role in conformally invariant two-dimensional models with continuous symmetries [3,9], as well as in string theories [4,10,11].

The Wess-Zumino models on group manifolds, describing string propagation in the group manifold, are typical models realizing such an algebra. Their supersymmetric version has been studied recently [12], and a new structure, that of a (N=1) super Kac-Moody algebra, has emerged. This algebra is essential in describing superstring propagation in a group manifold.[§]

In this section, we derive the transformation properties of the fields under a super Kac-Moody algebra, and we solve the projective Ward identities for the 2- and 3-point functions. We focus our attention on the degenerate representations of

§ A phenomenological study of these models will be presented in the next section.

the algebra that appear in the super Wess-Zumino model, and we derive the linear differential equations for the correlation functions of the degenerate fields. We solve these equations for the 3-point function, obtaining constraints on the dimensions of the fields that may exist in the theory. We thus show that the operator algebra of the degenerate fields closes in the same way as in the ordinary Wess-Zumino model. We also solve these equations to determine the 4-point function, which is the first non-trivial Green function. Some implications are also discussed.

The super Kac-Moody algebra is generated by the current superfield $\mathcal{J}^a \equiv \psi^a(z)+\theta J^a(z)$. In terms of the fourier modes of the supercurrent, this algebra is

$$\begin{aligned} [J^a_m, J^b_n] &= i f^{abc} J^c_{m+n} + k\delta^{ab} m \delta_{m+n,0} \quad , \\ [J^a_m, \psi^b_r] &= -i f^{abc} \psi^c_{m+r} \ , \quad \{\psi^a_r, \psi^b_s\} = \delta^{ab}\delta_{r+s,0} \quad , \end{aligned} \tag{4.2.1}$$

where f^{abc} are the structure constants of a semisimple Lie group G . We also have $(J^a_m)^\dagger = J^a_{-m}$ and $(\psi^a_r)^\dagger = \psi^a_{-r}$. One distinguishes between two sectors, the NS sector, where $\psi^a(z)$ is anti-periodic on the cylinder, and the R sector, where $\psi^a(z)$ is periodic on the cylinder. In this section, we shall only consider the NS sector. It is convenient to pass from the cylinder to the plane through the super-analytic transformation $(\ln z, z^{-1/2}\theta) \rightarrow (\sigma_0 + i\sigma_1, \theta)$. Then in the NS sector fermionic fields are single-valued whereas in the R sector the fermionic fields are double-valued on the plane.

A theory invariant under the algebra (4.2.1) also has an N=1 superconformal invariance. The generators of the superconformal algebra can be constructed from those of the super Kac-Moody algebra (in the Sugawara form),

$$L_n \equiv \frac{1}{2} \sum_{r\in Z+\frac{1}{2}} \left(r+\frac{1}{2}\right) : \psi^a_{n-r}\psi^a_r : + \frac{1}{2k} \sum_{m\in Z} : \tilde{J}^a_{n-m} \tilde{J}^a_m : \ , \tag{4.2.2a}$$

$$G_r \equiv -\frac{1}{\sqrt{k}} \sum_{m\in Z} : \psi^a_{r-m} \tilde{J}^a_m : + \frac{1}{6\sqrt{k}} i f^{abc} \sum_{m\in Z, r'\in Z+\frac{1}{2}} : \psi^a_{r-m}\psi^b_{m-r'}\psi^c_{r'} : \tag{4.2.2b}$$

$$\tilde{J}^a_m \equiv J^a_m - \frac{i}{2} f^{abc} \sum_{r\in Z+\frac{1}{2}} : \psi^b_{m-r}\psi^c_r : \ , \tag{4.2.2c}$$

where $\tilde{J}^a_m$ is the "bosonic" current. It can easily be seen that

$$[\tilde{J}^a_m, \psi^b_r] = 0 \quad , \tag{4.2.3a}$$

$$[\tilde{J}^a_m, \tilde{J}^b_n] = i f^{abc} \tilde{J}^c_{m+n} + \left(k - \frac{c_A}{2}\right) \delta^{ab} m \delta_{m+n,0} \quad . \tag{4.2.3b}$$

Demanding unitarity on the bosonic part, we obtain $k \geq c_A/2$. The Fock in-vacuum $|0\rangle$ is defined as the state annihilated by the generators $J^a_n (n \geq 0)$ and $\psi^a_r (r > 0)$. We end up with the semidirect product of the N=1 superconformal algebra and the super Kac-Moody algebra defined by the commutation relations (4.2.1) together with:

$$\begin{aligned}
&[L_m, L_n] = (m-n) L_{m+n} + \frac{\hat{c}}{8}(m^3 - m)\delta_{m+n,0} \quad , \\
&[L_m, G_r] = \left(\frac{m}{2} - r\right) G_{m+r} \quad , \quad [L_m, J^a_n] = -n J^a_{m+n} \quad , \\
&[L_m, \psi^a_r] = -\left(\frac{m}{2} + r\right) \psi^a_{m+n} \quad , \quad \{G_r, G_s\} = 2L_{r+s} + \frac{\hat{c}}{2}\left(r^2 - \frac{1}{4}\right)\delta_{r+s,0} \quad , \\
&[G_r, J^a_m] = \sqrt{k} m \psi^a_{m+r} \quad , \quad \{G_r, \psi^a_s\} = -\frac{1}{\sqrt{k}} J^a_{r+s} \quad ,
\end{aligned} \tag{4.2.4}$$

where $\hat{c} = \left(1 - \frac{c_A}{3k}\right) D$, D being the dimension of the group G. We will focus on the left sector of the theory, the full theory being the direct product of the left and right sectors. The highest-weight vectors of this algebra (primary states) are labeled by the eigenvalues of the zero modes L_0, J^a_0:

$$L_0 |R, i\rangle = \Delta |R, i\rangle \quad , \quad J^a_0 |R, i\rangle = (T^a_R)_{ij} |R, j\rangle \quad , \tag{4.2.5}$$

where R denotes an irreducible representation of the group G . We also have $L_n|R\rangle = J^a_n|R\rangle = G_r|R\rangle = \psi^a_r|R\rangle = 0$, for $n, r > 0$. These states are generated by the action of superfield operators, called primary superfields, on the in-vacuum,

$$|R_i\rangle \equiv \Phi^R_i(0)|0\rangle \quad , \tag{4.2.6a}$$

where

$$\Phi^R_i(z, \theta) = \phi^R_i(z) + \theta \psi^R_i(z) \quad . \tag{4.2.6b}$$

The algebra acts on the primary superfields as follows:

$$[L_m, \Phi_i^R(z,\theta)] = z^{m+1}\partial_z \Phi_i^R(z,\theta) + (m+1)z^m \left(\Delta + \frac{1}{2}\theta\frac{\partial}{\partial\theta}\right)\Phi_i^R(z,\theta) \,, \quad (4.2.7a)$$

$$[G_r, \Phi_i^R(z,\theta)] = z^{r+\frac{1}{2}}\left(\frac{\partial}{\partial\theta} - \theta\frac{\partial}{\partial z}\right)\Phi_i^R(z,\theta) - 2\Delta\left(r+\frac{1}{2}\right)z^{r-\frac{1}{2}}\theta\Phi_i^R(z,\theta) \ , \quad (4.2.7b)$$

$$[J_m^a, \Phi_i^R(z,\theta)] = z^m (T_R^a)_{ij}\Phi_j^R(z,\theta) \ , \quad (4.2.7c)$$

$$[\psi_r^a, \Phi_i^R(z,\theta)] = \frac{1}{\sqrt{k}} z^{m-\frac{1}{2}}(T_R^a)_{ij}\theta\Phi_j(z,\theta) \ . \quad (4.2.7d)$$

Here Δ is the conformal weight of the superfield Φ, defined in equation (4.2.5). The algebra above follows from the transformation of the superfield under the superconformal group and the Jacobi identities.

The theory is invariant under the global superconformal group, OSP(2|1), which is generated by the operators $G_{\pm 1/2}$, $L_{\pm 1}$, L_0, due to the fact that the vacuum is also OSP(2|1) invariant. We can derive appropriate Ward identities for the correlation functions reflecting the invariance mentioned above. The procedure is to insert a generator of OSP(2|1) in a correlation function acting on the in-vacuum and move it to the left using the commutation relations (4.2.7). Let us consider the 2-point function, whose form is fixed by the Ward identities from global superconformal invariance [13],

$$\langle 0|\Phi_i^{R_1}(z_1,\theta_1)\Phi_j^{R_2}(z_2,\theta_2)|0\rangle = \frac{A_{ij}}{z_{12}^{\Delta_{12}}}\delta_{\Delta_1\Delta_2} \ , \quad (4.2.8)$$

where $z_{12} = z_1 - z_2 - \theta_1\theta_2$, and $\Delta_{12} = \Delta_1 + \Delta_2$.

The vacuum is also invariant under global G-transformations, that is, the zero mode J_0^a annihilates the vacuum. The Ward identity for the zero mode of $J^a(z)$ implies

$$(T_{R_1}^a)_{ik}A_{kj} + (T_{R_2}^a)_{jk}A_{ik} = 0 \ , \quad (4.2.9)$$

with a solution

$$A_{ij} \sim \langle R_1, R_2, i, j | 1, 0 \rangle \quad , \tag{4.2.10}$$

which is the Clebsch-Gordan coefficient of the projection of $R_1 \times R_2$ on the singlet representation.

The 3-point function is constrained by the superconformal invariance to have the form [13]

$$\langle 0 | \Phi_i^{R_1}(z_1, \theta_1) \Phi_j^{R_2}(z_2, \theta_2) \Phi_k^{R_3}(z_3, \theta_3) | 0 \rangle = \frac{A_{ijk}}{z_{12}^{\Delta_{12}}} z_{13}^{\Delta_{13}} z_{23}^{\Delta_{23}} (1 + a\hat{\eta}) \quad , \tag{4.2.11}$$

where

$$\hat{\eta} = (z_{12} z_{13} z_{23})^{-1/2} (\theta_1 z_{23} - \theta_2 z_{13} + \theta_3 z_{12} + \theta_1 \theta_2 \theta_3) \tag{4.2.12}$$

is the only combination of the coordinates that is invariant under the global superconformal group OSP(2|1) squaring to zero. Thus a is an extra undetermined Grassmann parameter.

The current Ward identity is in this case:

$$(T^a_{R_1})_{il} A_{ljk} + (T^a_{R_2})_{jl} A_{ilk} + (T^a_{R_3})_{kl} A_{ijl} = 0 \quad , \tag{4.2.13}$$

with the solution,

$$A_{ijk} \sim \langle R_1, R_2, i, j | R_3, k \rangle \quad , \tag{4.2.14}$$

being the appropriate Clebsch-Gordan coefficient. The condition for the 3-point function to be non-zero is that the primary superfield Φ_3 be contained in the operator product of Φ_1, and Φ_2. Then the z-independent part of the 3-point function is the operator product coefficient multiplying Φ_3 in the expansion of the product $\Phi_1 \times \Phi_2$. Let us remark that, unlike the non-supersymmetric case, there are two operator-product coefficients to be determined here, one corresponding to the overall normalization, the other corresponding to the free parameter multiplying $\hat{\eta}$.

To proceed further, we shall make use of the existence of null states in the theory. The Sugawara form of the superconformal generators implies that the following state is null:

$$|\chi\rangle = \left[\sqrt{k}G_{-1/2} + T^a_R\psi^a_{-1/2}\right]|R\rangle \quad . \tag{4.2.15}$$

It is easy to verify that $|\chi\rangle$ is annihilated by all the raising operators, provided that its dimension is, $\Delta = \frac{c_R}{2k}$.

As we discussed in the previous section, the existence of degenerate representations in a theory is of prime importance, because in such a case the correlation functions of a degenerate superfield satisfy additional linear (super)differential equations which allow one to determine them completely.

