UNIVERSITY OF SOUTHAMPTON

SECOND QUANTISED STRING THEORY

by

David Laurence Gee

A Thesis Submitted for the Degree of

Doctor of Philosophy

Department of Physics

August 1989

CONTENTS

UNIVERSITY OF SOUTHAMPTON

ABSTRACT

FACULTY OF SCIENCE

PHYSICS

Doctor of Philosophy

SECOND QUANTISED STRING THEORY

by David Laurence Gee

An attempt is made to construct a non-perturbative, second quantised framework for string theory by describing worldsheets implicitly as the solution surfaces of $D-2$ functions on D-dimensional space-time. The formalism in fact generalises to extended objects of arbitrary dimension. The worldsheets are not parameterised and hence modular invariance and duality should be incorporated at a fundamental level, thus avoiding the overcounting of vacuum graphs encountered with string field theory. Couplings to background fields are introduced which reduce to the usual Polyakov couplings on choosing a parameterisation of the worldsheet. The formalism has a natural $GL(D-2,\mathbb{R})$ gauge invariance which is awkward to fix owing to the existence of a delta function in the action. BRST ghosts can be introduced however it is not clear how to remove all the gauge invariance while the delta function remains in the action. Replacing the delta function with a more general function of the $D-2$ implicit functions yields gauge fixed equations of motion however it is no longer clear how to introduce ghosts. A hamiltonian quantisation is presented however ambiguities arise for cases of greater complexity than the particle in two dimensions. The formalism appears to furnish a topological field theory for vanishing metric and may prove useful in investigating the conjectured unbroken phase of general relativity.

PREFACE

This work was carried out in collaboration with Dr T.R. Morris and has been accepted for publication in Nuclear Physics.

ACKNOWLEDGEMENTS

I have been fortunate in having Tim Morris as a supervisor during the course of this work and it has been a pleasure to work with him. His inspiring enthusiasm and endless encouragement, even in the face of alarming ignorance, were invaluable. I should thank my friends and colleagues in Southampton who have suffered me for varying lengths of time and who have conspired to make my stretch here an enjoyable and stimulating one. I am especially grateful for the companionship and encouragement of those old friends and relatives on whom I have freely imposed throughout what has at times been a trying three years.

I thank my parents for early guidance however I particularly wish to acknowledge the selfless love and support I received from my mother, the importance of which I was too blind to recognise during her lifetime.

I am greatly indebted to Peter Stewart not only for the tremendous support and guidance he has given me for many years but also for being a source of continual encouragement and indispensible wisdom. This thesis is dedicated to Peter.

I thank the Science and Engineering Research Council for financial aid.

The most merciful thing in the world, I think, is the inability of the human mind to correlate all its contents. We live on a placid island of ignorance in the midst of black seas of infinity, and it was not meant that we should voyage far. The sciences, each straining in its own direction, have hitherto harmed us little; but some day the piecing together of dissociated knowledge will open up such terrifying vistas of reality, and of our frightful position therein, that we shall either go mad from the revelation or flee from the deadly light into the peace and safety of a new dark age.

H.P. Lovecraft

INTRODUCTION

The origin of string theory (the study of one dimensional extended objects) in the dual resonance model of hadrons and the discoveries which led to its candidature as a theory of everything are well documented.[1,2] Briefly, it was the appearance of tachyons, the identification of the graviton in closed string spectra and the fact that unphysical degrees of freedom only decoupled in 26 or 10 space-time dimensions (for bosonic and fermionic strings respectively) that suggested this reinterpretation of string theory. It has been vigorously studied as a theory of everything ever since the discovery by Green and Schwarz of consistent supersymmetric theories, with phenomenologically sound gauge groups, which were free from anomalies and divergences (up to at least one loop order).[3]

Despite all the attractive features of string theory, however, there are many problems that it does not and cannot address largely due to the fact that it is defined perturbatively. This leads one to believe that it is just an approximation to some deeper and more fundamental theory. It may be that string theory is not the correct route along which to proceed in order to find the ultimate theory of nature—it might in the end just provide some useful insights into the quantisation of gravity. Such a conjecture is not productive; more knowledge can be gained by building on established foundations. In accordance with this philosophy, this work describes an attempt to construct a second quantised string theory which can, in principle, explore some of the avenues closed to the first quantised theory.

In chapter 1 some of the major phenomenological and theoretical triumphs of first quantised string theory are briefly mentioned. More attention is paid to the failures of string theory and it becomes clear that a non-perturbative, second quantised framework is required. Various attempts at such a framework are described along with their drawbacks however particular attention is paid to the main contender, the naïve generalisation of string theory; string field theory. Accordingly particular attention is paid in showing that field theory is not the answer either. The chapter concludes with a summary of some of the properties that a fundamental theory of everything ought to encompass.

Chapter 2 reviews and discusses a particular proposal for second quantised string theory which describes worldsheets implicitly as the solutions to certain functions. Consistency with the standard sigma model is extended to include all massless background field couplings and those at the first mass level. A prescription is given for computing higher mass couplings in the formalism. Partly as a check on the formalism and partly because it is not obvious at first sight, the constant dilaton coupling is explicitly shown to be a topological invariant.

Gauge fixing is investigated in chapter 3 using BRST methods and a rather unconventional one made possible by discoveries in chapter 2. The measure is discussed in detail and it is shown how the theory generates the usual first quantised partition function. A hamiltonian quantisation of the model is briefly attempted.

Some speculative ideas are presented in chapter 4 along with some conclusions.

References

[1] J. Scherk, *Rev. Mod. Phys.* **47** (1975) 123;
J.H. Schwarz, *Phys. Rep.* **89** (1982) 223.

[2] M.B. Green, J.H. Schwarz and E. Witten, "Superstring Theory," (C.U.P., 1988).

[3] M.B. Green and J.H. Schwarz, *Phys. Lett.* **149B** (1984) 117, *Nucl. Phys.* **B198** (1982) 441 and *Phys. Lett.* **151B** (1985) 21.

CHAPTER 1

FIRST QUANTISED STRING THEORY

String theory has generated considerable interest in recent years because it appeared to provide for the first time a theory unifying all the forces of nature. However evidence is mounting which suggests that, in its perturbative formulation, string theory cannot fulfill this promise. This chapter reviews the rise and fall of perturbative string theory and discusses some of its successors.

In section 1 some of the successes of string theory are noted briefly, the intention being to give some idea of why it is worth pursuing string theory into a non-perturbative and second quantised regime. The major reason for the interest in string theory is that it seems to provide for the first time a theory of quantum gravity. In preparation for some work in section 4 on the one loop string vacuum graph, a (non-covariant) quantisation of the free bosonic string is presented.

The shortcomings of string theory as a theory of everything are detailed in section 2. Different viewpoints are considered and they point unanimously to a second quantised theory.

Some of the alternative proposals to perturbative string theory, all of which are still connected with strings (or at least two dimensional surfaces), are mentioned in section 3. These satisfy to a greater or lesser extent some of the objections raised in section 2. The most popular approach has been to follow the particle analogy and construct a string field theory. Some of the more widely studied versions are considered, particularly that due to Witten[1,2] which, in its cubic form,[3] comes close to satisfying some of the philosophical criticisms raised in section 2.

Section 4 is devoted to yet more criticisms, this time of string field theory. Some of the reasons why field theory may not be the answer are presented and one of the more persuasive, the infinite overcounting of the torus contribution, is dwelt on.

A summary of the criteria a would-be theory of everything should satisfy is given in section 5 and the way is paved for the introduction of a candidate second quantised string theory in chapter 2.

1.1 Successes of String Theory

Probably the most important property which recommends string theory as a theory of unification, and justifies the thorough investigation it has enjoyed, is that it incorporates gravity at a fundamental level. In fact, of the two types of string theory, open and closed, open strings contain closed ones (in for example the one loop non-planar diagram (Fig. 1.1.1)) and the closed string spectrum contains the graviton (strictly speaking it contains a massless spin 2 symmetric tensor which has a coupling consistent with general covariance). Hence string theory not only includes gravity but is inconsistent without it. Indeed it contains gravity in a finite and anomaly free way thus yielding a viable theory of quantum gravity. This is remarkable considering that string theory is just a conservative extension of the relativistic quantum mechanics of particles in which it is very hard, if not impossible, to consistently incorporate gravity. String theory also correctly contains the gauge structure of Yang-Mills, necessary to describe low energy physics, and therefore makes it inseperable from gravity. String theory also predicts the dimension of space-time.

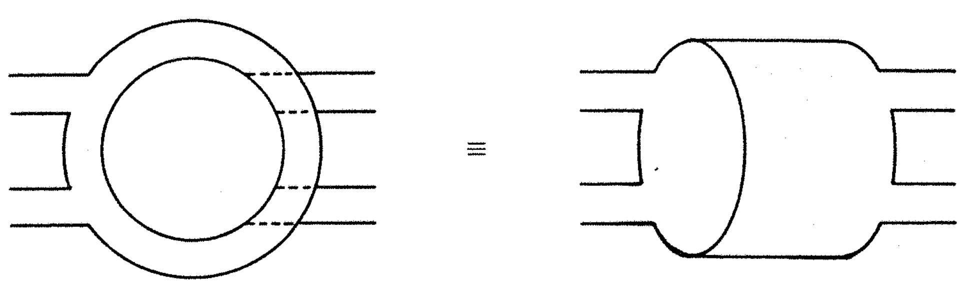


Fig. 1.1.1

While this all looks very promising, the important question of whether it in fact describes the real world must be addressed. Not long after Green and Schwarz sparked off the current interest in string theory the heterotic string[4] was formulated. This is a hybrid consisting of the right moving modes of the closed $D = 10$ superstring and the left moving modes of the $D = 26$ closed bosonic string. Sixteen of the twenty-six bosonic dimensions are compactified on a sixteen dimensional torus producing a ten dimensional theory with gauge group $E_8 \otimes E_8$. Realistic compactifications were soon found[5] yielding the familiar four dimensional space-time with six of the original ten dimensions belonging to a compact manifold; i.e., $M^{10} \to M^4 \times K^6$. The curvature of this compact manifold broke the gauge group to one in which the Standard Model and quarks and gluons could be identified.[6] Further it was found that the topology of the internal space determined things like the number of generations and Yukawa couplings.[7] For the heterotic string many solutions were found by compactifying on Calabi-Yau spaces[8] or on orbifolds.[9] Attempts at constructing realistic models based on the type II string (purely closed) were doomed after it was shown that no classical vacua of the type II superstring could reproduce the Standard Model. Therefore, at least at the classical (tree) level, the heterotic string is of the most interest. All that then remained was to find a special solution out of the many possibilities and introduce supersymmetry breaking.

It is quite remarkable that, merely by considering the fundamental objects of nature to be extended and string like rather than point like and by proceeding in the same manner as for particles, such a rich structure can be obtained. In fact, as shall be seen later, some of this structure suggests that investigations beyond the standard conceptual framework of physics may be fruitful.

The first quantisation of the free bosonic string is presented here partly for completeness but mainly because it will be required in the calculation of the one loop vacuum graph later. Although the BRST approach in conformal field theory is probably the most elegant method of quantising the string and does not need zeta function regularisation, the staid old hamiltonian method will be employed here. This choice is made since it easily lends itself to the light-cone gauge in which all the symmetry is completely fixed and no Fadeev-Popov ghosts are required.

The quadratic action for the free string propagating in D-dimensional space-time is

$$S = -\frac{T}{2}\int d^2\sigma\sqrt{g}g^{ab}\partial_a X^\mu \partial_b X^\nu G_{\mu\nu}(X) \tag{1.1.1}$$

where g_{ab} is an auxiliary field which plays the rôle of a two dimensional metric. $T = 1/2\pi\alpha'$ is the string tension. X^μ are the space-time coordinates of the string and are functions of the worldsheet parameter $\sigma^a = (\tau, \sigma)$, $a = 0, 1$. Flat space-time is considered where $G_{\mu\nu} = \eta_{\mu\nu}$ with positive signature. Varying (1.1.1) with respect to the auxiliary metric g_{ab} yields the field equation

$$\partial_a X \cdot \partial_b X - \tfrac{1}{2} g_{ab} g^{cd} \partial_c X \cdot \partial_d X = 0. \tag{1.1.2}$$

This will be imposed as the constraint $T_{ab} = 0$ where the energy momentum tensor is defined by

$$T_{ab} = -\frac{2}{T\sqrt{g}}\frac{\delta\mathcal{L}}{\delta g^{ab}}.$$

Defining

$$G_{ab} \equiv \partial_a X \cdot \partial_b X = \tfrac{1}{2} g_{ab} g^{cd} \partial_c X \cdot \partial_d X = \lambda g_{ab}$$

where $\lambda = \frac{1}{2}\mathrm{tr}(g^{-1}G)$, the determinant of G_{ab} is

$$|\det G_{ab}| = G = \lambda^2 |\det g_{ab}| = \lambda^2 g.$$

Therefore the string lagrangian density can be rewritten

$$\tfrac{1}{2}\sqrt{g} g^{ab} G_{ab} = \sqrt{G}$$

and the action (1.1.1) becomes

$$S = -T \int d^2\sigma \sqrt{\dot{X}^2 X'^2 - (\dot{X} \cdot X')^2}$$

where dot and prime refer to differentiation with respect to τ and σ respectively. This is the action originally proposed by Nambu.[10]

Action (1.1.1) has the following local invariances:

$$\text{Reparameterisations:} \quad \delta X^\mu = \xi^a \partial_a X^\mu \tag{1.1.3}$$

$$\delta g_{ab} = \nabla_a \xi_b + \nabla_b \xi_a \tag{1.1.4}$$

$$\text{Weyl scaling:} \quad \delta g^{ab} = \Lambda(\sigma) g^{ab}. \tag{1.1.5}$$

These invariances can be used to fix the three components of the worldsheet metric leaving a residual symmetry generated by a combined reparameterisation and Weyl scaling for which

$$\nabla_a \xi_b + \nabla_b \xi_a = \Lambda g_{ab}. \tag{1.1.6}$$

This is conformal invariance. The constraints (1.1.2) can be incorporated by introducing ghosts or by explicitly solving them in for example the non-covariant light-cone formulation. It is the latter approach that will be used here.