To illustrate the above statement, consider the 3-point function with one of the fields, $\Phi^{R_3}_i$ say, being degenerate . Taking advantage of the invariance of the correlation functions under global superconformal transformations, we can perform a translation and a supersymmetry transformation, to bring it into the form

$$\begin{aligned} F_{ijk} &\equiv \langle 0|\Phi^{R_1}_i(\tilde{z}_1,\tilde{\theta}_1)\Phi^{R_2}_j(\tilde{z}_2,\tilde{\theta}_2)\Phi^{R_3}_k(0,0)|0\rangle \\ &= \frac{A_{ijk}}{(\tilde{z}_1-\tilde{z}_2-\tilde{\theta}_1\tilde{\theta}_2)^{\Delta_{12}}\tilde{z}_1^{\Delta_{13}}\tilde{z}_2^{\Delta_{23}}}(1+a\tilde{\eta}) \quad , \end{aligned} \tag{4.2.16}$$

where

$$\begin{aligned} \tilde{z}_1 &= z_1 - z_3 - \theta_1\theta_3 \quad , \quad \tilde{\theta}_1 = \theta_1 - \theta_3 \quad , \\ \tilde{z}_2 &= z_2 - z_3 - \theta_2\theta_3 \quad , \quad \tilde{\theta}_2 = \theta_2 - \theta_3 \quad . \end{aligned} \tag{4.2.17}$$

Using the fact that the field Φ_{deg}, corresponding to the null state $|\chi\rangle$, has vanishing correlation functions with all other fields, we obtain

$$\langle 0|\Phi^{R_1}_i(\tilde{z}_1,\tilde{\theta}_1)\Phi^{R_2}_j(\tilde{z}_2,\tilde{\theta}_2)\left(\sqrt{k}G_{-1/2}\delta_{kl} + (T^a_{R_3})_{kl}\psi^a_{-1/2}\right)\Phi^{R_3}_l(0,0)|0\rangle = 0 \quad . \tag{4.2.18}$$

Commuting the generator of the algebra through to the left using eq.(4.2.7), we arrive at the following super-equation for the 3-point function (we drop the tildes from now on):

$$k\sum_{i=1}^{2}\left(\frac{\partial}{\partial\theta_i}-\theta_i\frac{\partial}{\partial z_i}\right)F_{ijk}+\frac{\theta_1}{z_1}(T^a_{R_3})_{km}(T^a_{R_1})_{il}F_{ljm}+\frac{\theta_2}{z_2}(T^a_{R_3})_{km}(T^a_{R_2})_{jl}F_{ilm}=0\ . \tag{4.2.19}$$

Eq.(4.2.19) implies that the odd part of the correlation function is zero ($a=0$) , and also

$$k\Delta_{13}A_{ijk}+(T^a_{R_3})_{km}(T^a_{R_1})_{il}A_{ljm}=0\ , \tag{4.2.20a}$$

$$k\Delta_{23}A_{ijk}+(T^a_{R_3})_{km}(T^a_{R_2})_{il}A_{ljm}=0\ . \tag{4.2.20b}$$

Using the current Ward identities (eq.(4.2.13)), it is easy to show that eqs.(4.2.20a) and (4.2.20b) are equivalent. We therefore only consider eq.(4.2.20b). After some straightforward algebra, it follows from eq.(4.2.13) that

$$(T^a_{R_2})_{jl}(T^a_{R_3})_{km}A_{ilm}=\frac{1}{2}(c_{R_1}-c_{R_2}-c_{R_3})A_{ijk}\ . \tag{4.2.21}$$

Consequently, if the fields Φ_{R_2} and Φ_{R_3} belong to degenerate representations, i.e., if $\Delta_2=\frac{c_{R_2}}{2k}$ and $\Delta_3=\frac{c_{R_3}}{2k}$, then $\Delta_1=\frac{c_{R_1}}{2k}$. This proves the closure under operator-product expansion of the degenerate representations of the semi-direct product of the superconformal and the super Kac-Moody algebras. Since any 3-point function of secondary fields is related, via the superconformal and G-Ward identities, to the 3-point function of the corresponding primary superfields, our results apply to any 3-point function. This fact is important for the construction of a superstring theory on a group manifold, since it implies that the corresponding vertex operators form a closed algebra and the amplitudes factorize onto physical intermediate states [14].

When $c_A=2k$, the representations of the super Kac-Moody algebra possess additional null states that are constructed out of the modes J^a_{-n} ,ψ^a_{-r} $(n,r>0)$. In this case, the central charge of the "bosonic" Kac-Moody algebra (eq.(4.2.3b)) vanishes, so we are left with only the free fermions that realize the supersymmetry non-linearly.

The states that remain to be considered are the proper null highest-weight vectors of the Kac-Moody algebra. These are obtained by the action of lowering operators of the Kac-Moody algebra on primary states. The operator algebra of those representations has been discussed in the previous section [4].♮ Combining the results of ref.[3] with ours, we have complete knowledge of the minimal system of representations of the super Wess-Zumino theory. In fact, the theory is exactly solvable in the sense that all the correlation functions satisfy a super-equation of the form (4.2.19) and are therefore computable in principle. Below, we present an explicit evaluation of the 4-point function, which contains non-trivial information on the non-vanishing operator-product coefficients of the operator algebra. OSP(2|1) invariance implies that the 4-point function is of the form

$$\begin{aligned} F_{ijkl} &\equiv \langle 0|\Phi_i^{R_1}(z_1,\theta_1)\Phi_j^{R_2}(z_2,\theta_2)\Phi_k^{R_3}(z_3,\theta_3)\Phi_l^{R_4}(z_4,\theta_4)|0\rangle \\ &= A_{ijkl}^{K}\prod_{I<J}(z_{IJ})^{\gamma_{IJ}}[f_K(x)+yg_K(x)] \ , \end{aligned} \tag{4.2.22}$$

where x, y are the two independent commuting combinations of the coordinates invariant under OSP(2|1),

$$x=\frac{z_{12}z_{34}}{z_{13}z_{24}} \ , \ y=x+\frac{z_{14}z_{23}}{z_{13}z_{24}}-1 \ , \ y^2=0 \ , \tag{4.2.23a}$$

and

$$\gamma_{IJ}=\gamma_{JI} \ , \ \sum_{I\neq J}\gamma_{IJ}=-2\Delta_J \ . \tag{4.2.23b}$$

Using the current Ward identities for the 4-point function, we can compute the group coefficient:

$$A_{ijkl}^{K}\sim\sum_{R,R',m,m'}\langle R_1,R_2,i,j|R,m\rangle\langle R',m'|R_3,R_4,k,l\rangle\langle R,R',m,m'|1_K,0\rangle \ , \tag{4.2.24}$$

where the index K labels the singlets in the product. The equation satisfied by the 4-point function can be derived in the same way as eq.(4.2.19) (The variables here

♮ The selection rules derived in this case state that all non-integrable representations decouple.

are the tilded ones (cf. eq.(4.2.17)):

$$k \sum_{i=1}^{3} \left(\frac{\partial}{\partial \theta_i} - \theta_i \frac{\partial}{\partial z_i} \right) F_{ijkl} + (T^a_{R_4})_{lm} \left\{ \frac{\theta_1}{z_1} (T^a_{R_1})_{in} F_{njkm} + \frac{\theta_2}{z_2} (T^a_{R_2})_{jn} F_{inkm} + \frac{\theta_3}{z_3} (T^a_{R_3})_{kn} F_{ijnm} \right\} = 0 \ . \tag{4.2.25}$$

We shall present the solution to this equation for the simplest non-trivial case, namely $G = SU(2)$ and R_1, R_2, R_3, R_4 all being the fundamental representation of $SU(2)$. Other cases do not require new techniques but considerably more labor. There are two singlets in the product above, so we can write:

$$F_{ijkl} = F_1(x,y)\delta_{ij}\delta_{kl} + F_2(x,y)\delta_{ik}\delta_{jl} \ . \tag{4.2.26}$$

In this case, eq.(4.2.25) is a 2×2 matrix equation. Using the identity

$$(T^a)_{ij}(T^a)_{kl} = \frac{1}{2}\left(\delta_{il}\delta_{jk} - \frac{1}{2}\delta_{ij}\delta_{kl}\right) \ , \tag{4.2.27}$$

we can reduce (4.2.25) to two independent equations in F_1 and F_2, respectively, which are of the hypergeometric type. Their solutions, to lowest order in θ_i, are:

$$F_1(x) = \left(\frac{1-x}{x^3}\right)^{1/4k} F\left(\frac{1}{2k}, -\frac{1}{2k}, 1 - \frac{1}{k}, x\right) \ , \tag{4.2.28a}$$

$$F_2(x) = [x(1-x)]^{1/4k} F\left(\frac{3}{4k}, \frac{1}{4k}, 1 + \frac{1}{k}, x\right) \ . \tag{4.2.28b}$$

It is straightforward to uncover the θ-dependence of the 4-point amplitude and normalize it correctly by factorizing it on 3-point functions.

The equation above has a very simple power-law solution in the special case where there is only one singlet contained in the product. Then,

$$(T^a_{R_4})_{lm}(T^a_{R_1})_{kn}F_{ijnm} = k a_{14} F_{ijkl} \ , \tag{4.2.29}$$

and similarly for the products of T_{R_4} with T_{R_2} and T_{R_3}, where we introduce constants a_{14}, a_{24} and a_{34} . Using the Ward identity (4.2.13), we can show that $a_{14} + a_{24} + a_{34}$

$= -\frac{c_{R_4}}{k} = -2\Delta_4$. Apart from the trivial solution, eq.(4.2.25) has two other solutions

$$\gamma_{14} = a_{14} \ , \ g(x) = 0 \ , \ f(x) = Cx^{a_{34}-\gamma_{34}} \ , \tag{4.2.30a}$$

$$\gamma_{14} = a_{14} - 1 \ , \ f(x) = 0 \ , \ g(x) = Cx^{a_{34}-\gamma_{34}-1} \ , \tag{4.2.30b}$$

We can always eliminate γ_{12} by absorbing it into a redefinition of the function f or g. Then, in the first case, eq.(4.2.30a), the exponents are determined to be:

$$\begin{gathered} \gamma_{14} = a_{14} \ , \ \gamma_{13} = -2\Delta_1 - a_{14} \ , \ \gamma_{34} = \Delta_1 + \Delta_2 - \Delta_3 - \Delta_4 \ , \\ \gamma_{24} = \Delta_3 - \Delta_2 - \Delta_4 - a_{14} \ , \ \gamma_{23} = \Delta_1 + \Delta_4 - \Delta_2 - \Delta_3 + a_{14} \ . \end{gathered} \tag{4.2.31}$$

In the second case, eq.(4.2.30b), they are given by eq.(4.2.31) if we make the substitution $a_{14} \to a_{14} - 1$. The constants a_{IJ} can be determined using group theory for each specific case. The evaluation of higher correlation functions proceeds in a similar manner.

Ordinary Wess-Zumino models at their critical point describe the critical behavior of quantum statistical chains with an arbitrary spin and continuous internal symmetry [9]. It would be interesting to see if some of these models are in fact supersymmetric, or if there are other critical systems that realize the semidirect product of the superconformal and the super Kac-Moody algebra.

4.3 Model building

Having studied the N=1 super Wess-Zumino model, we now turn to a discussion of how to build phenomenologically relevant models. As is well known, superstrings can only exist in ten space-time dimensions. In order to make contact with our four-dimensional world, the extra six dimensions have to be compactified to an internal space $\mathcal{K}$. So far, the most promising candidates for such a space $\mathcal{K}$ are the Calabi-Yau manifolds. Unfortunately, they are very hard to work with, as very little is known about their structure. Other promising candidates are the orbifolds [15] and torus compactification [16].

Another scheme of compactification is the one in which $\mathcal{K}$ is a group manifold. As we discussed in section 4.1, the interesting feature of this approach is that the model can be solved exactly, in a manner analogous to the treatment of critical systems by Belavin et al.[7]. In this section, we discuss the possibility of obtaining a consistent phenomenologically relevant model, by incorporating world-sheet supersymmetry. We study all possible Lie groups, and give a list of all the groups that can lead to consistent string theories. We also discuss the conditions that are needed for modular invariance. Unfortunately, simple GSO projections do not give rise to modular-invariant theories that possess a non-trivial massless spectrum in space-time dimension $d \geq 4$. In fact, world-sheet supersymmetry is broken in the Ramond sector of the group manifold, since there is no ground state of the Lorentz Kac-Moody algebra with the appropriate dimension. This leads to a breaking of space-time supersymmetry. We therefore discuss generalized GSO projections that may give rise to a massless sector. However, one then has to change the boundary conditions of the bosonic fields, making it more difficult to check modular invariance.