Choosing the gauge where $g_{ab} = \eta_{ab} = \mathrm{diag}(-1, 1)$, the equation of motion for $X^\mu(\sigma)$ is found to be the two dimensional wave equation

$$\Box X^\mu \equiv (\partial_\sigma^2 - \partial_\tau^2) X^\mu = 0. \tag{1.1.7}$$

This implies that $X^\mu(\sigma)$ can be expanded as the sum of left and right moving modes

$$X^\mu(\sigma) = X_L^\mu(\sigma^-) + X_R^\mu(\sigma^+)$$

where $\sigma^{\pm} = \tau \pm \sigma$. In these coordinates $\partial_{\pm} = \frac{1}{2}(\partial_\tau \pm \partial_\sigma)$ and $\eta_{+-} = \eta_{-+} = -\frac{1}{2}$ with the other components zero. Before a general solution to (1.1.7) (which becomes $\partial_+ \partial_- X^\mu = 0$ in the above coordinates) is written down, a few comments are needed about boundary conditions. Invariance of (1.1.1) under small changes of X demands that $X'|_{\sigma=0,\pi} = 0$ for open strings. Such boundary terms vanish automatically for closed strings due to their periodicity condition $X^\mu(\tau,\sigma) = X^\mu(\tau,\sigma+\pi)$. The closed string solutions to (1.1.7) (compatible with these periodicity requirements) are

$$
\begin{aligned}
X^\mu_R &= \tfrac{1}{2}x^\mu + \tfrac{1}{2}l^2 p^\mu \sigma^- + \tfrac{i}{2} l \sum_{n\neq 0} \frac{\alpha^\mu_n}{n} e^{-2in\sigma^-} \\
X^\mu_L &= \tfrac{1}{2}x^\mu + \tfrac{1}{2}l^2 p^\mu \sigma^+ + \tfrac{i}{2} l \sum_{n\neq 0} \frac{\tilde{\alpha}^\mu_n}{n} e^{-2in\sigma^+}
\end{aligned}
\tag{1.1.8}
$$

and for open strings the solution is

$$X^\mu = x^\mu + l^2 p^\mu \tau + il \sum_{n\neq 0} \frac{\alpha^\mu_n}{n} e^{-in\tau} \cos(n\sigma) \tag{1.1.9}$$

which is compatible with the open string boundary conditions. A length parameter, $l = 1/\sqrt{\pi T}$, has been included on dimensional grounds. Hermiticity of X implies $\alpha^\dagger_n = \alpha_{-n}$.

Introducing the conjugate momentum $P_\mu(\tau,\sigma) = \delta\mathcal{L}/\delta\dot{X}^\mu$ enables Poisson brackets to be calculated and hence canonical commutation relations to be inferred:

$$[P^\mu(\tau,\sigma), X^\nu(\tau,\sigma')] = -i\delta(\sigma-\sigma')\eta^{\mu\nu} \tag{1.1.10}$$

with others zero. This relation implies the following non-zero commutation relations:

$$[p^\mu, x^\nu] = -i\eta^{\mu\nu} \tag{1.1.11}$$

$$[\alpha^\mu_n, \alpha^\nu_m] = [\tilde{\alpha}^\mu_n, \tilde{\alpha}^\nu_m] = n\delta_{n+m}\eta^{\mu\nu}. \tag{1.1.12}$$

The constraint equation (1.1.2), $T_{ab} = 0$, can be imposed as a condition on physical states in the quantum theory, analogous to the Gupta-Bleuler mechanism in QED, i.e., $T^+_{ab}|\text{phys}\rangle = 0$ where the plus sign refers to the positive frequency modes. Tracelessness of T_{ab} implies $T_{+-} = T_{-+} = 0$ and so the constraint (1.1.2) becomes

$$T_{\pm\pm} = \partial_\pm X \cdot \partial_\pm X = 0. \tag{1.1.13}$$

It is easiest, however, to solve these constraints at the classical level and then quantise. This can be achieved by using the residual diffeomorphism invariance (1.1.6). In worldsheet light-cone coordinates the action is

$$S = -2T \int d^2\sigma \partial_+ X \cdot \partial_- X \tag{1.1.14}$$

where the measure $d^2\sigma = \frac{1}{2}d\sigma^+ d\sigma^-$. The action is invariant under $\sigma^\pm \to \sigma'^\pm = f_{L,R}(\sigma^\pm)$ for a general function f as a consequence of Weyl invariance (since the above transformation is conformal). Space-time light-cone coordinates are now introduced (so that this remaining symmetry can be removed) by defining the choice:

$$X^\mu \to (X^\pm, X^i) \quad i = 1, \ldots, D-2 \tag{1.1.15}$$

where for a general vector $V^\pm = \frac{1}{\sqrt{2}}(V^0 \pm V^{D-1})$. The residual conformal invariance can be used to choose the functions $f_{L,R}(\sigma^\pm)$ to be

$$f_{L,R}(\sigma^\pm) = \frac{2X^+_{L,R}(\sigma^\pm) - x^+}{l^2 p^+}. \tag{1.1.16}$$

The $\pm$ indices on the σ refer to the worldsheet and correspond to left and right moving modes whilst the index on X, p and x is a space-time light-cone index. Equation (1.1.16) corresponds to the choice

$$\begin{aligned}\tau \to \frac{1}{2}(\sigma'^+ + \sigma'^-) &= \frac{1}{l^2 p^+}(X^+_L + X^+_R - x^+)\\ &= \frac{1}{l^2 p^+}(X^+ - x^+)\end{aligned}$$

which implies

$$X^+(\sigma) = x^+ + l^2 p^+ \tau. \tag{1.1.17}$$

This is conceptually satisfactory since it means that every point on the string has the same value of 'time'.

In these coordinates the constraints (1.1.13) become

$$\begin{aligned}T_{++} &= \partial_+ X \cdot \partial_+ X\\ &= \partial_+ X^i \partial_+ X^i - 2\partial_+ X^+ \partial_+ X^-\\ &= \partial_+ X^i \partial_+ X^i - l^2 p^+ \partial_+ X^- = 0\\ T_{--} &= \partial_- X^i \partial_- X^i - l^2 p^+ \partial_- X^- = 0\end{aligned}$$

so that

$$\partial_\pm X^- = \frac{1}{l^2 p^+} \partial_\pm X^i \partial_\pm X^i. \tag{1.1.18}$$

When this is solved both X^+ and X^- will have been determined in terms of X^i and hence eliminated leaving only the transverse oscillations. Integrating (1.1.18) with respect to σ from 0 to π gives

$$\frac{\pi l^4}{2} p^+ p^- = \int_0^\pi d\sigma (\partial_+ X^i)^2 = \int_0^\pi d\sigma (\partial_- X^i)^2$$

which is satisfied automatically for open strings but yields the constraint

$$N - \tilde{N} = 0 \tag{1.1.19}$$

for closed strings. $N = \frac{1}{2} \sum_{n \neq 0} \alpha^i_{-n} \alpha^i_n$ (and similarly for $\tilde{N}$) is the number operator measuring the excitation of the string.

The closed and open string hamiltonians are

$$\begin{aligned} H_c &= 2\pi T^2 \int_0^\pi d\sigma (T_{++} + T_{--}) \\ &= -2p^+ p^- + (p^i)^2 + 4\pi T(N + \tilde{N}) \\ &= p^\mu p_\mu + 4\pi T(N + \tilde{N}) \end{aligned} \tag{1.1.20}$$

$$H_o = p^\mu p_\mu + 2\pi T N \tag{1.1.21}$$

respectively. In order to solve the constraint equation, $H|\Psi\rangle = 0$, and find the spectrum of states, the hamiltonian must be normal ordered to isolate the zero point energy. ($|\Psi\rangle$ will be discussed below). Inserting the commutation relations for the α's (1.1.12) between vacua gives (in light-cone coordinates)

$$\sum_{i=1}^{D-2} \langle 0|\alpha^i_n \alpha^i_{-n} - \alpha^i_{-n} \alpha^i_n |0\rangle = n = \sum_{i=1}^{D-2} \left(|\alpha^i_{-n}|0\rangle|^2 - |\alpha^i_n|0\rangle|^2 \right)$$

and similarly for $\tilde{\alpha}^i$. Assuming that α^i_n and α^i_{-n} annihilate different states then for $n > 0$ this implies $\alpha^i_n|0\rangle = 0$. Hence α^i_n can be regarded as annihilation operators for $n > 0$ and creation operators for $n < 0$. A Fock space vacuum will be used such that $|0\rangle = \prod_{n>0} |0\rangle_n$ where $\alpha^i_n|0\rangle_n = 0$. The number operators can now be normal ordered:

$$\begin{aligned} N &= \tfrac{1}{2} \sum_{n \neq 0} \alpha^i_{-n} \alpha^i_n \\ &= \tfrac{1}{2} \sum_{n>0} \alpha^i_{-n} \alpha^i_n + \tfrac{1}{2} \sum_{n<0} \alpha^i_n \alpha^i_{-n} \\ &= \sum_{n>0} \alpha^i_{-n} \alpha^i_n + \frac{D-2}{2} \sum_{n>0} n. \end{aligned} \tag{1.1.22}$$

The anomalous term in (1.1.22) can be regularised using the Riemann zeta function. $\zeta(s) = \sum_{n>0} n^{-s}$ is valid for $\mathrm{Re}(s) > 1$ however it has an analytic continuation to $s = -1$ where $\zeta(-1) = -1/12$. The need for this regularisation only arises when dealing explicitly with oscillators and is unnecessary when using, for example, conformal field theory. Inserting this value into the expressions for the hamiltonians (1.1.20, 1.1.21) gives

$$H_c = p^2 + 4\pi T(\sum_{n>0} \alpha^i_{-n}\alpha^i_n + \sum_{n>0} \tilde{\alpha}^i_{-n}\tilde{\alpha}^i_n - \frac{D-2}{12}) \equiv -\Box + M_c^2$$

$$H_o = p^2 + 2\pi T(\sum_{n>0} \alpha^i_{-n}\alpha^i_n - \frac{D-2}{24}) \equiv -\Box + M_o^2$$

for closed and open strings respectively. The normalisation implied in (1.1.12) is rather unusual so operators a^i_n and $a^{i\dagger}_n$ are introduced which satisfy the usual harmonic oscillator like commutation relations. Defining

$$\alpha^i_n = \sqrt{n}a^i_n, \qquad n > 0$$

$$\alpha^i_{-n} = \sqrt{n}a^{i\dagger}_n, \qquad n > 0$$

gives the more familiar commutation relations

$$[a^i_n, a^{j\dagger}_m] = \delta_{n-m}\delta^{ij}$$

with others zero. a^i_n and $a^{i\dagger}_n$ are annihilation and creation operators respectively.

States are now constructed from the vacuum by acting on it with $a^{i\dagger}_n$, for example the first excited state of the open string is $a^{i\dagger}_1|0\rangle$. A momentum label is to be understood in the vacuum, i.e., $|0\rangle = |0;p\rangle$. This first excited state is a $D-2$ component vector of $SO(D-2)$ and should be massless (so that it transforms correctly under a Lorentz boost). Acting on this state with the open string mass operator yields

$$\frac{1}{4\pi T}M_o^2 a^{i\dagger}_1|0\rangle = 2\sum_{n>0} n a^{j\dagger}_n[a^j_n, a^{i\dagger}_1]|0\rangle - \frac{D-2}{12}a^{i\dagger}_1|0\rangle$$

$$= (2 - \frac{(D-2)}{12})a^{i\dagger}_1|0\rangle$$

which vanishes for $D = 26$. Therefore the light-cone treatment gives a Lorentz invariant quantum theory if $D = 26$. Now that this famous result has been obtained, open strings will no longer be considered since they are not required in the calculation of the torus vacuum graph.

The string state $|\Psi\rangle$ mentioned earlier can be expanded (for closed strings) as a sum of component fields acting on the Fock space vacuum:

$$|\Psi\rangle = \Psi_0|0\rangle + \Psi_{1,ij}a^{i\dagger}_1\tilde{a}^{j\dagger}_1|0\rangle + \Psi_{2,ij}a^{i\dagger}_2\tilde{a}^{j\dagger}_2|0\rangle + \Psi_{2,ijk}a^{i\dagger}_1 a^{j\dagger}_1\tilde{a}^{k\dagger}_2 + \cdots \qquad (1.1.23)$$

where the component fields depend on the string coordinate. $|\Psi\rangle$ can thus be thought of as a functional of the string coordinate, $\Psi[X(\sigma)]$. Terms in just $a^{i\dagger}$ or $\tilde{a}^{i\dagger}$ and mixed terms such as $a^{i\dagger}_1\tilde{a}^{j\dagger}_2$ are eliminated from (1.1.23) by the (quantum) constraint

$N - \tilde{N}|\Psi\rangle = 0$. Under $SO(D-2)$ $\Psi_{1,ij}$ transforms as a traceless symmetric tensor $\oplus$ antisymmetric tensor $\oplus$ scalar. These are massless fields since (in $D = 26$)

$$M_c^2 a_1^{i\dagger} \tilde{a}_1^{j\dagger}|0\rangle = 4\pi T((1+1)-2) a_1^{i\dagger} \tilde{a}_1^{j\dagger}|0\rangle = 0.$$

The first of these is the graviton, the antisymmetric tensor field is $B_{\mu\nu}$ and the scalar is the dilaton field.

Although string theory appears to be an infinite tower of excitations, it will become clear later that it cannot just be represented by an infinite tower of fields with an infinite gauge invariance.

1.2 Failure of Perturbative String Theory

The inadequacies of first quantised string theory, as defined perturbatively, as a theory of everything fall into three main categories; phenomenological, theoretical and philosophical. Although presented separately, the criticisms mentioned are not all entirely unconnected. Some of the arguments below, particularly those in the final subsection are rather vague and qualitative but are underpinned by the more concrete ones; the less well defined objections are included as they tend to be of a more physical nature. Although most of the following complaints are directed at string theory some are in fact general and thus must be addressed by *any* fundamental theory be it string oriented or not.

i. Phenomenological Considerations

Working under the assumption that string theory is weakly coupled (i.e., it is in the perturbative regime) and that there is a realistic vacuum, Dine and Seiberg[11] have shown that the nonlinear sigma model[12] corresponding to the string theory is strongly coupled. Further considerations[13] concerning the dilaton field (which determines the string coupling) indicated that the full string theory must also in fact be strongly coupled. The arguments revolved around the dilaton effective potential and required that the theory did not describe 10-dimensional flat space and that there was no massless dilaton. The existence of a massless dilaton would give rise to an infinite range force and such a force (mediated by a scalar) does not appear to have been observed. It would also imply that all string theories had degenerate vacua which were therefore indistinguishable. This leaves two possibilities; either the effective potential is negative, in which case the theory is strongly coupled anyway, or there is a minimum at finite dilaton expectation value in the positive potential. However such a minimum would arise in a region in which the lowest (non-trivial) order calculation does not give the shape of the potential correctly. Therefore higher orders are at least as important implying that such a vacuum is strongly coupled.

ii. Theoretical Considerations

Further arguments for the invalidity of string perturbation theory came from Gross and Periwal[14] who demonstrated that, by considering the behaviour of the string integrand with genus compared to that of the volume of moduli space, not only did string perturbation theory diverge but that it was also not even Borel summable. This was triumphed as very good news since if the perturbation series was summable then effects that were true order by order would also be true of the full theory. This would imply the existence of a massless dilaton and that supersymmetry could not be broken. Thus even a sum to all orders of string perturbation theory is ruled out and so a non-perturbative framework must be sought. A very interesting development arose out of considerations of the close connection between conformally invariant 2-dimensional non-linear sigma models describing string propagation on some background and classical string physics. The condition for conformal invariance (vanishing of the renormalisation group beta functions) can be expressed as the equations of motion of a space-time effective action for the background fields.[15] Such actions and beta functions have been calculated, in the bosonic case,[16] up to 3 and 4 loop order. The background equations of motion will be altered by quantum string loop effects and it is not unreasonable to expect that they will have an expansion in $e^{-2\Phi_0}$ (for closed strings) and that they will be derivable from a space-time effective action. Φ_0 is the constant background dilaton field. If these are the conformal invariance conditions of some generalised 2-dimensional field theory then this may give insight into the deeper symmetries of string theory. The renormalisation group beta functions have been generalised to include the torus (e^0) contribution[17] and it was found that divergences obstructing, for example BRST invariance, cancelled between inequivalent worldsheets (e.g., a torus divergence is cancelled by a counterterm on the sphere weighted by $e^{-2\Phi_0}$). This implies a new family of string theories consistent only on all worldsheets at once which suggests the need for a second quantised string theory.

iii. Philosophical Considerations

The beauty of general relativity is that Einstein built it on a *priori* physical principles and it would be very satisfying to build the ultimate theory describing nature in such a way. Attempts have been made to do this, for example Kaku[18] proposed invariance under DiffS^1/S^1 (diffeomorphisms of strings into themselves) as a fundamental principle. This however presupposes that the fundamental objects are strings. Perhaps the only requirement is that nature be predictable which might lead to ideas such as the space-time invariance of physical laws. It is much more likely that, should the ultimate theory be found, it will require a *posterori* justifications and understanding. For example the symmetry underlying invariances such as BRST and general covariance promises to be very deep and perhaps inconceivable outside 10 or 26 dimensions. Perhaps a much deeper theory is required in any case to explain the origin of the huge gauge invariance of string theory (and the possibly even larger one of string field theory). More reasonable demands of any theory claiming to be fundamental are that it should not presuppose any particular background geometry or indeed topology.[19] For example, it is unsatisfactory for a theory of closed strings to depend on a background generated by a condensate of gravitons. The geometry and topology of space-time (and also perhaps even its dimension) should emerge as solutions to the equations of motion; they should become dynamical variables. String theory has taken a first step towards this goal by predicting the dimension of space-time (albeit at first sight incorrectly).

iv. Consequences for Couplings, Confinement and SSB

The success of QCD suggests that at some level any theory which purports to be fundamental must describe its confining nature. Being non-Borel summable, like string theory, an accurate theoretical description of QCD effects will probably only be obtainable from a non-perturbative theory. This rules out perturbative string theory and indeed on physical grounds it is hard to imagine a perturbative theory being able to accurately reproduce all the structure of strongly coupled effects. On a similar note, it has been conjectured[20] that there might exist a confining phase of general relativity in which general covariance is unbroken. Such a phase is only likely to be accessible from a non-perturbative framework. Incidentally, the investigation of exact solutions to Einstein's equations in the classical limit of general relativity (for example the Schwarzschild solution) would probably require the use of non-perturbative tools. Spontaneous symmetry breaking, in for example the Glashow-Salam-Weinberg model of electro-weak interactions, occurs due to an incorrect choice of vacuum. The 'correct' vacuum cannot be reached by perturbing about an incorrect one; its discovery lies in the non-perturbative sector. A serious theory must also address the breaking of supersymmetry and perhaps the breaking of general covariance, as noted above. Another issue concerning vacua and peculiar to strings is that, from a phenomenological viewpoint, there are many different string theories (corresponding to quantisation about different classical solutions) not all of which can be fundamental. The problem is not to find more solutions but rather to find the dynamics which chooses a particular vacuum (and hence solution). If these are indeed just perturbations of the same underlying theory[21,22] then the need for a non-perturbative theory of strings (assuming that the string route is correct) is obvious.

The arguments presented in this section should leave one in no doubt that some sort of non-perturbative theory is required if the above problems are to be resolved. The theory should also be second quantised to get away from the single string description which is probably just a convenient way of expressing the deeper structure sought.

A first quantised approach is intuitively inappropriate for describing condensates (of gravitons) and hence for describing gravity. Also, to describe an infinite number of worldsheets (as is necessary if the theory is to be consistent on all of them simultaneously) a single string picture is hopelessly inadequate.

1.3 Alternatives to Perturbative String Theory

i. Universal Moduli Space

Much of the beauty of string theory is due to conformal geometry. String processes are described in terms of functional integrals over Riemann surfaces of increasing topological complexity, only a few dual diagrams contributing at each order of perturbation theory. After gauge fixing Weyl rescalings and diffeomorphisms, one only wants to integrate over metrics inequivalent under these symmetries; i.e., moduli space. Letting $\mathcal{G}_\Sigma = \{\text{space of all metrics on a Riemann surface } \Sigma\}$, one can write moduli space as

$$\mathcal{M}_\Sigma = \frac{\mathcal{G}_\Sigma}{\text{Weyl} \times \text{diffeomorphisms}}.$$

Non-perturbative effects might be be incorporated by considering universal moduli space[23] which not only includes the Riemann surfaces that correspond to the complete perturbation expansion but also infinite genus Riemann surfaces. It is these latter surfaces which contain the non-perturbative information. Such an extension was considered since in the above formalism one is only required to integrate over the moduli space of the correct genus to describe a certain process.[24] The simplicity in having to consider only a few diagrams for a particular genus surface is lost in for example string field theory where the surface is seemingly arbitrarily sliced up into vertices and propagators. It then requires of the order of $n!$ Feynman diagrams to describe an n-loop process. Closed strings are more difficult to describe however, since a natural cell decomposition of Riemann surfaces without boundaries or marked points is less obvious. At present universal moduli space seems to provide a rather abstract statement of what non-perturbative string theory should be.

ii. Kähler Geometry of Loop Space

Yet another interesting development was the interpretation of the (doubled) string Fourier components as the coordinates of an infinite dimensional Kähler space.[25] The space of all maps of the circle into space-time can be interpreted as the configuration space of the closed string. The subset of this space corresponding to strings which begin and end at the origin is a Kähler manifold. The Kähler potential, K, describes the geometry of this space and plays the rôle of a closed string field. Loops should be invariant under differentiable mappings into themselves (i.e., invariant under the group DiffS^1). Pure rotations (corresponding to $\sigma \mapsto \sigma + \text{constant}$) are trivial and so one only need consider the quotient space

$$\mathcal{S} = \mathrm{Diff}S^1/S^1$$

which is generated by the Virasoro operators. The importance of $\mathcal{S}$ (which is also Kähler) in string theory has been noted[18] for example it has an invariant cohomology, the BRST cohomology. Bowick and Rajeev proposed as a fundamental principle the requirement that the Kähler space given as the sum of bundles over $\mathcal{S}$ be Ricci flat. This is reminiscent of the Einstein field equations, however the space in question is now infinite-dimensional and complex. For flat space-times $D = 26, 10$ is regained for bosonic strings and superstrings respectively. Perhaps the requirement of Ricci flatness is equivalent to the conformal invariance of an associated two-dimensional sigma model (and hence plays the rôle of classical field theory equations of motion). This formalism does not however incorporate interactions and is hence intrinsically limited.

iii. Renormalisation Group

The sigma model approach to string theory is based on the observation that the equations of motion require conformal invariance of the background 2-dimensional quantum field theory. Since the configuration space of string theory is maybe, in some sense, the space of all 2-dimensional field theories, it seems natural to formulate string theory on this space.[26] This "theory space" should necessarily contain the generalised 2-D field theory alluded to in subsection 1.2.*ii*, the conformal invariance conditions of which should correspond to the loop corrected equations of motion of a space-time effective action. Such effects have been investigated by generalising the renormalisation group beta functions to

$$\beta = \beta_c + \hbar\beta_1 + \hbar^2\beta_2 + \cdots$$

where β_c is the classical (tree level) beta function. Cutoffs must be incorporated since the space will contain non-renormalisable field theories. Theory space is believed to be an infinite dimensional smooth 'manifold' and different cutoffs are thought to correspond to different parameterisations of the space. The reason for this latter point is that most cutoffs make reference to a particular point in theory space and hence a choice of cutoff corresponds to a choice of coordinate patch. Different cutoff theories are related by the renormalisation group and in order to study its flow in theory space, the symmetries and structure of the space must be understood. Knowledge of these are also required to find transition functions between different coordinate patches and hence do more than just generate perturbation theory in the vicinity of classical solutions. Such a non-perturbative framework is necessary in order to search for candidate vacua which is related to the global aspects of theory space. A good enough understanding of theory space is still lacking.

iv. String Field Theory

Initial attempts at formulating string field theory were in connection with dual resonance models and were hence essentially aimed at hadronic physics.[27] The light-cone method, in which strings join at their ends, was used since it overcame some of the problems associated with string field theory such as the overcounting caused by summing diagrams equivalent under a duality transformation. This method was introduced by Mandelstam[28] who built on the light-cone quantisation of the Nambu string.[29] Another problem (of *covariant* second quantisation) was how to define a canonical momentum, $\Pi[X(\sigma)]$, conjugate to the string field $\Phi[X(\sigma)]$. In the light-cone formalism $\Pi[X(\sigma)]$ can be unambiguously defined whereas in the covariant approach the conjugate fields are not in one to one correspondence.