The models studied describe propagation of closed superstrings in $M^d \times G$, where M^d is a d-dimensional Minkowski space and G is a D-dimensional semi-simple Lie group. We consider both type-II and heterotic string models.† Our discussion includes models containing fermions that realize the supersymmetry non-linearly among themselves.

I. Type-II strings

The action of the models we are about to study is a sum of two terms. The first term describes propagation of closed superstrings in a d-dimensional Minkowski space. The second term describes an $N = 1$ supersymmetric Wess-Zumino model [18]. The world sheet is a compact surface. The degrees of freedom are bosonic coordinates $x^\mu(\sigma)$ ($\mu = 1, ..., d, \sigma = (\sigma_0, \sigma_1)$) and their fermionic super-partners $\psi^\mu(\sigma)$. The bosonic field g take values on a certain semi-simple Lie group G of dimension D.

† A partial study of type-II models has been done by E. Bergshoeff et al.[17]. Here, we generalize their discussion.

Their supersymmetric partners are free fermions, $\chi^a(\sigma)$ $(a = 1, ..., D)$. The operators in the $N = 1$ super Wess-Zumino model realize the super Kac-Moody algebra based on G. It is a minimal model in the sense that it contains a finite number of primary superfields transforming under a representation R of the Kac-Moody algebra with dimension $\Delta_R = \frac{c_R}{2k}$. The operator algebra closes properly, as discussed in the previous section [4,19].

The full theory is a direct product of a left- and a right-moving sector. We shall concentrate on the left-moving sector. The right-moving sector can be treated similarly. (Quantities in that sector will be denoted by a bar.) The theory is invariant under an $N = 1$ superconformal transformation generated by the operators

$$\begin{aligned} L_n =& \frac{1}{2}\sum_m : a^\mu_{n-m} a_{\mu m} : + \frac{1}{2k}\sum_m : \tilde{J}^a_{n-m}\tilde{J}^a_m : \\ &+ \frac{1}{2}\sum_s \left(s + \frac{1}{2}\right) : (\psi^\mu_{n-s}\psi_{\mu s} + \chi^a_{n-s}\chi^a_s) : \quad , \end{aligned} \tag{4.3.1a}$$

$$\begin{aligned} G_r =& -\sum_m : \psi^\mu_{r-m} a_{\mu m} : - \frac{1}{\sqrt{k}}\sum_m : \chi^a_{r-m}\tilde{J}^a_m : \\ &+ \frac{1}{6\sqrt{k}} i f^{abc} \sum_{m,s} : \chi^a_{r-m}\chi^b_{m-s}\chi^c_s : \quad , \end{aligned} \tag{4.3.1b}$$

where $\tilde{J}^a_m = J^a_m - \frac{i}{2} f^{abc}\sum_s : \chi^b_{m-s}\chi^c_s :$ is the "bosonic" current. The modes a^μ_m, J^a_m, ψ^μ_s and χ^a_s are the fourier components of x^μ, J^a, ψ^μ and χ^a, respectively. The current J has been defined in eq.(4.1.2). The indices r and s are half-integers when we are in the Neveu-Schwarz (NS) sector and they are integers if we are in the Ramond (R) sector.

These operators satisfy the following commutation relations:

$$[L_m, L_n] = (m - n)L_{m+n} + \frac{\hat{c}}{8}(m^3 - m)\delta_{m+n,0} \quad , \tag{4.3.2a}$$

$$[L_m, G_r] = \left(\frac{m}{2} - r\right) G_{m+r} \quad , \tag{4.3.2b}$$

$$\{G_r, G_s\} = 2L_{r+s} + \frac{\hat{c}}{2}\left(r^2 - \frac{1}{4}\right)\delta_{r+s,0} \quad , \tag{4.3.2c}$$

where $\hat{c} = d + \frac{2}{3}\frac{(2k-c_A)D}{2k} + \frac{1}{3}D$. For a conformally invariant theory, we need to have $\hat{c} = 10$. When G is a product of Lie groups G_i ($G = G_1 \times \cdots \times G_p$), of dimensions D_i, respectively, then there is an integer k_i ($i = 1, \ldots, p$) corresponding to each one of them. Thus, in general, the condition $\hat{c} = 10$ becomes

$$d + \frac{2}{3}\sum_{i=1}^{p}\left(\frac{(2k_i - c_A^i)D_i}{2k_i} + \frac{1}{2}D_i\right) = 10 \quad . \tag{4.3.3}$$

This is a constraint on the dimensionality of space-time and the group manifold. In Table 1 we list all possible groups that are solutions of eq.(4.3.3) in space-time dimension $d = 4$. We have omitted groups that only contain factors of $SU(2)$, because they cannot give phenomenologically relevant models.

Table 1 Conformally invariant type-II superstring models in $M^d \times G$, where M^d is a d-dimensional Minkowski space and G is a semi-simple Lie group. $\hat{k} = k - \frac{c_A}{2}$ is the level of the Kac-Moody algebra (cf. eq.(4.2.3b)). We have bose-fermi equivalence whenever $\hat{k} = 0$. An $SU(2)$ factor of level $\hat{k} = 0$ can be replaced by a $U(1)$.

G	$\hat{k}$
$SO(5)$	2
$SU(2) \times SU(2) \times SU(3)$	4, 0, 0
$SU(2) \times SU(2) \times SU(3)$	1, 1, 0
$SU(2) \times SU(2) \times SU(3)$	0, 0, 1
$SU(3)$	5
$SU(3) \times SO(5)$	0, 0
$SU(2) \times SU(4)$	0, 0
$SU(2) \times SO(5)$	0, 1
$SU(2) \times SU(2) \times SO(5)$	0, 1, 0
$SU(2) \times SU(3)$	2, 1

In order to find the groups that lead to consistent theories, one has to construct a modular-invariant partition function. Modular invariance is an important ingredient in string theories. In the Polyakov formalism, it corresponds to invariance of the theory under the modular (mapping-class) group of the respective Riemann surface. The modular transformations are globally non-trivial transformations leaving the conformal structure fixed. From the operator point of view, another reason for insisting on modular invariance is that it ensures locality in a superconformal theory when one combines the Neveu-Schwarz and Ramond sectors [20]. We will examine the models discussed above and determine the constraints coming from modular invariance on the torus. The discussion is easier in the light-cone gauge [21]. The partition function of a system defined on the torus is the product of the bosonic partition function and a linear combination of products of fermionic partition functions. The bosonic part is modular invariant by itself, through the same line of arguments developed in ref.[4]. Therefore, we only have to worry about the fermionic degrees of freedom.

The modular group of the torus is generated by the transformations $\tau \longrightarrow \tau + 1$ and $\tau \longrightarrow -\frac{1}{\tau}$, where τ is the modular parameter of the torus. For a Majorana-Weyl fermion there are four distinct partition functions on the torus:

$$Z_{--} \equiv \mathrm{Tr}_{(NS)} e^{2\pi i\tau(L_0-\frac{1}{48})} = (2\pi)^{1/6}\left[\frac{\theta_3(0|\tau)}{\theta_1'(0|\tau)^{1/3}}\right]^{1/2} \quad , \tag{4.3.4a}$$

$$Z_{-+} \equiv \mathrm{Tr}_{(NS)}(-)^F e^{2\pi i\tau(L_0-\frac{1}{48})} = (2\pi)^{1/6}\left[\frac{\theta_4(0|\tau)}{\theta_1'(0|\tau)^{1/3}}\right]^{1/2} \quad , \tag{4.3.4b}$$

$$Z_{+-} \equiv \mathrm{Tr}_{(R)} e^{2\pi i\tau(L_0+\frac{1}{24})} = (2\pi)^{1/6}\left[\frac{\theta_2(0|\tau)}{2\theta_1'(0|\tau)^{1/3}}\right]^{1/2} \quad , \tag{4.3.4c}$$

$$Z_{++} \equiv \mathrm{Tr}_{(R)}(-)^F e^{2\pi i\tau(L_0+\frac{1}{24})} = 0 \quad . \tag{4.3.4d}$$

The insertion of the fermion number operator $(-)^F$ changes the boundary conditions for the fermions in the time direction from antiperiodic to periodic. In the path-integral formalism, the four different partition functions (eq.(4.3.4)) correspond to the four different spin structures (boundary conditions in the σ_0 and σ_1 directions) of a Majorana-Weyl fermion on the torus.

Let us focus our attention on the left-moving sector. Considering the modular transformation properties of the θ-functions, we easily see that a group of N fermions with the same spin structure can only have a modular-invariant partition function if

N is a multiple of eight. In this case, the partition function is given by the standard GSO projection. More generally, we can define quantities:

$$F(\tau) = \mathrm{Tr}\mathcal{O}e^{2\pi i\tau :L_0:} \quad , \tag{4.3.5}$$

where we have inserted a modular-invariant operator $\mathcal{O}$. Then, in an obvious notation, the following expression is modular-invariant (up to a phase, which is canceled by the contribution of the right-moving sector):

$$F_N(\tau) = F_{--}(\tau) + (-)^n F_{-+}(\tau) + (-)^n F_{+-}(\tau) + \alpha F_{++}(\tau) \quad , \tag{4.3.6}$$

where $N = 8n$, n being an integer. The coefficient α can be 0 or ± 1. Factorization and modular invariance at two loops excludes the value $\alpha = 0$ [21]. These quantities, $F(\tau)$, contribute to a multi-loop expansion of correlation functions of vertex operators.

The factor α is relevant only when we have split the fermions in two or more groups. In that case, it differentiates between type-IIA and type-IIB theories. If each group contains N_i fermions, so that $\sum_i N_i = N$, then we can obtain a modular-invariant partition function by performing a GSO projection in each group separately. Then, eq.(4.3.6) can be replaced by a more general expression:

$$F_N(\tau) = \prod_i \{F^i_{--}(\tau) + (-)^{n_i} F^i_{-+}(\tau) + (-)^{n_i} F^i_{+-}(\tau) + \alpha_i F^i_{++}(\tau)\} \quad , \tag{4.3.7}$$

where $N_i = 8n_i$, n_i being an integer. Thus, in order to implement these projections, it is necessary that the number of fermions in both the left- and right-moving sectors be a multiple of 8.

We can relax this condition by considering modular-invariant partition functions that do not correspond to a product of independent GSO projections. They are linear combinations of products of partition functions in both the left- and right-moving sectors. Thus, a multitude of different possibilities emerge. It is not clear whether one can derive general constraints by considering the most general form of the partition function. Therefore, each case has to be studied separately.

When n is odd, the corresponding partition functions Z_{--} and Z_{-+} represent space-time bosons and Z_{+-} and Z_{++} space-time fermions, whereas when n is even, they all represent space-time bosons. In general, a set of N fermions generates highest-weight irreducible representations of the $SO(N)$ affine algebra. By splitting the fermions in groups and choosing independent spin structures, the $SO(N)$ symmetry gets broken down to $\prod_i SO(N_i)$.

To compute the spectrum of these theories, one first has to define the mass operator. We easily find that [17]

$$m^2 = L_0 + \bar{L}_0 - 1 + \frac{\epsilon_1 N_1 + \epsilon_2 N_2}{16} \quad , \tag{4.3.8}$$

where ϵ_i $(i = 1,2)$ is 0(1) if the N_i fermions are in the NS (R) sector. The ground states form representations of the Kac-Moody algebra of G with dimension $\Delta = \bar{\Delta} = \frac{c_R}{2k}$. It can be seen that under a standard GSO projection (i.e., if the first (second) group consists of the left (right) movers only), we cannot obtain a non-trivial massless sector. However, in any other case, we have to redefine the supersymmetry generators G_r (eq.(4.3.1b)), because not all fermionic modes can contribute. If we want the G_r's to obey the same commutation relations (eqs.(4.3.2b,c)), we have to twist the boundary conditions of the bosonic fields. This can be done consistently, provided that we can divide the group G by a subgroup H so that G/H is a symmetric space. Indeed, suppose that T^A are the generators of H $(A = 1, \ldots, \dim(H))$ and T^I $(I = \dim(H) + 1, \ldots, D)$ are the rest of the generators. Then, $f^{IJK} = f^{IAB} = 0$, and we can define a new supersymmetry generator whose modes will be:

$$\begin{aligned} G_r = & - \sum_m : \psi^\mu_{r-m} a_{\mu m} : - \frac{1}{\sqrt{k}} \sum_m : \left(\chi^A_{r-m} \tilde{J}^A_m + \chi^I_{r_m} \tilde{J}^I_m \right) : \\ & + \frac{1}{6\sqrt{k}} \sum_{m,s} : \left(f^{ABC} \chi^A_{r-m} \chi^B_{m-s} \chi^C_s + f^{AIJ} \chi^A_{r-m} \chi^I_{m-s} \chi^J_s \right) : \ , \end{aligned} \tag{4.3.9}$$

where we now allow twisted boundary conditions for the $\tilde{J}^I$'s, so that the G_r's obey unambiguous boundary conditions.