After string theory came to be viewed as (a stepping stone towards) a theory of everything, however, a covariant formulation of string field theory became essential in order that it could describe effects such as spontaneous symmetry breaking. Such a formulation was introduced by Siegel[30] and was based on the previous light cone work. It was unsatisfactory for a number of reasons (for example loop integrals were more infra-red divergent) and so, in line with more recent methods, he derived a free gauge-fixed second quantised open string action[31] using the BRST transformations obtained for the first quantised string by Kato and Ogawa.[32] In the closed string case[33] he found that the extra fields needed to fully describe the classical covariant theory were contained in the BRST-quantised string field (for example the trace of the graviton was in the normal statistics part of the ghost sector).

The gauge invariant free string action has been determined to be generally of the form[34,35]

$$S = (\Phi|Q|\Phi) \tag{1.3.1}$$

for some appropriate Q. The inner product assumes integration and multiplication laws. This style of action has the gauge invariance

$$\delta\Phi = Q\Lambda \tag{1.3.2}$$

for nilpotent Q. The content of (1.3.1) was hinted at when Banks and Peskin[35] discovered that their open string action (with Q the BRST charge) contained a linearised

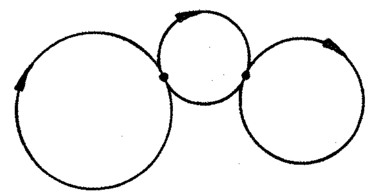


Fig. 1.3.1

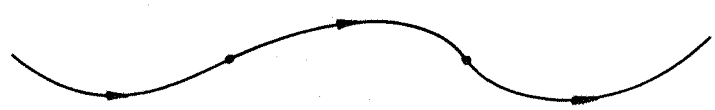


Fig. 1.3.2

version of Yang-Mills and that their closed string action contained a linearised form of the classical effective action[15] (giving the classical conformal invariance conditions on the background fields). The corresponding gauge invariances were found in (1.3.2). This is encouraging indeed and suggests that string field theory may lead in the right direction. This is probably to be expected since string field theory is effectively just a rewrite of string theory anyway. In depth treatments of light-cone based interacting open and closed string field theories[36-43] revealed much of interest. For example it was found that symmetries such as BRST and gauge invariance were indeed connected[39] and that the ortho-symplectic group may be significant[37,38] in unifying space-time and gauge invariances. Despite these discoveries, problems remained with the light-cone derived theories such as their complexity and in particular the necessity of having a string length parameter, α.[41] Although it has been shown[37,41] that on-shell physical amplitudes are independent of the string lengths (up to a conservation factor $\delta(\sum_r \alpha_r)$) and that they are thought to be purely gauge artifacts,[39,42] there are still $\delta(0)$ type divergences in the closed string sector[43] which require regularisation. These are in fact avoided in the $OSP(26,2/2)$ formalism[38] where the string lengths are cancelled by fermionic degrees of freedom.

A more geometric formulation for open string field theory, without the need for a string length parameter, was proposed by Witten.[1] As can be seen from fig. 1.3.1, closed string interactions are commutative (since there is only one way to join two closed strings) but can be non-associative. Thus for closed strings $A \cdot (B \cdot C) = A \cdot (C \cdot B) \neq (A \cdot C) \cdot B = 0$. On the other hand, fig. 1.3.2 illustrates that open string interactions are non-commutative but associative (unless the strings are parameterised). This is because the two ways of joining two open strings are not necessarily equivalent ($U \cdot V \neq V \cdot U$). Based on this observation, Witten, generalising Yang-Mills, proposed the following gauge invariance on the (open) string field $\Phi[X(\sigma), \text{ghosts}]$:

$$\delta\Phi = Q\Lambda + \Phi * \Lambda - \Lambda * \Phi. \tag{1.3.3}$$

Λ is the gauge parameter and the $*$ operation can be viewed as a generalisation of matrix multiplication and wedge product. The only (non-topologically invariant) action consistent with (1.3.3) and the axioms of non-commutative geometry:

$$Q(\Phi * \Psi) = Q\Phi * \Psi + (-)^{|\Phi|}\Phi * Q\Psi \tag{1.3.4}$$

$$\int \Phi * \Psi = (-)^{|\Phi||\Psi|} \int \Psi * \Phi \tag{1.3.5}$$

$$\int Q\Phi = 0 \quad \forall\Phi \tag{1.3.6}$$

is the Chern-Simons three-form action:

$$S = \tfrac{1}{2} \int \Phi * Q\Phi + \tfrac{2}{3}\Phi * \Phi * \Phi. \tag{1.3.7}$$

Here $\int$ is a map from string fields to the complex numbers (a sort of combined matrix trace and integral). $|\Phi| = 0, 1$ for Grassman even and odd Φ respectively which is just like introducing a $\mathbb{Z}_2$ grading. Incidentally the above axioms (1.3.4 to 6) were also used in ref. [44]. A closed string field theory would require a non-associative geometry.

The radical difference between this and the old light-cone based approaches to string field theory is in the form of the interactions. Reparameterisation invariance is discarded in favour of BRST invariance (which has already been seen to be relevant)[31] and the mid-point of the string, $\sigma = \frac{\pi}{2}$, is singled out. The definition of Q in this theory has deliberately been left until now; it is the usual BRST charge. The new type of interaction considers the string as two halves, left and right. The star operation merges the left half of one string with the right half of another to produce a third string again with left and right pieces (fig. 1.3.3). It is clear that the length of the string is irrelevant, thus avoiding the need for an extra length parameter. The kinetic part of action (1.3.7) is of the same form as (1.3.1) and is thus consistent with previous work, however there are intriguing differences between Witten's action and light-cone ones. Since physical states have ghost number $-\frac{1}{2}$ and the BRST operator Q has ghost number 1, (1.3.7) implies that the star operation and the integral have ghost numbers $\frac{3}{2}$ and $-\frac{3}{2}$ respectively. Hence there can be no quartic interaction (as this would have non-zero ghost number) unlike in the light-cone theory.[27] The ghost number for the physical states comes from the physical state condition $Q|\text{phys}\rangle = 0$ which, when expressed in oscillators,[45] implies $c_n|\text{phys}\rangle = 0$ for $n > 0$. In general $c_{n>-p}|p-\frac{1}{2}\rangle = 0$ for a ghost number $p - \frac{1}{2}$ vacuum and this implies that the physical vacuum has ghost number $-\frac{1}{2}$.

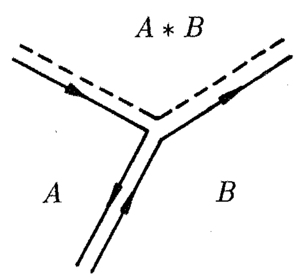


Fig. 1.3.3

Despite the absence of manifest duality and the existence of a preferred point, the string mid-point, it has been explicitly demonstrated that Witten's string field theory reproduces the correct dual amplitudes and is reparameterisation invariant and BRST gauge-invariant.[46] At the quantum level it has been shown that closed string poles appear correctly,[47–52] however this then raises the question of the existence of external closed strings as required by unitarity.

Most of the several attempts at constructing closed string field theories[42–44,51–57] fall foul of the problem mentioned in the previous section of dependence on the background. Simply generalising the usual Chern-Simons type action to closed strings by reinterpreting the open string fields as closed ones leads to problems with ghost number. One finds that unless everything has twice the ghost number as for the open string case, the integral and multiplication law depend on the surface for their definition.[53] This can be overcome by modifying the kinetic operator to have ghost number 2,[43,44,53,54] however since the new kinetic operator still contains the BRST charge, the problem of background dependence persists.

There have been attempts to remove the explicit background dependence contained in the BRST charge[3,58] by transforming the string fields in such a way that only an interaction part remains in the action. In general this leaves a cubic action of the form:

$$S = \tfrac{1}{3} \int \Phi * \Phi * \Phi \tag{1.3.8}$$

with equation of motion

$$\Phi * \Phi = 0. \tag{1.3.9}$$

Another method[40-42] used a mechanism involving $OSP(26, 2/2)$ to remove the closed string vacuum graphs (which generate the full effective action for the background fields). Closed strings could presumably then be incorporated in a way which does not permit a background field interpretation. The light-cone style formulation of ref. [58] suffers from that old bane the string length parameter, α. This time ambiguities arise in taking the limit $\alpha \to 0$ in expressions such as $\Phi * \Phi$.

The cubic version of Witten's string field theory[3] has been more widely studied and has in fact been shown to be free of any background dependence.[59] Its large symmetry structure has also been investigated.[48,60,61] It was suggested that closed string effects might be included in the open string action without the need for explicit closed string fields or extra open string interactions.[2,51,52] This was verified and it was also shown that general covariance could be identified in the open string gauge invariance and that closed string backgrounds could be created.[50-52] Purely closed interactions described in such a theory are of $O(\lambda^4)$ instead of $O(\lambda^2)$ indicating that a separate closed string theory is required. Here λ is the cubic open string coupling constant and is related to the cubic closed string coupling, g, by $\lambda^2 = g$. The non-associative behaviour[51] of closed string interactions was exploited to construct a purely closed string action from open string fields.[55] The open string sector was eliminated by only considering the parts with associativity anomalies.

Despite all this progress, a really satisfactory closed string field theory is still lacking. String field theory is still under investigation and, although progress has been made towards a theory independent of the geometry and topology of space-time,[19,59] there is increasing evidence that such a conservative extension of string theory is insufficient. This is the subject of the next section.

1.4 Failure of String Field Theory

Attempts to gain insight into the underlying symmetries of string theory have suggested that field theory is not in fact a suitable framework in which to describe it. Investigations of the high energy scattering of strings,[62] motivated by the assumption that certain effects, for example a broken symmetry, should be easier to recognise at high energy, revealed that all non-vanishing fixed angle scattering amplitudes were equal up to a constant. The emergence of this strange symmetry, relating every particle to every other, suggests that some sort of phase transition might occur in entering a high energy regime. Moreover it was found that the behaviour of these amplitudes fell off too fast to be described by an effective local field theory.[63] Considerations of the thermal ensemble,[64] again in an attempt to investigate deeper symmetry structures, have revealed that the number of degrees of freedom of string theory is much less than that for any field theory. It also appeared that above a certain temperature, analogous to the deconfining temperature in QCD, the notion of strings propagating on smooth surfaces must be forsaken. Worldsheets should be replaced by some sort of discrete surface. In fact it has been shown that continuum strings can be described by discrete field theory.[65]

String theory cannot naïvely reduce to a sum of field theories because processes inequivalent in field theory are equivalent in string theory due to duality. Thus field theory suffers from an immediate overcounting of scattering amplitudes. Further evidence to support the idea that string theory is more than just a sum of field theories comes from calculations of the first contribution to the second quantised closed string, the torus vacuum graph. This contribution will be calculated in field theory using results derived earlier in the light-cone gauge and will then be compared to the answer obtained directly from the Polyakov path integral.[66]

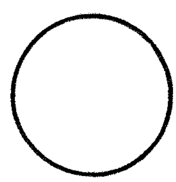
As noted earlier, string theory is equivalent to an infinite tower of fields Ψ_A, where the index A not only stands for labels distinguishing the different fields but also the $SO(D-2)$ indices which are obtained in the light-cone gauge. In this gauge the action for these fields can be written

$$S = -\tfrac{1}{2}\sum_A \int d^D x\, \Psi_A(-\Box + M_A^2)\Psi_A \tag{1.4.1}$$

which yields the correct equation of motion, the first quantised Schrödinger equation. The calculations will be done in euclidean space since considerable simplification is afforded by this choice. Being a purely free action, (1.4.1) does not suffer from the problems associated with introducing interactions into closed string field theory. Treating (1.4.1) as a quantum field theory, the generating functional can be written

$$Z = \prod_A \int \mathcal{D}\Psi_A e^{-S} = e^{-\Gamma} \tag{1.4.2}$$

where Γ is the effective action (generator of one particle irreducible diagrams). Usually (1.4.2) would be written $Z = e^{-W}$ where W is the generator of connected graphs however it is related to Γ by a term involving sources of which there are none in this case. Since there are no interactions either, the only contribution to Γ is the Feynman diagram:



(1.4.3)

which corresponds to a sum of particle loops each representing a string going round a loop with a particular excitation. This is expected to be equivalent to a closed string going round a loop with all fluctuations being considered at once. This is what the Polyakov path integral measures; the integral of all maps of the torus into space-time.

The above diagram has meaning in terms of background fields. When the metric is expanded around flat space, $G_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}$, the $\Box \doteq \nabla_\mu \nabla^\mu$ term in (1.4.1) generates graviton interaction terms which cause coupling to the background. Taking these into account, Γ can be shown to give the one loop contribution to the cosmological constant.[17] (1.4.3) can be regarded as the unexpanded diagram immersed in the full background metric.