The generators of conformal symmetry (eq.(4.3.1a)) remain the same. However, the mass operator changes, owing to the fact that the bosons obeying anti-periodic boundary conditions contribute to the intercept an additional $-\frac{1}{24}$. Thus, eq.(4.3.8) is modified accordingly (in the sector where the $\tilde{J}^I$'s obey anti-periodic boundary conditions, the additional contribution is $\frac{(2k-c_A)D'}{96k}$, where $D' = D - \dim(H)$).

The question of modular invariance has now become more involved, because the partition function for bosons obeying twisted boundary conditions is different. Therefore, the arguments of ref.[4] are not readily applicable. However, in the case where the bosons are equivalent to fermions (i.e., $\hat{k} = k - \frac{c_A}{2} = 0$), one only has to consider the fermionic partition function. This has been done [12,23] and consistent theories have been obtained. Unfortunately, these models do not contain the standard $SU(3)\times SU(2)\times U(1)$ model. In fact, as Dixon et al. have argued [11], it is impossible to obtain the standard model with type-II strings. Twisting the boundary conditions for the bosonic fields offers a possibility of bypassing their argument.

In the case where the left and right sectors are treated in the same way, the models are trivially anomaly-free, since they are non-chiral. Even though the Ramond ground state transforms as a Weyl spinor, upon reduction to d-dimensions, both chiralities arise, a problem already encountered with Kaluza-Klein scenarios. A method of obtaining chiral fermions, by making use of asymmetric orbifolds, has been proposed in ref.[11]. An application of this method to models with $\hat{k} \neq 0$, may lead to theories that will contain the standard model.

We finally note that it is possible to enlarge the list of Lie groups (Table 1) if we allow part of the internal degrees of freedom to form a Virasoro algebra with central charge $\hat{c} < 1$. All of these models have been classified and completely solved [24]. It can be shown that the only possible values of $\hat{c}$ are:

$$\hat{c} = 1 - \frac{8}{m(m+2)} , \tag{4.3.10}$$

where m is an integer ($m \geq 4$). This approach has been studied by Dixon, et al.[11]. They were able to obtain consistent models with gauge groups G_2, $SU(3) \times SU(3)$

and proper subgroups of the groups listed in Table 1.

II. Heterotic strings

We now turn to a discussion of heterotic strings on group manifolds.‡ The left-moving sector consists of superstrings propagating in $M^d \times G$. We shall write G as the product of two Lie groups: $G = G_1 \times G_2$, of dimensions D_1 and D_2, respectively ($D_1 + D_2 = D$). The components of G_1 have non-vanishing levels, whereas G_2 corresponds to level $k = 0$. The right-moving sector consists of bosonic strings propagating in $M^{d+\bar{d}} \times G_1$, where the extra $\bar{d}$ dimensions curl up to form an internal space through toroidal compactification. We shall use the equivalence of bosons and fermions in two dimensions to replace the extra $\bar{d}$ bosonic fields by $2\bar{d}$ fermions.

Thus, the left-moving sector is identical to the one described for the type-II strings. In the right-moving sector, however, the fourier components of the energy-momentum tensor are

$$\bar{L}_n = \frac{1}{2}\sum_m : \bar{a}^\mu_{n-m}\bar{a}_{\mu m} : + \frac{1}{2k}\sum_m : \bar{J}^a_{n-m}\bar{J}^a_m : + \frac{1}{2}\sum_s \left(s + \frac{1}{2}\right) : \bar{\lambda}^I_{n-s}\bar{\lambda}^I_s : \quad , \tag{4.3.11}$$

where $\bar{\lambda}^I_s$ ($I = 1, ..., 2\bar{d}$) are the fourier components of the $2\bar{d}$ Majorana fermions corresponding to the curled-up dimensions, and $a = 1, \ldots, D_1$. These operators satisfy the commutation relations

$$[\bar{L}_m, \bar{L}_n] = (m - n)\bar{L}_{m+n} + \frac{\bar{c}}{12}(m^3 - m)\delta_{m+n,0} \quad , \tag{4.3.12}$$

where $\bar{c} = d + \bar{d} + \frac{(2k - c_A)D_1}{2k}$. Since the right-moving sector contains only bosonic strings, the condition for conformal invariance is $\bar{c} = 26$, i.e.,

$$26 = d + \bar{d} + \frac{(2k - c_A)D_1}{2k} \quad . \tag{4.3.13}$$

For conformal invariance of the total theory, both eq.(4.3.3) and eq.(4.3.13) have to

‡ Sezgin has studied heterotic strings on group manifolds [22], and found Lie groups that lead to conformally invariant models. Here, we extend his analysis by studying modular invariance of these models.

be satisfied.* Using these two equations, we find

$$2\bar{d} = d + D + 22 \quad . \tag{4.3.14}$$

A complete list of groups that lead to conformally invariant theories is given in Table 2.

Table 2 Conformally invariant heterotic string models in which the left-movers propagate in $M^d \times G_1 \times G_2$, M^d being a d-dimensional Minkowski space and G_1, G_2 groups of vanishing and non-vanishing level, respectively. The fourth column shows the rank of the group K, formed by the compactified dimensions in the right-moving sector.

G_1	G_2	rank(K)
$SO(5)$	–	18
$SU(2)$	$SU(2) \times SU(3)$	20
$SU(2) \times SU(2)$	$SU(3)$	20
$SU(3)$	$SU(2) \times SU(2)$	20
$SU(3)$	–	17
–	$SU(3) \times SO(5)$	22
–	$SU(2) \times SU(4)$	22
$SU(2) \times U(1)$	$SO(5)$	20
$SU(2)$	$SO(5) \times SU(2)$	21
$SU(2) \times U(1)$	$SU(3)$	19
$SU(3) \times U(1)$	$SU(2)$	19
$SU(3) \times U(1) \times U(1)$	–	18
$U(1)$	$SU(4)$	21

Using the GSO projections (eq.(4.3.7)), we see that a necessary condition for modular invariance is that the total number of fermions in both the left- and right-moving sectors be a multiple of 8. Unfortunately, it is impossible to find a theory with

* When the group G_1 is a product of semi-simple Lie groups, then the last term on the right-hand side of eq.(4.3.13) has to be replaced by a sum of similar terms over the components (cf. eq.(4.3.3)).

a non-trivial massless spectrum if this additional constraint is satisfied. We therefore have to consider projections that mix the two sectors.

The extra coordinates in the right-moving sector are compactified to a manifold K (see Table 2). One has to choose appropriate boundary conditions for the fermions comprising K, to obtain a modular-invariant model. Certain groups have been studied by Kawai et al.[16], and consistent theories have emerged. They only studied level $\hat{k} = 0$ cases. In the other cases, the complications we encountered in type-II strings (having to twist the boundary conditions of the bosonic fields) will arise.

In all the models, vertex operators can be constructed in the standard way. Tree and loop amplitudes can easily be computed in the operator formalism. We can also construct open string models. (It is possible to construct the Wess-Zumino term even with open string boundary conditions.)[♮] In order to obtain gauge fields, though, one has to introduce Chan-Paton factors.

♮ We are indebted to A. P. Polychronakos for enlightening us on this point.

References

1. M. B. Green and J. H. Schwarz, *Phys. Lett.* **149B** (1984) 117.

2. J. H. Schwarz, Caltech preprint (1987), CALT-68-1432.

3. V. Knizhnik and A. B. Zamolodchikov, *Nucl. Phys.* **B247** (1984) 83.

4. D. Gepner and E. Witten, Princeton preprint, April 1986.

5. J. Wess and B. Zumino, *Phys. Lett.* **37B** (1971) 95.

6. E. Witten, *Comm. Math. Phys.* **92** (1984) 451;

   V. G. Kac, "*Infinite-dimensional Lie Algebras,*" Birkhäuser, Boston 1983;

   E. J. Lepowsky, "*Vertex Operators in Mathematics and Physics,*" Springer-Verlag, New York 1984.

7. A. A. Belavin, A. M. Polyakov and A. B. Zamolodchikov, *Nucl. Phys.* **B241** (1984) 333.

8. D. I. Olive, Imperial College preprint (1985), ICTP-84-85-14.

9. I. Affleck, *Nucl. Phys.* **B265** [FS15] (1986) 409;

   E. B. Kiritsis, Caltech preprint (1987), CALT-68-1424.

10. D. J. Gross, J. Harvey, E. J. Martinec and R. Rohm, *Nucl. Phys.* **B256** (1985) 253.

11. L. J. Dixon, V. S. Kaplunovsky and C. Vafa, SLAC preprint (1987), SLAC-PUB-4282.

12. T. L. Curtright and C. K. Zachos, *Phys. Rev. Lett.* **53** (1984) 1799;

    E. Abdalla and M. C. B. Abdalla, *Phys. Lett.* **152B** (1985) 59;

    R. Rohm, *Phys. Rev.* **D32** (1985) 2849;

13. Z. Qiu, *Nucl. Phys.* **B270** [FS16] (1986) 205.

14. E. B. Kiritsis and G. Siopsis, Caltech preprint (1987), CALT-68-1395.

15. L. J. Dixon, J. A. Harvey, C. Vafa and E. Witten, *Nucl. Phys.* **B261** (1985) 678; *Nucl. Phys.* **B274** (1986) 285;

S. Hamidi and C. Vafa, *Nucl. Phys.* **B279** (1987) 465.

16. R. Bluhm, L. Dolan and P. Goddard, Rockefeller Univ. preprint;

H. Kawai, D. C. Lewellen and S.-H. H. Tye, *Phys. Rev. Lett.* **57** (1986) 1832;

Cornell preprints CLNS 86/751 (1986) and CLNS 87/760 (1987).

17. E. Bergshoeff, S. Randjbar-Daemi, A. Salam, H. Sarmadi and E. Sezgin, *Nucl. Phys.* **B269** (1986) 77.

18. P. Di Vecchia, V. G. Knizhnik, J. L. Petersen and P. Rossi, *Nucl. Phys.* **B253** (1985) 701.

13. Z. Qiu, *Nucl. Phys.* **B270** [FS16] (1986) 205.

19. E. B. Kiritsis and G. Siopsis, *Phys. Lett.* **184B** (1987) 353;

J. Fuchs, *Nucl. Phys.* **B286** (1987) 455.

20. D. Friedan, E. Martinec and S. Shenker, *Nucl. Phys.* **B271** (1986) 93.

21. N. Seiberg and E. Witten, *Nucl. Phys.* **B276** (1986) 272.

22. E. Sezgin, *Trieste H. E. Workshop* 1985, 350.

23. I. Antoniadis, C. Bachas, C. Kounnas and P. Windey, *Phys. Lett.* **171B** (1986) 51.

24. D. Friedan, Z. Qiu and S. Shenker, *Phys. Rev. Lett.* **52** (1984) 1575.

5. String field theory

5.1 Interactions of strings

There are two ways to describe interactions of strings. The first is by considering Riemann surfaces of arbitrary genus as world-sheets [1]. The total partition function is then a sum over all Riemann surfaces,

$$Z = \sum_g c(g) \int_g \mathcal{D}g^{ab}\mathcal{D}x^\mu e^{-S_a[x^\mu, g^{ab}]} \ , \tag{5.1.1}$$

where

$$S_a[x^\mu, g^{ab}] = \int_{\Sigma_g} d^2\sigma \sqrt{-\det g^{ab}}\, g^{ab}\partial_a x^\mu \partial_b x_\mu \ , \tag{5.1.2}$$

and we integrate over surfaces Σ_g (representing the world-sheet) that are topologically equivalent (of the same genus g). We then have to sum over all topologies, weighed by coefficients $c(g)$ that depend on the genus.†

The second approach is by formulating a string field theory. The central problem there is to find a suitable interaction term in the action. Such a term was first proposed by Kaku and Kikkawa, and Cremmer and Gervais [3]. However, they worked in the light-cone gauge in which the required symmetries of the theory were not manifest. A covariant approach has been advocated by Witten [4] for the case of open bosonic strings. This section is devoted to a description of his method.