As in normal field theory, (1.4.2) can be rewritten

$$e^{-\Gamma} = \prod_A \mathrm{Det}^{-\frac{1}{2}}(-\Box + M_A^2)$$

and hence

$$\begin{aligned} \Gamma &= \tfrac{1}{2} \sum_A \ln \mathrm{Det}(-\Box + M_A^2) \\ &= \tfrac{1}{2} \sum_A \mathrm{Tr} \ln(-\Box + M_A^2) \\ &= \tfrac{1}{2} \int d^D x \sum_A \left[\ln(-\Box + M_A^2)\right]_{xx}. \end{aligned} \tag{1.4.4}$$

Here Det (Tr) stands for the determinant (trace) of an infinite dimensional operator. The Fourier transform of (1.4.4) is

$$\begin{aligned}\Gamma &= \tfrac{1}{2}\int \frac{d^D p}{(2\pi)^D}\sum_A \ln(p^2+M_A^2)\\ &= -\tfrac{1}{2}\int \frac{d^D p}{(2\pi)^D}\int_0^\infty \frac{d\tau_2}{\tau_2}\sum_A e^{-\tau_2(p^2+M_A^2)}\end{aligned}$$

where the logarithm has been expressed as an integral over the (conventionally named) parameter τ_2. Also a factor $\int d^D x$, the space-time volume, has been factored out. Since the exponent in the above expression is negative definite (due to working in euclidean space) the momentum integral factorises into D Gaussian integrals which are easily performed giving

$$\begin{aligned}\Gamma &= -\tfrac{1}{2}\int_0^\infty \frac{d\tau_2}{\tau_2}(4\pi\tau_2)^{-\frac{D}{2}}\sum_A e^{-\tau_2 M_A^2}\\ &= -\tfrac{1}{2}\int_0^\infty \frac{d\tau_2}{\tau_2}(4\pi\tau_2)^{-\frac{D}{2}}\sum_A \langle A|e^{-\tau_2 M^2}|A\rangle. \end{aligned}\tag{1.4.5}$$

The exponent has been rewritten as the mass operator M^2 acting on the excited Fock space state $|A\rangle$ related to the field Ψ_A. These states have the normalisation $\langle A|B\rangle = \delta_{AB}$ and are subject to the constraint $N - \tilde{N}|A\rangle = 0$. It is convenient to incorporate this constraint explicitly and to sum over all possible Fock space states. Inserting the relation

$$\delta_{N-\tilde{N}} = \int_{-\frac{1}{2}}^{\frac{1}{2}} d\tau_1 e^{2\pi i\tau_1(N-\tilde{N})} \tag{1.4.6}$$

into (1.4.5) with M^2 replaced by $4\pi T(N+\tilde{N}-(D-2)/12)$ and τ_2 rescaled to $\tau_2/2T$ produces the following expression

$$\Gamma = -\frac{T^{\frac{D}{2}}}{2}\int_0^\infty \frac{d\tau_2}{\tau_2}\int_{-\frac{1}{2}}^{\frac{1}{2}} d\tau_1(2\pi\tau_2)^{-\frac{D}{2}}\sum_A \langle A|e^{2\pi i\tau(N-\beta)}e^{2\pi i\bar{\tau}(\tilde{N}-\beta)}|A\rangle \tag{1.4.7}$$

where the complex number $\tau = \tau_1 + i\tau_2$ and $\beta = (D-2)/24$ have been introduced. It is evident from (1.4.7) that the left and right moving sectors factorise and can thus be treated separately. Also since

$$e^{2\pi i\tau N} = \prod_{i=1}^{D-2}\prod_{n>0} e^{2\pi i\tau n a_n^{i\dagger}a_n^i}$$

each mode (i.e., each n, i) can be considered separately too. The sum $\sum_A$ then corresponds to a sum over all possible contributions of the form $|r\rangle = (a_n^{i\dagger})^r|0\rangle$, restriction

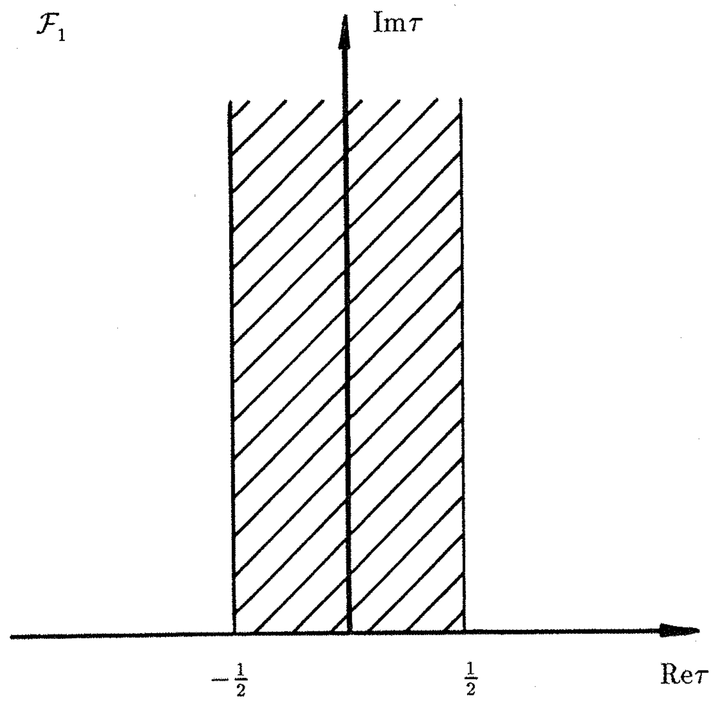


Fig. 1.4.1

(1.4.6) ensuring the correct number of left and right movers. Hence

$$\begin{aligned}\sum_A \langle A|e^{2\pi i\tau N}|A\rangle &= \prod_{i=1}^{D-2}\prod_{n>0}\sum_{r>0}\langle r|e^{2\pi i\tau N}|r\rangle \\ &= \prod_{i=1}^{D-2}\prod_{n>0}\sum_{r>0} e^{2\pi i\tau nr} \\ &= \prod_{n>0}(1-e^{2\pi i n\tau})^{-(D-2)}.\end{aligned}$$

Inserting this into (1.4.7) along with the left moving contribution yields

$$\Gamma = -\frac{T^{\frac{D}{2}}}{2}\int_0^\infty \frac{d\tau_2}{\tau_2}\int_{-\frac{1}{2}}^{\frac{1}{2}} d\tau_1 (2\pi\tau_2)^{-\frac{D}{2}}|\prod_{n>0}(1-e^{2\pi i n\tau})|^{-2(D-2)}e^{\pi\tau_2\frac{(D-2)}{6}}. \tag{1.4.8}$$

This can be expressed more concisely in terms of the Jacobi theta function

$$\theta_1'(0|\tau) = 2\pi e^{\frac{i\pi\tau}{4}}(\prod_{n>0}(1-e^{2\pi i n\tau}))^3.$$

Inserting this into the above equation for Γ gives

$$\Gamma = -\frac{T^{\frac{D}{2}}}{2}\int_{\mathcal{F}_1} d^2\tau (2\pi)^{-\frac{D-8}{6}}\tau_2^{-\frac{D+2}{2}}|\theta_1'(0|\tau)|^{-\frac{2(D-2)}{3}} \tag{1.4.9}$$

where the region of integration, $\mathcal{F}_1$ (shaded part of fig. 1.4.1), is defined as

$$|\tau_1| < \frac{1}{2}, \qquad \tau_2 > 0.$$

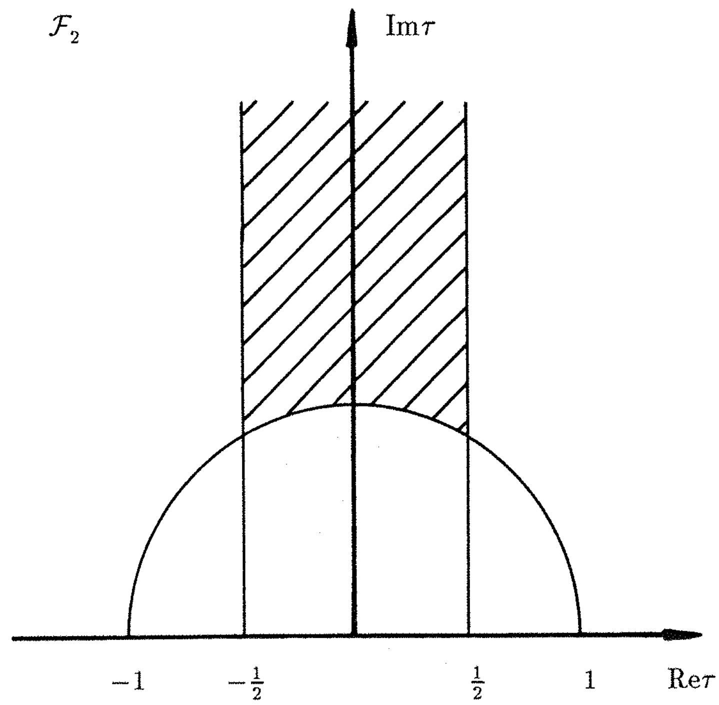


Fig. 1.4.2

Expression (1.4.9) is invariant under $\tau \to \tau + 1$ (i.e., $\tau_1 \to \tau_1 + 1$) as a consequence of (1.4.6). In $D = 26$ it is also invariant under a new, unexpected symmetry; $\tau \to -1/\tau$. These two transformations generate the modular group, a general transformation being of the form

$$\tau \to \frac{\lambda_{11}\tau + \lambda_{12}}{\lambda_{21}\tau + \lambda_{22}}$$

$\lambda \in SL(2, \mathbb{Z})$. Therefore (1.4.9) is an infinite-fold copy of the integral restricted to the smallest region consistent with these symmetries, the fundamental region $\mathcal{F}_2$ (shaded part of fig 1.4.2) defined by

$$|\tau_1| < \frac{1}{2}, \qquad \tau_2 > 0, \qquad |\tau| > 1.$$

The above field theory result will now be compared to the string theory result[66] obtained directly from the Polyakov path integral

$$\int_{\text{torus}} \frac{\mathcal{D}X^\mu \mathcal{D}g_{ab}}{V_{\text{Weyl}} V_{\text{diff}}} e^{-S}$$

where $S = \frac{T}{2} \int d^2\sigma \sqrt{g} g^{ab} \partial_a X^\mu \partial_b X_\mu$. The volumes of the local Weyl and diffeomorphism groups have been factored out. Working in $D = 26$ (so that the Weyl symmetry is still exact) the contribution to the cosmological constant is

$$-T^{13} \int_{\mathcal{F}_2} \frac{d^2\tau}{4\pi\tau_2^2} e^{4\pi\tau_2} (2\pi\tau_2)^{-12} \Big| \prod_{n>0} (1 - e^{2\pi i n\tau}) \Big|^{-48}$$

which is the same as (1.4.8) but integrated over $\mathcal{F}_2$. Thus the field theory calculation infinitely overcounts the string theory result due to the fact that the $\tau \to -1/\tau$ symmetry is not evident when starting from (1.4.1).

This symmetry is evident in the calculation of torus embeddings from the doubly periodic boundary conditions used when representing the torus by a square with opposite sides identified:

$$\begin{aligned} X^\mu(\sigma^0+m,\sigma^1+n) &= X^\mu(\sigma^0,\sigma^1) \\ g_{ab}(\sigma^0+m,\sigma^1+n) &= g_{ab}(\sigma^0,\sigma^1) \end{aligned} \tag{1.4.10}$$

$(m,n \in \mathbf{Z})$ which define a unit cell $0 < (\sigma^0,\sigma^1) < 1$. With these periodicity requirements a metric cannot be brought to a diagonal form by a diffeomorphism and Weyl rescaling but can be rewritten in the form

$$g = \begin{pmatrix} 1 & \tau_1 \\ \tau_1 & |\tau|^2 \end{pmatrix} .$$

$\tau = \tau_1 + i\tau_2$ and without loss of generality the choice $\tau_2 > 0$ can be made. This has fixed the local invariances but there are still global diffeomorphisms (ones which cannot be produced by continuous deformations from the identity) to be considered. These will generally be of the form

$$\sigma' = \boldsymbol{A}\sigma + c$$

where $\boldsymbol{A}$ and c have constant elements. The periodicity (1.4.10) constrains $\boldsymbol{A}$ and c to have integer entries and $\det \boldsymbol{A} = 1$ (i.e., $\boldsymbol{A} \in SL(2,\mathbf{Z})$). c just corresponds to translations required to bring σ' back into the unit cell. In the coordinate system $\sigma' = \boldsymbol{A}^{-1}\sigma$, g' can be brought to the form

$$g = \begin{pmatrix} 1 & \tau_1' \\ \tau_1' & |\tau'|^2 \end{pmatrix}$$

where

$$\tau' = \frac{A_{11}\tau + A_{12}}{A_{21}\tau + A_{22}}$$

which is just the same as before.