Before building a field theory, it is necessary to develop a satisfactory first-quantized formalism. We start with a list of definitions to fix our notation. We expand the coordinates $x^\mu(\sigma)$, the ghost $c(\sigma)$ and the anti-ghost $b(\sigma)$ ($\mu = 0, 1, ..., 25, \sigma \in [0, \pi]$) in modes a_n^μ, c_n and b_n, respectively, satisfying (anti-)commutation relations

$$[a_n^\mu, a_m^\nu] = n\delta_{m+n,0}\eta^{\mu\nu} \ , \tag{5.1.3a}$$

† It has recently been shown [2] that if $c(g) = e^{-\chi(g)}$, where $\chi(g) = \frac{1}{4\pi}\int_{\Sigma_g}\sqrt{-\det g^{ab}}R$ is the Euler characteristic (R is the curvature of the surface Σ_g), then the theory is unitary.

$$\{b_m, c_n\} = \delta_{m+n,0} \ , \tag{5.1.3b}$$

with all other (anti-)commutators vanishing. a_0^μ is related to the center-of-mass momentum by $a_0^\mu = \sqrt{2}p^\mu$. Thus, if q^μ is the position of the center of mass, then $[q^\mu, p^\nu] = i\eta^{\mu\nu}$. The vacuum is defined as an eigenstate of a_0^μ: $a_0^\mu|p\rangle = \sqrt{2}p^\mu|p\rangle$. Modes with positive (negative) indices are creation (annihilation) operators. It is convenient to separate the zero modes of the ghost (c_0) and anti-ghost (b_0) from the rest of the string modes. They act on a two-dimensional space spanned by the kets $|+\rangle$ and $|-\rangle$, where $b_0|+\rangle = |-\rangle$, $c_0|-\rangle = |+\rangle$, $b_0|-\rangle = c_0|+\rangle = 0$. We also define bras $\langle+|$ and $\langle-|$ that are annihilated by negative-index modes. An inner-product is defined by $\langle+|-\rangle = \langle-|+\rangle = 1$, $\langle+|+\rangle = \langle-|-\rangle = 0$. Thus, a general state $|A\rangle$ can be written as $|A\rangle = |A_+\rangle + |A_-\rangle$, where $|A_+\rangle$ ($|A_-\rangle$) is constructed from the $|+\rangle$ ($|-\rangle$) vacuum. We can also define a functional $A[z(\sigma)]$, where $z = (x^\mu, b, c)$, by $A[z(\sigma)] = \langle z|A\rangle$. A $\mathbf{Z}_2$ grading is imposed on the string fields by assigning a number $(-)^{|A|}$ to each functional A, where $GA = (|A| - \frac{1}{2})A$, G being the ghost-number operator:

$$G = c_0 b_0 + \sum_{n\neq 0} : c_{-n} b_n : -\frac{1}{2} \ . \tag{5.1.4}$$

Instead of the fermion ghost fields, $b(\sigma)$ and $c(\sigma)$, one can introduce a single bosonic field $\phi(\sigma)$. The connection with the fermionic fields is

$$c(\sigma) =: e^{i\phi(\sigma)} : \ , \quad b(\sigma) =: e^{-i\phi(\sigma)} : \ . \tag{5.1.5}$$

The action for $\phi(\sigma)$ is [4]

$$S_b[\phi] = \frac{1}{2\pi} \int d^2\sigma \sqrt{\det g^{ab}} (g^{ab} \partial_a \phi \partial_b \phi - 3iR\phi) \ , \tag{5.1.6}$$

where R is the curvature of the world-sheet (cf. eq.(5.1.2) for the coordinates x^μ).

Next, we introduce the BRS operator

$$Q = \sum L_m c_{-m} - \frac{1}{2} \sum (m-n) : c_{-m} c_{-n} b_{m+n} : -c_0 \ , \qquad (5.1.7)$$

where L_m is the fourier component of the stress-energy tensor (cf. eq.(3.5.24)). Q can also be written as $Q = \int_0^\pi d\sigma j^0$, where j^a is a conserved current [4]:

$$D_a j^a = 0 \ . \qquad (5.1.8)$$

Physical states obey the condition $Q|A\rangle = 0$ and we identify states that differ by $Q|B\rangle$, for some state $|B\rangle$. We also require that they have ghost number $G = -\frac{1}{2}$ (therefore, they are bosons under our $\mathbf{Z}_2$ grading). This is a consistent formalism, provided that $Q^2 = 0$. As Kato and Ogawa [5] found,

$$Q^2 = \frac{D-26}{24} \sum (m^3 - m) : c_m c_{-m} : \ . \qquad (5.1.9)$$

Therefore, a necessary consistency condition is $D = 26$, as expected.

Passing to Second Quantization, we observe that the equation $Q|A\rangle = 0$ can be viewed as an equation of motion coming from the action

$$S_0 = \langle A|Q|A\rangle \ . \qquad (5.1.10)$$

We wish to write this as the integral of some quantity B. Consider a surface Σ bounded by line elements S_1 and S_2, perpendicular to each other. Assume further that Σ is flat near S_1 and S_2.§ On S_1 we choose free-string boundary conditions. On S_2 we choose data specified by $z(\sigma)$, where $z = (x^\mu, \phi)$. We then define $\int A$ by

$$\int A = \int \mathcal{D}z A[z] \int_\Sigma \mathcal{D}z' e^{-S[z']} \ , \qquad (5.1.11)$$

where $S[z] = S_a[x^\mu] + S_b[\phi]$ (cf. eqs.(5.1.2) and (5.1.6)). It is easy to check that

§ An example of such a surface is a quarter-sphere bounded by semi-circles S_1 and S_2.

$\int QA = 0$, by using eq.(5.1.8). It follows that

$$\int \mathcal{D}zA[z] \int\limits_{\Sigma} \mathcal{D}z' e^{-S[z']} \int\limits_{\Sigma} d^2\sigma D_a j^a = 0 \ . \tag{5.1.12}$$

Using Stoke's theorem, we can write $\int_\Sigma D_a j^a$ as a line integral of j^0 along the boundary (since the time direction is perpendicular to S_1 and S_2, which form the boundary). Therefore, $\int_\Sigma D_a j^a$ can be replaced by $Q = \int_0^\pi d\sigma j^0$ in eq.(5.1.12). It then follows from eq.(5.1.12) that

$$\int \mathcal{D}zQA[z] \int\limits_{\Sigma} \mathcal{D}z' e^{-S[z']} \int\limits_{\Sigma} d^2\sigma D_a j^a = 0 \ . \tag{5.1.13}$$

But the left-hand side of eq.(5.1.13) is just $\int QA$. Therefore,

$$\int QA = 0 \ . \tag{5.1.14}$$

Notice that the above arguments did not depend on the shape of the surface Σ. The only requirement was that S_1 and S_2 intersect at right angles and that Σ be flat near S_1 and S_2. It will be necessary, however, to choose a specific metric on Σ. We shall choose a metric that is flat everywhere apart from a singular point, where all the curvature is concentrated. Furthermore, we shall only be interested in the limit where the length of S_1 goes to zero. In that limit, Σ degenerates to S_2, where the left half of S_2 is identified with its right half. The singular point is then its mid-point and eq.(5.1.11) reduces to

$$\int A = \int \mathcal{D}x^\mu \mathcal{D}\phi e^{-\frac{3}{2}i\phi(\frac{\pi}{2})} \prod_{\mu=0}^{25} \prod_{\sigma=0}^{\pi/2} \delta(x^\mu(\pi-\sigma) - x^\mu(\sigma))\delta(\phi(\pi-\sigma) - \phi(\sigma)) \ . \tag{5.1.15}$$

The factor $e^{-\frac{3}{2}i\phi(\frac{\pi}{2})}$ comes from the term $e^{-\frac{3i}{2\pi}\int R\phi}$ in the ghost action (eq.(5.1.6)). The curvature is a δ-function at the singular point proportional to the deficit angle, which in this case is π.

In terms of oscillators, $\int A = \langle I|A\rangle$, where the "identity" state $|I\rangle$ is [6]

$$|I\rangle = b_+ b_- \exp\left\{\frac{1}{2}\sum_{n=1}^{\infty}\frac{(-)^n}{n}a^{\mu}_{-n}a^{\mu}_{-n}\right\}\exp\left\{\sum_{n=1}^{\infty}(-)^n b_{-n}c_{-n}\right\}|+\rangle \ , \qquad (5.1.16)$$

where $b_\pm = \sum(\pm i)^n b_n$. The factor b_+b_- is just the fermionic form of the insertion $e^{\frac{3}{2}i\phi(\frac{\pi}{2})}$ in eq.(5.1.15). We can check that $Q|I\rangle = 0$, which confirms that $\int QA = 0$.

Since Q acts as a derivation as far as integration is concerned, it is natural to try to find a multiplication operation between states that will obey Leibnitz's rule:

$$Q(A \star B) = (QA)\star B + (-)^{|A|}A\star QB \ . \qquad (5.1.17)$$

The $\star$ operation is defined as follows. Consider a surface Σ bounded by a hexagon with sides $S_1, \ldots, S_6$ that are straight lines. We also require that adjacent sides be perpendicular and that Σ be flat near the sides. On S_2, S_4 and S_6, we choose free-string boundary conditions. On S_1, S_3 and S_5, we choose data specified by $z_1(\sigma)$, $z_2(\sigma)$ and $z(\sigma)$, respectively. Then,

$$(A\star B)[z(\sigma)] = \int \mathcal{D}z_1 A[z_1]\int \mathcal{D}z_2 B[z_2]\int_{\Sigma} \mathcal{D}z' e^{-S[z']} \ . \qquad (5.1.18)$$

To prove Leibnitz' rule, we work as before. Using eq.(5.1.8), we obtain

$$\int \mathcal{D}z_1 A[z_1]\int \mathcal{D}z_2 B[z_2]\int_{\Sigma} \mathcal{D}z' e^{-S[z']}\int_{\Sigma} d^2\sigma D_a j^a = 0 \ . \qquad (5.1.19)$$

The last integral can be converted to a surface integral of j^0. The sides S_2 and S_4 give the first and second terms on the right-hand side of (5.1.17), respectively, whereas S_6 gives the left-hand side of (5.1.17). Thus, eq.(5.1.19) is equivalent to (5.1.17).

To proceed further, a particular choice of metric has to be made for Σ. As before, we choose a metric that is flat everywhere, except at singular points. We then shrink the sides S_2, S_4 and S_6 to zero length. The resulting degenerate surface has the shape of a **Y**, where the left half of a side has been identified with the right half of the next side that survived. The singular point is then the common mid-point of the three sides. Thus, we find

$$(A\star B)[z(\sigma)] = \int \mathcal{D}z_1\mathcal{D}z_2 A[z_1]B[z_2]e^{\frac{3}{2}i\phi(\frac{\pi}{2})} \times \prod_{\sigma=0}^{\pi/2} \delta(z_1(\pi-\sigma)-z_2(\sigma))\delta(z_2(\pi-\sigma)-z(\sigma))\delta(z(\pi-\sigma)-z_1(\sigma)) \ . \tag{5.1.20}$$

Notice that this time we have an insertion of $e^{\frac{3}{2}i\phi(\frac{\pi}{2})}$, because the deficit angle for the hexagon is $-\pi$ (cf. the case of a quarter sphere (eq.(5.1.15), where the deficit angle is π).