Thus the modular group arises out of considerations of global deformations of the torus. The generator $\tau \to \tau + 1$ corresponds to a Deyn twist $\sigma^1 \to \sigma^1 + 2\pi$ under which one would expect a closed string wavefunction to be invariant (being constrained by $N - \tilde{N}|\Psi\rangle = 0$). However the transformation $\tau \to -1/\tau$ corresponds to the interchange of σ^0 and σ^1 and has no analogue in particle theory which is the reason for the overcounting noted above. Considering the torus as $T^2 \cong S^1 \times S^1$ shows that the transformation is quite reasonable, it being just the interchange of the S^1. The fact that field theory overlooks this suggests that string theory must be fundamentally 'stringy' i.e., that it should not be describable in terms of particles or indeed some particle analogue such as an infinite tower of fields. The generalisation of modular invariance to higher genera involves more complicated groups generated by twists round and interchange of non-contractable cycles. Each order of perturbation theory generates a different group which suggests that these invariances should be included at a fundamental level if calculations are to avoid infinitely overcounting vacuum graphs. Indeed the transformation $\tau \to -1/\tau$ interchanges heavy and light particles which violates usual effective field theory notions.

1.5 Summary of Chapter 1

Prescriptions for calculating string scattering amplitudes have been extensively studied and are all based on (some generalisation of) the first quantised description which is to integrate over configurations of a world sheet with fixed topology weighted by e^{-S}, where S is the Nambu action.[10] These are presumed to be the perturbation expansion (in the string coupling constant) of some deeper system of rules but at present only a few tantalising clues as to what this might be have been discovered.[20,62,64,65] A deeper system of rules is needed at least in the sense that they do not require a perturbation expansion in order to be defined. Technically this is required to be able, for example, to determine whether the ground state of a string theory is a compactification[13] and whether any of the myriad of alternative perturbative string theories are in fact different non-perturbatively (rather than just different perturbations of the same theory).[21,22] Indeed it is not even known how consistently to define the theory when a background solution of the full quantum equations of motion does not obey the classical (tree level) equations required for worldsheet (super) conformal invariance. It would be very surprising if this situation did not occur in general and it is very likely that the solution to this question lies outside the present realm of perturbation theory. Within first quantised perturbation theory, some fascinating but incomplete intuitive understanding of this last question has been gained.[17]

Any quantum theory that can be defined independently of its first quantised expansion has a right to be called "second quantised." An alternative but equivalent definition is to regard a theory where the number of particles (or strings) is not *a priori* specified but becomes a quantum variable as second quantised. But for strings it seems that such a theory will not be in the traditional form of quantum field theory. There are many reasons for suspecting that a string field theory is not the answer. First of all, it appears that string field theory offers no solution to the problem of backgrounds that do not satisfy the tree level consistency equations. Perhaps only the Witten open string field theory action[1] can be consistently defined away from flat space but even here any background must satisfy the tree level equations of motion.[59] In fact a more general cubic action formulation has been derived[3,55] but here also it is not yet known how to define it in a phase where the background fails to satisfy the classical conformal invariance conditions. Another, well known, problem is the one loop vacuum contribution which is given by integrating over configurations of a torus in the first quantised formulation. This automatically has meaning in a non zero background as the first quantum correction to the background field effective action, however the result from closed string field theory is a sum over its component modes and infinitely over-counts the first quantised result.[66] In first quantised language this can be traced to an ignorance of some world sheet global diffeomorphisms (modular transformations) which is unfortunate since without identification under modular transformations non-zero closed string vacuum contributions cannot be finite. This last problem will be encountered with any local field theory which uses as quantum variables the component fields of the string, thus a proper theory will have to choose some other form for quantisation.

It was seen in section 1 that first quantised string theory comes fairly close to fulfilling the phenomenological requirements of a theory of unification. Modulo the inclusion of supersymmetry breaking it is possible to construct a model, the heterotic string, that contains the Standard Model. String theory also satisfies one of the philosophical requirements mentioned in section 2, namely that it predicts the dimension of space-time (albeit to be 10 or 26). The property which makes string theory so exciting is that, at the very least, it provides a consistent theory of quantum gravity. Coupled with the above observations, it was this that led people to consider it a theory of everything. However this turned out to be a little premature since there are areas, such as confinement, that perturbative string theory cannot hope to describe, its perturbation series being non-summable. String theory should be viewed as a low energy, perturbative limit of some deeper theory. This theory would inherit the beneficial aspects which string theory enjoys.

It should have become clear that, in addition to being non-perturbative, a second quantised theory is needed. A theory describing single strings would never be able adequately to describe the infinite condensates necessary to determine (i.e., generate) its own space-time geometry and topology. A second quantised theory should in principle be able to do this (and indeed in the case of string field theory it comes remarkably close).[19,59] The fascinating suggestion that string theory should be consistent on all worldsheets (i.e., all topologies) simultaneously[17] encourages the pursuit of a second quantised theory and in fact implies that a theory where the number and topology of worldsheets become quantum variables may be appropriate. This is reminiscent of topological invariants such as the Euler number which has meaning independent of the metric used to construct it. Perhaps the same is true for worldsheets and strings.

It is desirable, at least in retrospect, that the ultimate theory can be derived from a few fundamental physical or philosophical principles. This would suggest that one should start by considering a fundamental object of arbitrary dimension (i.e., a p-brane) or its trajectory and let the theory determine that dimension. Hence whether the fundamental objects of nature can be considered as particles, strings or membranes etc., may well appear as a solution to some equation of motion or as some consistency condition.

As was noted in the previous section, the downfall of string field theory stemmed from its ignorance of modular invariance. Therefore, in order to avoid the same problems of over counting, a second quantised theory must incorporate modular invariance (and duality) at a fundamental level. This can be achieved by forgoing the introduction of a worldsheet parameterisation.

Although it seems that second quantised string theory will not use the traditional form of quantum fields (otherwise it would suffer the same problem of overcounting as string field theory), provisionally it is necessary to be able sometimes to express the background (especially space-time) in the traditional form if there is to be any hope of understanding the results. Rather than aim at an all encompassing theory of everything immediately, it is probably more useful to find one which makes such contact with previous knowledge yet still incorporates some of the above criteria. Such a dichotomous theory is presented in the next chapter.

References

[1] E. Witten, *Nucl. Phys.* **B268** (1986) 253.

[2] E. Witten, *Nucl. Phys.* **B276** (1986) 291.

[3] G.T. Horowitz, J. Lykken, R. Rohm and A. Strominger, *Phys. Rev. Lett.* **57** (1986) 283.

[4] D.J. Gross, J. Harvey, E. Martinec and R. Rohm, *Phys. Rev. Lett.* **54** (1985) 502, *Nucl. Phys.* **B256** (1985) 253 and **B267** (1986) 75.

[5] P. Candelas, G.T. Horowitz, A. Strominger and E. Witten, *Nucl. Phys.* **B258** (1985) 46.

[6] E. Witten, *Phys. Lett.* **155B** (1985) 151 and *Nucl. Phys.* **B258** (1985) 75;
J. Breit, B. Ovrut and G. Segre, *Phys. Lett.* **158B** (1985) 33;
A. Sen, *Phys. Rev. Lett.* **55** (1985) 33.

[7] A. Strominger and E. Witten, *Commun. Math. Phys.* **101** (1985) 341.

[8] D. Gepner, *Nucl. Phys.* **B296** (1988) 757.

[9] L. Dixon, J. Harvey, C. Vafa and E. Witten, *Nucl. Phys* **B261** (1985) 678 and **B274** (1986) 385.

[10] Y. Nambu, Lectures at the Copenhagen Symposium (1970).

[11] M. Dine and N. Seiberg, *Phys. Rev. Lett.* **55** (1985) 366.

[12] See for example C.M. Hull, "Lectures on Sigma Models and Strings," presented at 1986 7^{th} Inst. for Theoretical High Energy Physics, London, England;
P. Howe, "Two Dimensional Nonlinear Sigma Models," lectures given at Super Field Theories Workshop, Vancouver, Canada 1986.

[13] M. Dine and N. Seiberg, *Phys. Lett.* **162B** (1985) 299.

[14] D.J. Gross and V. Periwal, *Phys. Rev. Lett.* **60** (1988) 2105.

[15] C.G. Callan, E. Martinec, M. Perry and D. Friedan, *Nucl. Phys.* **B262** (1985) 593.

[16] A.P. Foakes and N. Mohammedi, *Phys. Lett.* **198B** (1987) 395 and *Nucl. Phys.* **B306** (1988) 343;
S.J. Graham, *Phys. Lett.* **197B** (1987) 543;
R.R. Metsaev and A.A. Tseytlin, *Phys. Lett.* **185B** (1987) 52;
I. Jack, D.R.T. Jones and N. Mohammedi, *Phys. Lett.* **220B** (1989) 171.

[17] C.G. Callan, C. Lovelace, C.R. Nappi and S.A. Yost, *Nucl. Phys.* **B288** (1987) 525;
W. Fischler and L. Susskind, *Phys. Lett.* **171B** (1986).

[18] M. Kaku, *Phys. Rev.* **D38** (1988) 3052.

[19] G.T. Horowitz and D.M. Witt, *Phys. Lett.* **199B** (1987) 176.

[20] E. Witten, *Commun. Math. Phys.* **117** (1988) 353 and *Phys. Lett.* **206B** (1988) 601.

[21] D.J. Gross, in *XXIV Int. Conf. on High Energy Physics*, eds. R. Kotthaus and J.H. Kühn (Springer-Verlag, 1988) 310 and Princeton preprint PUPT-1108 C88/08/04 (1988).

[22] E. Witten, "Search for Higher Symmetry in String Theory," IAS preprint IASSNS-HEP-88/55 (1988).

[23] D. Friedan and S. Shenker, *Nucl. Phys.* **B281** (1987) 509.

[24] S.B. Giddings, E. Martinec and E. Witten, *Phys. Lett.* **176B** (1986) 362;
L. Alvarez-Gaumé, C. Gomez and C. Reina, *Phys. Lett.* **190B** (1987) 55.

[25] M.J. Bowick and S.G. Rajeev, *Phys. Rev. Lett.* **58** (1987) 535.

[26] T. Banks and E. Martinec, *Nucl. Phys.* **B294** (1987) 733.

[27] M. Kaku and K. Kikkawa, *Phys. Rev.* **D10** (1974) 1110,1823.

[28] S. Mandelstam, *Nucl. Phys.* **B64** (1973) 205.

[29] P. Goddard, J. Goldstone, C. Rebbi and C.B. Thorn, *Nucl. Phys.* **B56** (1973) 109.

[30] W. Siegel, *Phys. Lett.* **142B** (1984) 276.

[31] W. Siegel, *Phys. Lett.* **151B** (1984) 391.

[32] M. Kato and K. Ogawa, *Nucl. Phys.* **B212** (1983) 443.

[33] W. Siegel, *Phys. Lett.* **151B** (1984) 396.

[34] W. Siegel and B. Zwiebach, *Nucl. Phys.* **B263** (1986) 105;
K. Itoh, T. Kugo, H. Kunitomo and H. Ooguri, *Prog. Theor. Phys.* **75** (1986) 162;
A. Neveu, H. Nicolai and P.C. West, *Phys. Lett.* **167B** (1986) 307.

[35] T. Banks and M.E. Peskin, *Nucl. Phys.* **B264** (1986) 513.

[36] H. Hata, K. Itoh, T. Kugo, H. Kunitomo and K. Ogawa, *Phys. Lett.* **172B** (1986) 186, 195.

[37] H. Hata, K. Itoh, T. Kugo, H. Kunitomo and K. Ogawa, *Phys. Rev.* **D34** (1986) 2360.

[38] A. Neveu and P.C. West, *Nucl. Phys.* **B293** (1987) 266.

[39] H. Hata, K. Itoh, T. Kugo, H. Kunitomo and K. Ogawa, *Phys. Rev.* **D35** (1987) 1318.

[40] A. Neveu and P.C. West, *Nucl. Phys.* **B268** (1986) 125.

[41] A. Neveu and P.C. West, *Phys Lett.* **168B** (1986) 192.

[42] A. Neveu and P.C. West, *Nucl. Phys.* **B278** (1986) 601.

[43] H. Hata, K. Itoh, T. Kugo, H. Kunitomo and K. Ogawa, *Phys. Rev.* **D35** (1987) 1356.

[44] J. Lykken and S. Raby, *Nucl. Phys.* **B278** (1986) 256.

[45] M.E. Peskin, "Introduction to String and Superstring Theory 2," Lectures presented at 1986 Theoretical Advanced Study Institute in Particle Physics, Santa Cruz.

[46] D.J. Gross and A. Jevicki, *Nucl. Phys.* **B283** (1987) 1 and **B287** (1987) 225.

[47] G. Siopsis, *Phys. Rev.* **D39** (1989) 534.

[48] Z. Qiu and A. Strominger, *Phys. Rev.* **D36** (1987) 1794.

[49] D.Z. Freedman, S.B. Giddings, J.A. Shapiro and C.B. Thorn, *Nucl. Phys.* **B298** (1988) 253 and references therein.

[50] M. Srednicki and R.P. Woodard, *Nucl. Phys.* **B293** (1987) 602.

[51] G.T. Horowitz and A. Strominger, *Phys. Lett.* **185B** (1987) 45;
A. Strominger, *Phys. Lett.* **187B** (1987) 295.

[52] A. Strominger, *Phys. Rev. Lett.* **58** (1987) 629.

[53] S. Sen and R. Holman, *Phys. Rev. Lett.* **58** (1987) 1304.

[54] J. Feng, "A Proposal for Closed Bosonic String Field Theory," Caltech preprint CALT-68-1466 (1987).

[55] A. Strominger, *Nucl. Phys.* **B294** (1987) 93.

[56] M.A. Awada, *Nucl. Phys.* **B286** (1987) 78 and **B287** (1987) 249.

[57] M.A. Awada, *Phys. Lett.* **180B** (1986) 45.

[58] H. Hata, K. Itoh, T. Kugo, H. Kunitomo and K. Ogawa, *Phys. Lett.* **175B** (1986) 138.

[59] T.R. Morris, *Nucl. Phys.* **B297** (1988) 141;
T.R. Morris and B. Spence, *Nucl. Phys.* **B316** (1989) 113.

[60] G.T. Horowitz, J. Morrow-Jones, S.P. Martin and R.P. Woodard, *Phys. Rev. Lett.* **60** (1988) 261;
J. Feng, "New Solutions and Symmetries of the Purely Cubic Bosonic String Field Theory," Caltech preprint CALT-68-1475 (1988).

[61] G.T. Horowitz and S.P. Martin, *Nucl. Phys.* **B296** (1988) 220.

[62] D.J. Gross and P.F. Mende, *Nucl. Phys.* **B303** (1988) 407;
D.J. Gross, *Phys. Rev. Lett.* **60** (1988) 1229.

[63] D.J. Gross and P.F. Mende, *Phys. Lett.* **197B** (1987) 129.

[64] J.J. Atick and E. Witten, *Nucl. Phys.* **B310** (1988) 291.

[65] M. Karliner, I. Klebanov and L. Susskind, *Int. J. Mod. Phys.* **A3** (1988) 1981; I. Klebanov and L. Susskind, *Nucl. Phys.* **B309** (1988) 175.

[66] J. Polchinski, *Commun. Math. Phys.* **104** (1986) 37.

CHAPTER 2

From First to Second Quantised String Theory

In [1] a description of bosonic orientable closed string worldsheets in flat D-dimensional space-time was given in terms of the simultaneous zeroes of $D-2$ functions $f^i(x)$ (where x is the space-time coordinate). The functions can describe any number of worldsheets simultaneously, however the worldsheets are hidden from view allowing the theory to be defined independently of the first quantised expansion (see chapter 3) and thus satisfying the conditions necessary for a second quantised theory. Indeed once properly quantised, the functions f^i will no longer have precise zeroes and the description in terms of worldsheets dissolves. Note that a parameterisation of the worldsheet is avoided (this is to be contrasted with string field theory). Thus the theory manifestly incorporates duality and identifications under modular invariance. Also it is difficult to see how this formulation could do other than solve the problem of consistency on all worldsheets simultaneously. This is because, provided that the description can be consistently quantised, one cannot avoid considering all topologies simultaneously. This is a *bona fide* second quantised string theory unlike string field theory where one first constructs a classical field action by requiring consistency with first quantised string tree level and then one attempts to second quantise.

In this formulation the functions' dynamics are governed by a space-time action which possesses a local $GL(D-2,\mathbb{R})$ invariance. It reduces to the Nambu action for each worldsheet, together with any couplings to background fields. In ref. [1] couplings to a general background metric and massless antisymmetric tensor ($B_{\mu\nu}$) were given. The dilaton coupling and the coupling to any of the massive component fields were not given. The primary purpose of this chapter is to fill this gap. Unfortunately this is not straightforward since the standard couplings of these fields are given in the Polyakov[2] description in terms of an auxiliary worldsheet metric (or equivalent definitions through gauge fixing). However with just a general background metric, massless antisymmetric tensor and *constant* dilaton field the auxiliary worldsheet metric may be equivalently replaced by the induced metric in the sigma-model action. In section 2 an expression for the constant dilaton coupling is found (in terms of the background space-time metric and the functions f^i) by solving for the worldsheet locally in terms of worldsheet coordinates $\sigma^a (a = 0, 1)$ and constructing the induced metric. As is well known, the constant dilaton couples via the Euler number and is thus a topological invariant independent of the space-time embedding and space-time metric. This is far from obvious in the expression derived in section 2. Thus in section 3 independence under perturbations of the f^i (corresponding to small changes in the embedding) and independence of small changes in the space-time metric are explicitly demonstrated.

Although no more than the above may be deduced by a direct comparison with the Polyakov action, the result of section 2 suggests a simple proposal for the general dilaton coupling which is given in section 4. The coupling possesses local $GL(D-2, \mathbb{R})$ invariance and is readily generalised to local $GL(D-2, \mathbb{R})$ invariant expressions for all other couplings as well. However it is argued that the correct method for coupling the non-constant dilaton and massive modes will involve an explicit breaking of $GL(D-2, \mathbb{R})$. Section 4 demonstrates that all local worldsheet functionals of string position and induced metric may be turned into local functionals of f.

In section 5 the results are summarised. First however, in section 1, the results derived in [1] are reviewed and various points mentioned only briefly there are expanded upon, in particular the analysis in a general background metric is completed and a worldsheet projector, which plays an important rôle in the coupling of other background fields, is introduced.

2.1 Implicit Function Approach

In this section the method of describing bosonic closed string worldsheets as the solutions of implicit functions is described. The discussion will be for p-branes where 'brane' is just a term for a general extended object and p is the dimension of the brane's trajectory. The string case can be obtained by setting $p = 2$. For simplicity (arbitrary) closed D-dimensional (euclidean) space-time manifolds, $\mathcal{A}$, will be considered.

It was originally shown in [1] that the flat space action

$$S = \lambda \int_{\mathcal{A}} d^D x \ \delta(f) |df^1 \wedge \cdots \wedge df^{D-p}| \tag{2.1.1}$$

reduced (for $p = 2$) to the Nambu action on choosing a parameterisation of the worldsheet. In the case of strings, $\lambda = 1/2\pi\alpha'$. The proof will be generalised to arbitrary $p < D$ and a general metric later. In curved space with metric $G_{\mu\nu}$

$$|df^1 \wedge \cdots \wedge df^{D-p}| = \{ f^1_{\mu_1} \cdots f^{D-p}_{\mu_{D-p}} f^1_{[\nu_1} \cdots f^{D-p}_{\nu_{D-p}]} G^{\mu_1\nu_1} \cdots G^{\mu_{D-p}\nu_{D-p}} \}^{\frac{1}{2}} \tag{2.1.2}$$

where $f_\mu \equiv \nabla_\mu f = \partial f / \partial x^\mu$ and in general $f_{;\mu\cdots\nu} = \nabla_\nu \ldots \nabla_\mu f$. No factorial weighting is implied in (anti)symmetrisation.

The $D - p$ f^i are smooth functions on D-dimensional space-time, the simultaneous solutions of which define p-dimensional worldsurfaces, $\mathcal{M}^\alpha$ where α labels disconnected topologies. The p-brane worldsurface is hence defined by

$$f^i(x) = 0 \qquad i = 1, \ldots, D - p. \tag{2.1.3}$$

It is of some concern as to whether *all* surfaces, even disconnected ones of any topology, can be described in the above manner however this follows from two theorems of Whitney. The first[3] states that any closed set in $\mathbb{R}^n$ occurs as the solution set $f^{-1}(0)$ for some smooth function $f : \mathbb{R}^n \to \mathbb{R}$ while the second, the famous embedding theorem, allows any submanifold of $\mathcal{A}$ to be considered a submanifold of $\mathbb{R}^n$ for sufficiently large n.

The $f^i(x)$ are restricted by the important condition

$$df^1 \wedge \cdots \wedge df^{D-p} \neq 0 \qquad \text{when} \quad f^i = 0 \tag{2.1.4}$$

which ensures that the df^i are non-vanishing and linearly independent at $f = 0$ (and hence that the delta function $\delta(f) \equiv \prod_{i=1}^{D-p} \delta(f^i)$ is well defined under the space-time integral). f without indices is used to denote the column vector with elements $f_1, f_2, \ldots, f_{D-p}$. The condition also restricts the surfaces to be orientable non-intersecting p-dimensional manifolds without boundary. This follows because the (Hodge) dual of (2.1.4) is a p^{th} rank totally antisymmetric tensor perpendicular to all the $df^i(x)$ which may be used to define a unique p-dimensional oriented tangent space to the solution set $\{x|f(x) = 0\}$ at each point x. This is not possible if any of the above conditions are relaxed. Note also that the description (2.1.3) of the surface is degenerate, there being many equivalent functions f that have the same zeroes. If f' is another such vector it must vanish when f vanishes and therefore without loss of generality the relation

$$f'(x) = \Omega(x)f(x) \tag{2.1.5}$$

may be written where Ω is a $(D-p)$-dimensional real matrix valued smooth function of x. It must also satisfy $\det\Omega(x) \neq 0$ for x away from the surface in order that f' have only the zeroes of f. In addition, requiring that (2.1.4) holds for f' implies that $\det\Omega(x) \neq 0$ when $f(x) = 0$ and hence $\det\Omega(x)$ is non-vanishing everywhere and Ω is invertible. Thus (2.1.5) satisfies the properties of an equivalence relation $f' \sim f$ as indeed it should. Physically (2.1.5) implies that there is a natural local $GL(D-p, \mathbb{R})$ gauge invariance inherent in this description. Next it will be shown that the action (2.1.1) respects this gauge invariance.

The action (2.1.1) can be written in a more useful form by noticing that, after antisymmetrising over internal rather than space-time indices in (2.1.2),

$$|df^1 \wedge \cdots \wedge df^{D-p}| = \sqrt{\det(f_\mu G^{\mu\nu} f_\nu^\top)} = \sqrt{\det \boldsymbol{M}} \equiv \sqrt{M} \tag{2.1.6}$$

where $\boldsymbol{M} = f^\mu f_\mu^\top$. The action can now be written

$$S = \lambda \int d^D x \, \sqrt{G}\sqrt{M}\, \delta(f). \tag{2.1.7}$$

In this form the action is readily seen to be $GL(D-p, \mathbb{R})$ invariant by noticing that any undifferentiated f^i that is integrated against the $\delta(f)$ can be ignored. Thus under $f \to \Omega f$, $f_\mu \to \Omega_\mu f + \Omega f_\mu$ and the action becomes

$$\begin{aligned} S \to & \lambda \int d^D x \, \sqrt{G} \det\Omega^{-1}\, \delta(f) \det(\Omega^\top\Omega)^{\frac{1}{2}} \sqrt{\det(f^\mu f_\mu^\top)} \\ = & \lambda \int d^D x \, \sqrt{G}\sqrt{M}\, \delta(f). \end{aligned}$$

The general curved space p-brane action (2.1.7) can be shown to be equivalent to a sum of Nambu actions (one for each surface $\mathcal{M}^\alpha$) by changing to a local coordinate system consisting of σ^a, local coordinates for one of the worldsurfaces described by f, and $D-p$ other linearly independent coordinates. For these $D-p$ coordinates the functions themselves may as well be used since (2.1.4) guarantees their linear independence. Transforming all tensors in the standard fashion (and using $\partial_a f^k = 0$ in the presence of $\delta(f)$) gives

$$S = \lambda \int d^p\sigma d^{D-p} f\, \delta(f)\sqrt{G}\det(G^{ij}).$$

Using the general relation derived in the appendix, G may be factored into a p-dimensional and $(D-p)$-dimensional determinant:

$$G = \det(\partial_a x^\mu G_{\mu\nu}\partial_b x^\nu)\det(G_{jk} - G_{jb}(G_{ab})^{-1}G_{ak})$$

where $(G_{ab})^{-1}$ is the inverse of the $p\times p$ matrix G_{ab}. Now $G_{kj} - G_{kb}(G_{ab})^{-1}G_{aj}$ can be shown to be the inverse of the $(D-p)\times(D-p)$ matrix G^{ij} by remembering that, since G^{ij} is the $(ij)^{th}$ component of $G^{\mu\nu}$, $G^{ij}G_{jk} = \delta^i_j - G^{ia}G_{ak}$. Hence one obtains

$$S = \lambda \int d^p\sigma \sqrt{\det(\partial_a x^\mu G_{\mu\nu}\partial_b x^\nu)}.$$

which equals the usual Nambu action when $p=2$.