In this singular limit, we can prove the following two extra properties:♮

$$\int A\star B = \int B\star A \ , \tag{5.1.21a}$$

$$(A\star B)\star C = A\star(B\star C) \ . \tag{5.1.21b}$$

In terms of oscillators, $|A\star B\rangle_1 = {}_1\langle A|_2\langle B|V_3\rangle_{123}$, where the three-string vertex is given by

$$|V_3\rangle_{123} = \int d^{26}p_1 d^{26}p_2 d^{26}p_3 \delta(p_1+p_2+p_3)\exp\left\{\sum_{r,s=1}^{3}\sum_{m,n=0}^{\infty} b^r_{-m}\tilde{N}^{rs}_{mn} n c^s_{-n}\right\} \times \exp\left\{\frac{1}{2}\sum_{r,s=1}^{3}\sum_{m,n=0}^{\infty} a^{\mu,r}_{-m}N^{rs}_{mn}a^{\mu,s}_{-n}\right\}|p_1,+\rangle_1|p_2+\rangle_2|p_3+\rangle_3 \ . \tag{5.1.22}$$

The explicit form of the Neumann coefficients N^{rs}_{mn} and $\tilde{N}^{rs}_{mn}$ can be found in ref.[6].

♮ Making this singular choice for the metric is a necessary ingredient in the proof of eqs.(5.1.21a) and (5.1.21b). However, the proof is still not free of criticism [7]. One can of course try to build a theory without making use of these equations. We shall not do this here, as things becomes considerably more complicated, without an apparent reward. Instead, we shall proceed by assuming the validity of these equations.

It is useful to define a two-string vertex by ${|V_2\rangle}_{12} = {}_3{\langle I|V_3\rangle}_{123}$. Explicitly,

$$
\begin{aligned}
{|V_2\rangle}_{12} = \int d^{26}p_1 d^{26}p_2 \delta(p_1+p_2) \exp\left\{\sum_{n=1}^{\infty} \frac{(-)^n}{n} a_{-n}^{\mu,1} a_{-n}^{\mu,2}\right\} \\
\times \exp\left\{\sum_{n=1}^{\infty}(-)^n (b_{-n}^1 c_{-n}^2 + b_{-n}^2 c_{-n}^1)\right\} \left({|p_1,+\rangle}_1 {|p_2,-\rangle}_2 + {|p_1,-\rangle}_1 {|p_2,+\rangle}_2\right) \ .
\end{aligned}
\tag{5.1.23}
$$

It also follows that $\int A \star B = {}_1{\langle A|}_2{\langle B|V_2\rangle}_{12}$. Because of the special form of the vacuum in ${|V_2\rangle}_{12}$ (eq.(5.1.23)), we have $\int A_+ \star B_+ = \int A_- \star B_- = 0$.

Having assembled all the necessary ingredients, it is now possible to define an action that will describe interactions among strings. The action is a Chern-Simons form [4]:

$$
S = \int \frac{1}{2} A \star QA + \frac{1}{3} A \star A \star A \ , \tag{5.1.24}
$$

and is invariant under the transformation

$$
\delta A = QA + A \star \epsilon - \epsilon \star A \ . \tag{5.1.25}
$$

The remainder of this chapter is devoted to the quantization of this action.

5.2 Lagrangian quantization

In this section we describe how to quantize Witten's string field theory using the lagrangian formalism and point out the shortcomings of the method. The ideal approach appears to be a generalization of the Faddeev-Popov procedure introduced by Fradkin et al.[8]. We shall first discuss the procedure in general, and then apply it to Yang-Mills theories and strings. In the following section we shall apply the Faddeev-Popov procedure that we discussed in section 2.1 to string field theory.

Suppose we are given a set of bosonic and fermionic fields Φ^A, where A can be both a discrete and a continuous index. Let Φ^*_A be a new set of fields of opposite statistics. In the phase space of Φ and Φ^* we define "anti-brackets" of functions $F(\Phi, \Phi^*)$ and $G(\Phi, \Phi^*)$ by

$$(F, G) = \frac{\partial F}{\partial \Phi^A} \frac{\partial G}{\partial \Phi^*_A} - \frac{\partial F}{\partial \Phi^*_A} \frac{\partial G}{\partial \Phi^A} \ . \tag{5.2.1}$$

The properties of anti-brackets do not coincide with those of Poisson brackets. The most bothersome property is that under an infinitesimal canonical transformation,

$$\delta \Phi^A = (\Phi^A, \Gamma)\alpha \ , \ \delta \Phi^*_A = (\Phi^*_A, \Gamma)\alpha \ , \tag{5.2.2}$$

the volume element of phase space changes by $2\Delta\Gamma\alpha$, where

$$\Delta = \sum_A \epsilon(\Phi^A) \frac{\partial^2}{\partial \Phi^A \partial \Phi^*_A} \ , \tag{5.2.3}$$

and $\epsilon(\Phi^A) = +(-)$ if Φ^A is a boson (fermion). We define dynamics through an action $W(\Phi, \Phi^*)$ that satisfies

$$(W, W) = 0 \ . \tag{5.2.4}$$

This is called the "master equation." Once W is found, one can proceed to define the quantum theory. This is done as follows. We introduce an action $\tilde{W}$ (to be specified later) and we choose an arbitrary fermionic function $\Psi(\Phi)$. We then define a path-integral by

$$Z_\Psi = \int \mathcal{D}\Phi e^{i\tilde{W}(\Phi, \frac{\partial \Psi}{\partial \Phi})} \ . \tag{5.2.5}$$

Under the BRS transformation:

$$\delta \Phi = (\Phi, \tilde{W})\Big|_{\Phi^* = \frac{\partial \Psi}{\partial \Phi}} \cdot \alpha \ , \tag{5.2.6}$$

where α is a fermionic variable, we have $\mathcal{D}\Phi \to \mathcal{D}\Phi(1 + 2i\Delta\tilde{W}\alpha)$, because of eq.(5.2.3), and $\delta\tilde{W} = (\tilde{W}, \tilde{W})\alpha$. Therefore, Z_Ψ is invariant under a BRS transformation, provided

$$\frac{1}{2}(\tilde{W},\tilde{W}) = i\Delta\tilde{W} \ . \tag{5.2.7}$$

Also, by choosing $\alpha = i\delta\Psi$ (gauge transformation), eq.(5.2.7) implies that Z_Ψ is independent of Ψ. If the measure is chosen to be BRS invariant, then eq.(5.2.7) reduces to eq.(5.2.4) and therefore we can set $\tilde{W} = W$. Otherwise, $\tilde{W}$ has to be expanded as a series in $\hbar$ (which we have suppressed) and eq.(5.2.7) becomes a recursion relation to determine the terms in the expansion of $\tilde{W}$.

One way to find the appropriate measure is by using the hamiltonian formalism, in which case the volume of phase space is invariant under canonical transformations. This can easily be done for gauge theories. The case of strings will be the subject of the next section, where an alternate approach avoiding this ambiguity will be discussed. For the remainder of this section, we shall ignore this problem. Thus, we assume that $\tilde{W} = W$.

We want W to correspond to a classical system described by an action $S(A)$, a function of the fields A^i ($i = 1,\dots,n$). Suppose that there are m constraints among these fields. We introduce ghosts η^a ($a = 1,\dots,m$), and we let the set of Φ^A's be the set $\{A^i, \eta^a\}$.‡ Then W is a function of A^i, η^a, A_i^* and η_a^*. However, this W is not unique. We can add extra fields λ^I and $\bar{\eta}^I$ and their partners λ_I^* and $\bar{\eta}_I^*$. The action then becomes

$$W' = W(A^i, \eta^a; A_i^*, \eta_a^*) + \lambda^I \bar{\eta}_I^* \ . \tag{5.2.8}$$

It is easily seen that W' satisfies $(W', W') = 0$.

As an example, consider a Yang-Mills theory with group G. Let f^{abc} be the

‡ It can be shown that this is the minimal choice of coordinates. Notice that there are $n - m$ degrees of freedom in the classical system. The system of Φ^A's appears to have $n + m$ degrees of freedom. However, this is not true, because the ghosts are effectively negative degrees of freedom. Thus, by introducing m of them, we reduce the number of degrees of freedom from n to $n - m$.

structure constants of G. We define W by

$$W = \int \frac{1}{2g^2}\mathrm{tr}(F_{\mu\nu}F_{\mu\nu}) + \mathrm{tr}(A^*_\mu D_\mu \eta) + f^{abc}\eta^*_a\eta_b\eta_c + \mathrm{tr}(\lambda\bar{\eta}^*_a) \ , \tag{5.2.9}$$

where, as usual, $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu + i[A_\mu, A_\nu]$. The transformation (5.2.6) is the ordinary BRS transformation:

$$\delta A_\mu = D_\mu\eta \cdot \alpha \ , \ \ \delta\eta^a = f^{abc}\eta_b\eta_c \cdot \alpha \ . \tag{5.2.10}$$

Let us now make the following choice:

$$\Psi = \mathrm{tr}(\bar{\eta}\partial_i A_i) \ . \tag{5.2.11}$$

Then, $A^* = \frac{\delta\Psi}{\delta A} = \partial\bar{\eta}$, $\eta^* = \frac{\delta\Psi}{\delta\eta} = 0$, $\bar{\eta}^* = \frac{\delta\Psi}{\delta\bar{\eta}} = \partial_i A_i$, and therefore

$$W(\Phi, \frac{\delta\Psi}{\delta\Phi}) = \int \frac{1}{2g^2}\mathrm{tr}F^2 + \mathrm{tr}(\bar{\eta}\partial_\mu D_\mu\eta) + \mathrm{tr}(\lambda\partial_i A_i) \ , \tag{5.2.12}$$

which is the Yang-Mills action in the Coulomb gauge. η and $\bar{\eta}$ are the ordinary Faddeev-Popov ghosts and λ is a Lagrange multiplier enforcing the constraint $\partial_i A_i = 0$.

In string theory, we let $\tilde{\Phi}$ be any Grassman odd string state and $\tilde{\Phi}^*$ its partner. We also introduce Lagrange multipliers λ and additional fields $\bar{\eta}$ and $\bar{\eta}^*$. We postulate the action [9] [10]

$$W = \int \frac{1}{2}\Phi \star Q\Phi + \frac{1}{3}\Phi \star \Phi \star \Phi + \lambda \star \bar{\eta}^* \ , \tag{5.2.13}$$

where $\Phi = \tilde{\Phi} + \tilde{\Phi}^*$. W reduces to Witten's action (eq.(5.1.16)) if $\tilde{\Phi}^* = \bar{\eta}^* = 0$ and $\tilde{\Phi}$ is restricted to ghost number $-1/2$. To check that $(W, W) = 0$, we can ignore the last term, because it has been designed to satisfy this condition. Also, $\frac{\delta W}{\delta \tilde{\Phi}} = \frac{\delta W}{\delta \tilde{\Phi}^*} = Q\Phi + \Phi \star \Phi$. Thus,

$$\begin{aligned}(W, W) &= \int (Q\Phi + \Phi \star \Phi) \star (Q\Phi + \Phi \star \Phi) \\ &= \int \left(\frac{1}{2} Q\Phi \star Q\Phi + Q\Phi \star \Phi \star \Phi + \frac{1}{2} \Phi \star \Phi \star \Phi \star \Phi \right)\end{aligned} \quad . \tag{5.2.14}$$

The first term vanishes because $Q^2 = 0$ (eq.(5.1.3)). The second term can be written as $\frac{1}{3} \int Q(\Phi \star \Phi \star \Phi)$ and therefore vanishes because of eq.(5.1.7). By using eq.(5.1.13a) and the fact that both Φ and $\Phi \star \Phi \star \Phi$ are fermions, we see that the third term also vanishes. Therefore, W satisfies eq.(5.2.4). By choosing

$$\Psi = \int b_0 \tilde{\Phi} \star \bar{\eta} \ , \tag{5.2.15}$$

we have $\tilde{\Phi}^* = \frac{\delta \Psi}{\delta \tilde{\Phi}} = b_0 \bar{\eta}$, which is equivalent to $b_0 \tilde{\Phi}^* = 0$ (since $\bar{\eta}$ does not appear in the action). Also, $\bar{\eta}^* = \frac{\delta \Psi}{\delta \bar{\eta}} = b_0 \tilde{\Phi}$. Therefore, the action becomes

$$W = \int \frac{1}{2} \Phi \star Q\Phi + \frac{1}{3} \Phi \star \Phi \star \Phi + \lambda \star b_0 \Phi \ , \tag{5.2.16}$$

which can serve as the starting point of perturbation theory. The last term simply enforces the constraint $b_0 \Phi = 0$, which is known as the Siegel gauge. An explicit calculation in four-tachyon scattering will be presented at the end of the next section.

5.3 Hamiltonian formalism

We now turn to a discussion of the quantization of Witten's sting field theory [4] via the hamiltonian formalism [11]. The action is $S = \int \mathcal{L}$, where the lagrangian density $\mathcal{L}$ is defined by [4]

$$\mathcal{L} = \frac{1}{2} A \star QA + \frac{1}{3} A \star A \star A \ . \tag{5.3.1}$$

This is different from the usual density in space-time; it is defined in the space of string modes. To obtain the momenta that are conjugate to the "coordinates" $A[z(\sigma)]$, we define the center-of-mass time-coordinate $q^0 = \int_0^\pi d\sigma x^0(\sigma)$ as the time of the system. Then the momenta are given by

$$P \equiv \frac{\delta \mathcal{L}}{\delta(a_0^0 A)} \ . \tag{5.3.2}$$

Although $|V_3\rangle_{123}$ (eq.(5.1.14)) contains arbitrarily high orders of the time derivative, a_0^0, the interaction term in the action, $S_{int} \equiv \frac{1}{3}\int A \star A \star A$, This can be seen by writing S_{int} in terms of functionals:

$$S_{int} = \int \mathcal{D}z_1 \mathcal{D}z_2 \mathcal{D}z_3 V_3[z_1(\sigma), z_2(\sigma), z_3(\sigma)] A[z_1(\sigma)] A[z_2(\sigma)] A[z_3(\sigma)] \ , \tag{5.3.3}$$

where $V_3[z_1, z_2, z_3] = {}_1\langle z|_2\langle z|_3\langle z|V_3\rangle_{123}$ represents an interaction potential that has absorbed all possible derivative terms.*

It is convenient to expand the BRS charge Q in the zero modes b_0, c_0, and the time derivative a_0^0. We obtain

$$Q = c_0 a_0^0 a_0^0 + 2K a_0^0 + c_0 \tilde{\Delta} + b_0 T_+ + \tilde{Q} \ . \tag{5.3.4}$$

The operators K, $\tilde{\Delta}$, T_+, and $\tilde{Q}$ (anti-)commute with each other by virtue of the nilpotency of Q in 26 dimensions. Also, $\{K, K\} = \frac{1}{2}T_+$, and $\tilde{Q}^2 + \tilde{\Delta}T_+ = 0$. Another important property of these operators is their "hermiticity," i.e., $\int A \star KB = (-)^{|A|} \int KA \star B$, and similarly for $\tilde{\Delta}$ and $\tilde{Q}$. They follow from the properties of Q as a derivation, eq.(5.1.10) and $\int QA = 0$.

A straightforward computation gives

$$P_- \equiv b_0 \frac{\delta \mathcal{L}}{\delta(a_0^0 A_-)} = a_0^0 A_- + 2b_0 K A_+ \ , \tag{5.3.5a}$$

---

* Notice that, with our choice of time, $V_3[z_1, z_2, z_3]$ is non-local in both space and time. For locality, we have to choose the coordinates of the mid-point of the string as the space-time coordinates [12]. However, for explicit calculations, it is necessary that we express the vertex (eq.(5.1.22)) in terms of a more convenient basis. Henceforth, we shall ignore the problem of non-locality, hoping to resolve it in the future.

$$P_+ \equiv c_0 \frac{\delta \mathcal{L}}{\delta(a_0^0 A_+)} = 0 \ . \tag{5.3.5b}$$

The factors b_0 and c_0 have been introduced for convenience. In deriving these two equations we have discarded surface terms. We have also made use of the 'hermiticity' of the operators K, $\tilde{\Delta}$ and $\tilde{Q}$. Thus, the phase space is endowed with a symplectic structure described by the two-form

$$\omega \equiv \int \delta(\Sigma_0)\delta A \star \delta P = \int \delta(\Sigma_0)\delta A_- \star (c_0 a_0^0 \delta A_- + 2K\delta A_+) \ , \tag{5.3.6}$$

where Σ_0 is the hypersurface $t = 0$. $\delta(\Sigma_0)$ is a δ-function which vanishes unless the center of mass of the string lies on the surface Σ_0. Notice that this is different from the one adopted by Witten [12]. The difference lies in the choice of time. ω is obtained by choosing the center-of-mass time-coordinate as the time of the system, whereas Witten's choice corresponds to time being the time coordinate of the mid-point of the string, i.e., $x^0(\frac{\pi}{2})$.

Since P_+ vanishes, it follows that A_+ is similar to A_0^a in Yang-Mills theories. It plays the role of a Lagrange multiplier, as can be seen by constructing the hamiltonian density,

$$\mathcal{H} \equiv P_- \star c_0 a_0^0 A_- - \mathcal{L} = \mathcal{H}_0 + A_+ \star \Gamma_- \ , \tag{5.3.7}$$

where

$$\mathcal{H}_0 = \frac{1}{2} P_- \star c_0 P_- + \frac{1}{2} A_- \star c_0 \tilde{\Delta} A_- - \frac{1}{3} A_- \star A_- \star A_- \ , \tag{5.3.8}$$

$$\Gamma_- = 2KP_- + \tilde{Q} A_- - b_0 c_0 (A_- \star A_-) \ . \tag{5.3.9}$$

The operator $b_0 c_0$ acts as a projection onto the space generated from the $|-\rangle$ vacuum. To derive the above equation, we used the fact that $X \equiv A_+ \star A_+ = 0$. To prove this, we have to make use of the explicit form of the three-string vertex (eq.(5.1.14)).

Since $|X\rangle_1 = {}_2\langle A_+|_3\langle A_+|V_3\rangle_{123}$ and $\langle +|+\rangle = \langle -|-\rangle = 0$, separating the zero modes of the ghosts in the vertex, we obtain

$$|X\rangle_1 = {}_2\langle A_+|_3\langle A_+|e^{N_b}e^{N_f}(1 + b_0^1 C^1)\frac{1}{2}\{C^2, C^3\}|+\rangle_1|-\rangle_2|-\rangle_3 \ , \qquad (5.3.10)$$

where $N_b = \sum a_{-m}^{\mu,r} N_{mn}^{rs} a_{-n}^{\mu,s}$, $N_f = \sum b_{-m}^r \tilde{N}_{mn}^{rs} n c_{-n}^s$ and $C^r = \sum \tilde{N}_{0n}^{rs} n c_{-n}^s$ $(m > 0)$. Since the c_n's anti-commute, we have $\{C^2, C^3\} = 0$ and therefore $|X\rangle_1 = 0$, which shows that only terms linear in A_+ can appear in the hamiltonian. In this formulation, there is no sense in setting $A_+ = 0$, because we lose the constraint $\Gamma_- = 0$. Thus we see that the gauge $A_+ = 0$ (known as the Siegel gauge) is similar to the $A_0 = 0$ gauge in Yang-Mills theories. However, unlike in gauge theories, Γ_- generates a transformation

$$\delta A_- = 2K\epsilon_- \qquad (5.3.11a)$$

$$\delta P_- = -\tilde{Q}\epsilon_- + b_0[A_-, \epsilon_-] \ , \qquad (5.3.11b)$$

where $[A, B] = A \star B - (-)^{AB} B \star A$, which does not leave the hamiltonian invariant. (The gauge parameter ϵ_- is of course independent of time ($a_0^0 \epsilon_- = 0$), and $\delta A_+ = 0$. Also, $(-)^{|\epsilon|} = -1$.) We therefore have to impose an additional constraint, $\tilde{\Gamma}'_- \equiv \{\Gamma_-, \mathcal{H}_0\} = 0$, where $\{A, B\}$ is the Poisson bracket of A and B. We easily find

$$\tilde{\Gamma}'_- = 2K\tilde{\Delta}A_- - 2b_0 K(A_- \star A_-) + \tilde{Q}P_- - b_0 c_0 [P_-, A_-] \ , \qquad (5.3.12)$$

generating the transformation

$$\delta A_- = \tilde{Q}\epsilon'_- + b_0 c_0 [A_-, \epsilon'_-] \ , \qquad (5.3.13a)$$

$$\delta P_- = K\tilde{\Delta}\epsilon'_- + b_0[P_-, \epsilon'_-] - 2b_0 K[A_-, \epsilon'_-] \ . \qquad (5.3.13b)$$

No further constraints need to be imposed, because

$$\{\tilde{\Gamma}'_-, \mathcal{H}_0\} = \tilde{\Delta}\Gamma_- + B(\tilde{Q}\Gamma_- + 2K\tilde{\Gamma}'_-) \ , \qquad (5.3.14)$$

where $B = \sum \beta_{-n} b_n$ and the coefficients β_{-n} $(n > 0)$ solve the system of equations: $\sum_n \tilde{N}_{0n}^{11} n \beta_{-n} = 1$ and $\sum_n \tilde{N}_{mn}^{11} n \beta_{-n} = 0$ $(m > 0)$. These constraints are first class constraints and they form a closed algebra, because

$$\{\Gamma_-, \tilde{\Gamma}'_-\} = \Gamma_- \ . \tag{5.3.15}$$

To quantize the theory, we couple the system to an external current J_+, by adding a term $J_+ \star A_-$ to the lagrangian. Thus, the second constraint (eq.(5.3.12)) becomes

$$\Gamma'_- \equiv \tilde{\Gamma}'_- + 2b_0 K J_+ = 0 \ . \tag{5.3.16}$$

We define the generating functional by

$$e^{-W[J_+]} = \int \mathcal{D}A_- \mathcal{D}P_- \mathcal{D}\lambda_+ \mathcal{D}\lambda'_+ e^{-\int P_- \star c_0 a_0^0 A_- - \mathcal{H}_0 + \lambda_+ \star \Gamma_- + \lambda'_+ \star \Gamma'_- + J_+ \star A_-} \ . \tag{5.3.17}$$

To compute Green functions, we have to fix the gauge. Exploiting the invariance under the transformation generated by Γ_-, we impose the condition $KA_- = 0$. Notice that this implies $T_+ A_- = 4K^2 A_- = 0$, and $GA_- = 0$, where G is the ghost number operator. This gauge ($T_+ A_- = GA_- = 0$) was introduced by Siegel and Zwiebach [13], in the case of a free theory. The inclusion of interactions does not change things, because the transformation properties of A_- (eq.(5.3.11a)) are not affected by the interactions. To eliminate the second gauge invariance (eq.(5.3.13)), we impose the constraint $\tilde{Q} A_- = 0$. These two constraints are implemented by the Faddeev-Popov procedure. Inserting the two factors

$$1 = \Delta_1[A_-] \int \mathcal{D}\epsilon_- \delta[K(A_- + 2K\epsilon_-)] \ , \tag{5.3.18a}$$

$$1 = \Delta_2[A_-] \int D\epsilon'_- \delta[\tilde{Q}(A_- + (\tilde{Q} + b_0 c_0 [A_-, \ \cdot \ ])\epsilon'_-)] \ , \tag{5.3.18b}$$

into the path-integral (eq.(5.3.17)), and performing two gauge transformations, we obtain

$$e^{-W[J_+]} = \int \mathcal{D}A_- \mathcal{D}P_- \mathcal{D}\lambda_+ \mathcal{D}\lambda'_+ \Delta_1[A_-]\Delta_2[A_-]\delta[KA_-]\delta[\tilde{Q}A_-]$$
$$\times \exp\left\{ - \int P_- \star c_0 a_0^0 A_- - H_0 + \lambda_+ \star \Gamma_- + \lambda'_+ \star \Gamma'_- + J_+ \star A_- \right\} \ . \tag{5.3.19}$$

The Faddeev-Popov determinants are $\Delta_1 = \det' K^2$ and $\Delta_2 = \det' \tilde{Q}(\tilde{Q} + b_0 c_0 [A_-, \cdot])$, where the prime denotes omission of the zero modes of the operators K and $\tilde{Q}$.

Integration over the lagrange multipliers, λ_+ and λ'_+ in eq.(5.3.19) produces two δ-functionals that enforce the constraints $\Gamma_- = 0$ and $\Gamma'_- = 0$, respectively. We can use these two constraints to integrate over the redundant degrees of freedom. Splitting the momentum P_- as $P_- = \Pi_- + \Pi'_- + \Pi''_-$, where $\tilde{Q}\Pi_- = K\Pi_- = K\Pi'_- = 0$, the constraints (eqs.(5.3.9) and (5.3.16)) become

$$\Gamma_- \equiv 2K\Pi''_- + b_0 c_0 (A_- \star A_-) = 0 \ , \tag{5.3.20a}$$
$$\Gamma'_- \equiv (\tilde{Q} + b_0 c_0 [A_-, \ \cdot \ ])(\Pi_- + \Pi''_-)$$
$$+ b_0 c_0 [\Pi'_-, A_-] - 2K(b_0 (A_- \star A_-) + J_+) = 0 \ . \tag{5.3.20b}$$

Using eqs.(5.3.20a) and (5.3.20b), we can express Π'_- and Π''_- in terms of Π_-. Explicitly,

$$\Pi'_- = \frac{1}{2} K^{-1} b_0 c_0 (A_- \star A_-)$$
$$+ (\tilde{Q} + b_0 c_0 [A_-, \ \cdot \ ])^{-1} b_0 ([\Pi_-, A_-] + 2K((A_- \star A_-) + J_+)) \ , \tag{5.3.21a}$$
$$\Pi''_- = -\frac{1}{2} K^{-1} b_0 c_0 (A_- \star A_-) \ . \tag{5.3.21b}$$

Therefore, integration over Π'_- and Π''_- gives rise to two factors, $(\det' K)^{-1}$ and $(\det'(\tilde{Q} + b_0 c_0 [A_-, \ \cdot \ ]))^{-1}$. These two factors, together with the two Faddeev-Popov determinants, give a factor of $\det'(K\tilde{Q})$ which is a constant, and can therefore be absorbed into the overall normalization of the generating functional. Hence, the final form of the generating functional (eq.(5.3.19)) is

$$e^{-W[J_+]} = \int \mathcal{D}A_-\mathcal{D}\Pi_- e^{-\int \Pi_-\star c_0 a_0^0 A_- - \tilde{\mathcal{H}}_0} \ , \tag{5.3.22}$$

where A_- and Π_- are annihilated by both K and $\tilde{Q}$. The hamiltonian $\tilde{\mathcal{H}}_0$ is a function of A_-, its conjugate momentum Π_- and the external current J_+:

$$\begin{aligned} \tilde{\mathcal{H}}_0 = &\frac{1}{2}\Pi_- \star c_0\Pi_- + \frac{1}{2}\Pi'_- \star c_0\Pi'_- + \frac{1}{2}\Pi''_- \star c_0\Pi''_- \\ &+\frac{1}{2}A_- \star c_0\tilde{\Delta}A_- - \frac{1}{3}A_- \star A_- \star A_- + J_+ \star A_- \end{aligned} \ , \tag{5.3.23}$$

where Π'_- and Π''_- are given by eqs.(5.3.21a) and (5.3.21b).

This form of the hamiltonian agrees with the minimal form of ref.[13], if interactions are ignored. To demonstrate this, consider the generating functional for the free theory,

$$e^{-W_0[J_+]} = \int \mathcal{D}A_-\mathcal{D}\Pi_- e^{-\int \Pi_-\star c_0 a_0^0 A_- - \frac{1}{2}\Pi_-\star c_0\Pi_- + \frac{1}{2}J_+\star b_0\tilde{\Delta}^{-1}J_+ - \frac{1}{2}A_-\star c_0\tilde{\Delta}A_- + J_+\star A_-} \ . \tag{5.3.24}$$

To derive this, we used $\Pi'_- = 2b_0\tilde{Q}^{-1}KJ_+$, $\Pi''_- = 0$, which follow from eqs.(5.3.21a) and (5.3.21b), respectively, when the interactions are switched off. Integrating over Π_- by completing the square in the exponent, we obtain

$$e^{-W_0[J_+]} = \int \mathcal{D}A_- e^{-\int \frac{1}{2}A_-\star c_0\Delta A_- + J_+\star A_- + \frac{1}{2}J_+\star b_0\tilde{\Delta}^{-1}J_+} \ , \tag{5.3.25}$$

where $\Delta \equiv a_0^0 a_0^0 + \tilde{\Delta} = \sum : a_{-n}^\mu a_n^\mu : -1 + \sum n : c_{-n}b_n :$. Thus, $\frac{1}{2}\int A_- \star c_0\Delta A_-$ is the free gauge-fixed action, in agreement with the results of ref.[13]. The free propagator is therefore Δ^{-1}. The first few eigenstates of Δ are $|p,-\rangle$ (with eigenvalue p^2-1), $a_{-1}^\mu|p,-\rangle$ (with eigenvalue p^2), $a_{-1}^\mu a_{-1}^\nu|p,-\rangle$, $a_{-2}|p,-\rangle$, $c_{-1}b_{-1}|p,-\rangle$ (with eigenvalue p^2+1), etc. Expanding a general state $|A_-\rangle$ in terms of these states, we obtain

$$\begin{aligned} |A_-\rangle = \int d^{26}p \Big(&\tilde{\phi}(p) + \tilde{A}_\mu(p)a_{-1}^\mu \\ &+\tilde{B}_{\mu\nu}(p)a_{-1}^\mu a_{-1}^\nu + \tilde{C}_\mu(p)a_{-2}^\mu + \tilde{\psi}(p)c_{-1}b_{-1} + \cdots\Big)|p,-\rangle \ . \end{aligned} \tag{5.3.26}$$

The constraint $KA_- = 0$ implies $A_0 = B_{0\mu} = C_0 = \psi = 0$. It is also easily seen that the second constraint, $\tilde{Q}A_- = 0$ implies $\partial_i A_i = \partial_i B_{ij} = \partial_i C_i = 0$ $(i,j = 1,...,25)$. Notice that for the massless gauge field A_μ, the second constraint is the Coulomb gauge.

We shall now compute the contribution of the lowest modes of $|A_-\rangle$ to the four-tachyon scattering amplitude. Our discussion will follow closely the discussion in ref.[10]. Consider on-shell tachyons, $|p_1,-\rangle$, $|p_2,-\rangle$, $|p_3,-\rangle$ and $|p_4,-\rangle$, where $p_1^2 = p_2^2 = p_3^2 = p_4^2 = 1$. Define the Mandelstam variables $s = -(p_1+p_2)^2$, $t = -(p_2+p_3)^2$ and $u = -(p_1+p_3)^2$. We shall concentrate on the $s-t$ dual diagram. The amplitude is

$$S = S_{-1} + S_0 + \cdots \ , \tag{5.3.27}$$

where $S_{-1}, S_0, \ldots$ are the contributions of an intermediate virtual tachyon, massless field, etc., respectively. The propagator for the tachyon is just

$$D_{-1}(p,p') \equiv \langle p, -|\Delta^{-1}|p', -\rangle = \frac{1}{p^2-1}\delta(p+p') \ . \tag{5.3.28}$$

It is a little harder to find the propagator $D_0^{\mu\nu}(p,p')$ for a massless field. Writing

$$D_0^{\mu\nu}(p,p') = \langle p, -|a_1^\mu \frac{\delta^2}{\delta J_+^2} W[J_+]\Big|_{J_+=0} a_{-1}^\mu |p', -\rangle \ , \tag{5.3.29}$$

after a little algebra we obtain

$$\begin{aligned} D_0^{00} &= \frac{1}{p_i p_i}\delta(p+p') \ , \ D_0^{0i} = 0 \ , \\ D_0^{ij} &= \frac{1}{p^2}\left(\delta_{ij} - \frac{p_i p_j}{p_i p_i}\right)\delta(p+p') \ . \end{aligned} \tag{5.3.30}$$

This is just the propagator of a photon in the Coulomb gauge. The other ingredients that are needed are the tachyon-tachyon-tachyon and tachyon-tachyon-photon interaction vertices. A straightforward calculation gives

$$V_{-1}(p_1,p_2,p_3) = {}_1\langle p_1, -|{}_2\langle p_2, -|{}_3\langle p_3, -|V_3\rangle_{123}$$

$$= \left(\frac{4}{3\sqrt{3}}\right)^{p_1^2+p_2^2+p_3^2} \delta(p_1+p_2+p_3) \ , \tag{5.3.31}$$

$$V_0^\mu(p_1,p_2,p_3) = {}_1\langle p_1,-|\,{}_2\langle p_2,-|\,{}_3\langle p_3,-|a_1^{\mu,3}V_3\rangle_{123}$$

$$= \left(\frac{4}{3\sqrt{3}}\right)^{p_1^2+p_2^2+p_3^2+1} \frac{1}{\sqrt{2}}(p_2-p_1)^\mu \delta(p_1+p_2+p_3) \ , \tag{5.3.32}$$

where we used $N_{00}^{rs} = \frac{1}{2}\ln(16/27)\delta^{rs}$ and $N_{10}^{31} = -N_{10}^{32} = (4/27)^{1/2}$, $N_{10}^{33} = 0$. Thus, using eqs.(5.3.28) and (5.3.31), we obtain

$$\begin{aligned} S_{-1} &= V_{-1}(p_1,p_2,p)D_{-1}(p)V_{-1}(p_3,p_4,p) \\ &\propto \left(\frac{16}{27}\right)^{-s-1} \frac{1}{s+1} \ . \end{aligned} \tag{5.3.33}$$

Similarly, eqs.(5.3.30) and (5.3.32) give

$$\begin{aligned} S_0 &= V_0^\mu(p_1,p_2,p)D_0^{\mu\nu}(p)V_0^\nu(p_3,p_4,p) \\ &\propto \left(\frac{16}{27}\right)^{-s} \frac{4+2t+s}{2s} \ . \end{aligned} \tag{5.3.34}$$

Therefore, the ratio of the residues of the poles at $s=0$ and $s=-1$ is $2+t$. It is equal to the ratio of the residues of the corresponding poles in the function $B(-s-1,-t-1)$ [10]. This provides an indication that our results agree with the results of the dual theory at tree level.

It would be interesting to extend these calculations by computing the poles of a loop diagram. This will allow a better understanding of how closed strings arise in an open-string field theory, and will resolve questions of unitarity.

References

1. L. Brink, P. Di Vecchia and P. Howe, *Phys. Lett.* **65B** (1976) 471;

   S. Deser and B. Zumino, *Phys. Lett.* **65B** (1976) 369;

   A. M. Polyakov, *Phys. Lett.* **103B** (1981) 207.

2. E. D'Hoker and S. B. Giddings, Princeton preprint (1987), PUPT-1046.

3. M. Kaku and K. Kikkawa, *Phys. Rev.* **D10** (1974) 1110, 1823;

   E. Cremmer and J.-L. Gervais, *Nucl. Phys.* **B76** (1974) 209; **B90** (1975) 420.

4. E. Witten, *Nucl. Phys.* **B268** (1986) 253.

5. M. Kato and K. Ogawa, *Nucl. Phys.* **B212** (1983) 443.

6. D. J. Gross and A. Jevicki, *Nucl. Phys.* **B283** (1987) 1.

7. G. T. Horowitz and A. Strominger, Princeton preprint (1987), Print-87-0006.

8. E. S. Fradkin and G. A. Vilkovisky, *Phys. Lett.* **55B** (1975) 224;

   I. A. Batalin and G. A. Vilkovisky, *Phys. Lett.* **102B** (1981) 27.

9. M. Bochicchio, Rome Univ. preprint (1986), ROME-527-1986;

   M. Bochicchio, Princeton preprint (1986), PUPT-1028.

10. C. B. Thorn, Princeton preprint (1986), Print-86-1334.

11. G. Siopsis, Caltech preprint (1987), CALT-68-1438.

12. E. Witten, *Nucl. Phys.* **B276** (1986) 291;

    C. Crnković and E. Witten, Princeton preprint (1986), Print-86-1309;

    C. Crnković, Princeton preprint (1986), PUPT-1033.

13. W. Siegel and B. Zwiebach, *Nucl. Phys.* **B263** (1986) 105.