Varying (2.1.7) with respect to f gives

$$\begin{aligned}\delta S &= \lambda \int d^D x\, \sqrt{G}(\sqrt{M}\,\delta f^{\top}\frac{\partial\delta(f)}{\partial f} - \nabla^\mu(\sqrt{M}\,\delta(f) f^{\top}_{\mu} \boldsymbol{N})\delta f) \\ &= -\lambda \int d^D x\, \sqrt{G}\sqrt{M}\,\delta(f)P^{\mu\nu} f^{\top}_{;\mu\nu}\boldsymbol{N}\,\delta f\end{aligned}$$

where $\boldsymbol{N} = \boldsymbol{M}^{-1}$ and $P_{\mu\nu} = G_{\mu\nu} - f^{\top}_{\mu}\boldsymbol{N} f_\nu$. Notice that the $\delta'(f)$ terms cancel and that the properties of the delta function were not explicitly utilised in the above. This also happens later when the equations of motion for the constant dilaton action are derived (in order to prove that it is topological) and will prove useful in section 2 of the next chapter in the context of gauge fixing. The equations of motion for f take the form

$$\delta(f)P^{\mu\nu} f_{;\mu\nu} = 0. \qquad (2.1.8)$$

Note that (2.1.6) implies $\boldsymbol{M}$ is invertible if and only if $df^1\wedge\cdots\wedge df^{D-p} \neq 0$ and hence (2.1.4) ensures that the equations of motion and $P_{\mu\nu}$ are well defined. When the above condition is satisfied, $P_{\mu\nu}$ is a projector onto the p-dimensional subspace orthogonal to all the f^i_μ and in particular projects onto the tangent space of the worldsurface when $f=0$. The properties of $P_{\mu\nu}$ are listed in the appendix along with some useful identities.

At first sight the equations of motion (2.1.8) look quite different from those derived in [1] (for strings), the covariantised form of which are

$$\frac{1}{GM}\delta(f)(*df^1\wedge\cdots\wedge df^{D-2})^{\sigma\mu}(*df^1\wedge\cdots\wedge df^{D-2})_\sigma{}^\nu\,\nabla_\mu\nabla_\nu f^i \equiv \delta(f)P'^{\mu\nu}\,f^i_{;\mu\nu} = 0. \tag{2.1.9}$$

A proof that these are equivalent to (2.1.8) appears in the appendix at the end of this chapter. Henceforth only strings ($p = 2$ and $\lambda = 1/2\pi\alpha'$) will be considered unless otherwise stated.

2.2 Constant Dilaton Coupling

The first step towards writing down a generalisation of the Polyakov couplings in the above formalism was made in [1] for the antisymmetric tensor, $B_{\mu\nu}$, and the metric. These will be reexpressed in terms of the worldsheet projector $P_{\mu\nu}$ in section 4 along with some other couplings. In this section an action for the constant dilaton field is calculated in terms of the functions $f^i(x)$ which reduces to the usual action

$$S_{\Phi_0} = \frac{\Phi_0}{4\pi}\int d^2\sigma\,\sqrt{g}R^{(2)} \tag{2.2.1}$$

on parameterising the worldsheet. $R^{(2)}$ is the worldsheet curvature in terms of the induced metric g_{ab}. This action governs the string theory coupling constant since in the Polyakov loop expansion graphs are weighted by $g^{-(2-2L)}$ where L is the number of loops of a particular order of perturbation theory. This weighting can be incorporated as an additional term in the action by writing $g^{-\chi} = \exp(-\chi \ln g)$ where $\chi = 2-2L$ is the Euler number of the L loop worldsheet. Interpreting $\ln g$ as the constant part of a field Φ, $\ln g = -\Phi_0$, the weighting becomes $\exp(-S_{\Phi_0})$ with $S_{\Phi_0} = -\chi\Phi_0$ which is just (2.2.1). The convention $R^{\mu}{}_{\nu\rho\sigma} = \Gamma^{\mu}_{\nu[\rho,\sigma]} - \Gamma^{\mu}_{\kappa[\rho}\Gamma^{\kappa}_{\sigma]\nu}$ is adopted.

Recalling the proof of equivalence between (2.1.7) and the Nambu action, the dilaton action can be shown to be of the form

$$S_{\Phi_0} = \frac{\Phi_0}{4\pi}\int d^Dx\,\sqrt{G}\sqrt{M}\,\delta(f)\,R \tag{2.2.2}$$

where R is the worldsheet curvature which will be written in terms of f and $P_{\mu\nu}$. Such an expression for $R = R(x)$ will be arrived at by changing to space-time normal (geodesic) coordinates at the point x (which lies in $\mathcal{M}$) and requiring a proximate point, $X^\mu = x^\mu + y^\mu(\sigma)$, to also lie in $\mathcal{M}$. σ^a $(a = 0,1)$ are two parameters for this solution and may be used as local worldsheet coordinates with $y^\mu(0) = 0$. X^μ and x^μ are solutions to (2.1.3) and therefore

$$f(X) - f(x) = y^\mu f_\mu + \tfrac{1}{2}y^\mu y^\nu f_{,\mu\nu} + \cdots = 0. \tag{2.2.3}$$

Resolving y^μ into components normal and tangent to $\mathcal{M}$ enables one to write

$$y^\mu = \sigma^\mu + \lambda^{\tau} f^\mu \tag{2.2.4}$$

with $\sigma^\mu = V_a^\mu\sigma^a$ and $V_a^\mu f_\mu = 0$. The σ^μ lie in the tangent space to $\mathcal{M}$ at x, V_a^μ is a $D\times 2$ matrix and λ is a $D-2$ vector. Substituting (2.2.4) into (2.2.3) determines the parameters λ to be

$$\lambda = -\frac{N}{2}f_{,\mu\nu}\sigma^\mu\sigma^\nu + O(\sigma^3)$$

and hence one finds

$$X^\mu = x^\mu + \sigma^\mu - \tfrac{1}{2} f^{\top}_{,\kappa\lambda} N f^\mu \sigma^\kappa \sigma^\lambda + \tfrac{1}{3} A^\mu_{(\lambda\rho\kappa)} \sigma^\lambda \sigma^\rho \sigma^\kappa + O(\sigma^4). \tag{2.2.5}$$

where a general third order term has been introduced. The next step is to calculate the worldsheet induced metric, $g_{ab} = G_{\mu\nu}(X) X^\mu_{,a} X^\nu_{,b}$, and thence $R(x) = R(X)|_{\sigma=0}$. The expansion of $G_{\mu\nu}(X)$ in the normal coordinates $y^\mu(\sigma) = \sigma^\mu + O(\sigma^2)$ is

$$G_{\mu\nu}(X) = G_{\mu\nu}(x) + \tfrac{1}{2}\sigma^\kappa \sigma^\lambda G_{\mu\nu,\kappa\lambda} + \cdots$$

which coupled with (2.2.5) gives

$$g_{ab} = V_a^\mu V_{\mu b} + (f^{\top}_{,\lambda\mu} N f_{,\kappa\nu} + \tfrac{1}{2} G_{\mu\nu,\kappa\lambda} + A_{\mu(\lambda\nu\kappa)} + A_{\nu(\lambda\mu\kappa)}) \sigma^\kappa \sigma^\lambda V_a^\mu V_b^\nu + O(\sigma^3). \tag{2.2.6}$$

Since $R_{abcd} = \frac{1}{4} R_{[ab][cd]}$, there will always be two indices in the symmetrised part of $A_{\mu(\lambda\nu\kappa)}$ that are antisymmetrised over when calculating R to zeroth order (terms arising from the part of the Riemann tensor bilinear in the connection will be of order σ^2). Therefore the leading contribution to R from this term is zero. From the above convention for the Riemann tensor it is clear that at $\sigma = 0$

$$\begin{aligned} R_{abcd} &= \tfrac{1}{2}(g_{a[c,d]b} - g_{b[c,d]a}) \\ &= V_a^\mu V_b^\nu V_c^\kappa V_d^\lambda (f^{\top}_{,\mu\lambda} N f_{,\nu\kappa} - f^{\top}_{,\mu\kappa} N f_{,\nu\lambda} + R_{\mu\nu\kappa\lambda}). \end{aligned}$$

As a consequence of $V_a^\mu f_\mu = 0$ and the fact that (as shown in the appendix) $P^{\mu\nu}$ is a projector and hence unique, one can write (at $\sigma = 0$) $V_a^\mu V^{\nu a} = P^{\mu\nu}$. Therefore

$$R = R_{ab}{}^{ab} = P^{\mu\kappa} P^{\nu\lambda} (f^{\top}_{;\mu[\lambda} N f_{;\kappa]\nu} + R_{\mu\nu\kappa\lambda}) \tag{2.2.7}$$

where the connections and a general coordinate system have been restored. After rearranging (2.2.7) (see appendix) the constant dilaton action can be written

$$S_{\Phi_0} = \frac{\Phi_0}{4\pi} \int d^D x \, \sqrt{G}\sqrt{M}\, \delta(f) (\nabla_\mu P^{\kappa[\nu} \nabla_\nu P^{\mu]\lambda} P^*_{\kappa\lambda} + P^{\mu\kappa} P^{\nu\lambda} R_{\mu\nu\kappa\lambda}) \tag{2.2.8}$$

where $P^*_{\mu\nu} = G_{\mu\nu} - P_{\mu\nu}$ projects orthogonal to the $P_{\mu\nu}$.

In the next section it is directly verified that the action (2.2.8) is topological, i.e., independent of small changes of f and the metric, and that the new dilaton coupling is indeed the same topological invariant as (2.2.1).

2.3 Topologicality

In the previous section a dilaton coupling was derived which, although it reduces to the usual coupling (2.2.1), it is not manifestly invariant under small changes of f (i.e., embedding). It is also not obvious that it is independent of small variations of the metric as it should also be if it is to be a topological invariant. This section addresses both of these points, proving them by explicit calculation. Prodigious use is made of the identities in the appendix throughout this section. The section is concluded with an example which illustrates that our coupling is the expected topological invariant, the Euler number.

Firstly the contribution to the f equation of motion from (2.2.8) is proven to be identically zero (i.e., that (2.2.8) is invariant under small variations of f). Varying (2.2.2) with respect to f one finds

$$\delta S_{\Phi_0} = \frac{\Phi_0}{4\pi} \int d^D x\, \delta(f) \sqrt{G}\sqrt{M}\,(\delta R - \delta f^\tau N f^\lambda R_{;\lambda} - \delta f^\tau N f_{;\kappa\lambda} P^{\kappa\lambda} R). \tag{2.3.1}$$

With R as defined in (2.2.8), the first term in the brackets is

$$\begin{aligned}
\delta R =& 2\nabla_\mu \delta P^{\rho[\nu} \nabla_\nu P^{\mu]\sigma} P^*_{\rho\sigma} - \nabla_\mu P^{\rho[\nu} \nabla_\nu P^{\mu]\sigma} \delta P_{\rho\sigma} + 2\delta P^{\mu\rho} P^{\nu\sigma} R_{\mu\nu\rho\sigma} \\
=& -2\{ \nabla_\mu (P^{\rho[\nu} \nabla_\nu P^{\mu]\sigma} f^\tau_\sigma N\, \delta f_\rho) - \nabla_\mu P^{\rho[\nu} \nabla_\nu P^{\mu]\sigma} P_{\rho\tau} f^\tau_\sigma N\, \delta f^\tau \} \\
& - \delta P^{\rho[\nu} [\nabla_\mu, \nabla_\nu] P^{\mu]\sigma} P^*_{\rho\sigma} \\
=& -2\{ (P^{\rho\tau} \delta f^\tau_\tau N f^{[\nu} \nabla_\nu P^{\mu]\sigma} + \delta f^\tau_\tau N f^\rho P^{\tau[\nu} \nabla_\nu P^{\mu]\sigma}) \nabla_\mu P_{\rho\sigma} - \delta P^{\mu\rho} P^{\nu\sigma} R_{\mu\nu\rho\sigma} \} \\
=& -2\{ \nabla_\mu (P^{\rho[\nu} \nabla_\nu P^{\mu]\sigma} f^\tau_\sigma N\, \delta f_\rho) + P^{\sigma\rho} \delta f^\tau_\rho N f^\nu P^{\mu\tau} R_{\sigma\tau\nu\mu} \\
& + \delta f^\tau_\tau N\, (f^{[\nu} \nabla_\nu P^{\mu]\sigma} P^{\rho\tau} \nabla_\mu P_{\rho\sigma} - f_\sigma \nabla^\rho P_{\rho\mu} \nabla_\nu P^{\sigma[\mu} P^{\nu]\tau}) \}
\end{aligned} \tag{2.3.2}$$

where $[\nabla_\mu, \nabla_\nu] P^{\rho\sigma} = P^{\tau(\sigma} R^{\rho)}{}_{\tau\nu\mu}$ has been used and it has been noted that

$$f_\sigma \nabla_\mu P^{\rho[\nu} \nabla_\nu P^{\mu]\sigma} P_\rho{}^\tau - f^\rho P^{\tau[\nu} \nabla_\nu P^{\mu]\sigma} \nabla_\mu P_{\rho\sigma} = f_\sigma \nabla^\rho P_{\rho\mu} \nabla_\nu P^{\sigma[\mu} P^{\nu]\tau}.$$

Similarly for the second, $R_{;\lambda}$, term in (2.3.1) one has

$$\begin{aligned}
\nabla_\lambda R = 2\{& \nabla_\mu (P^{\rho[\nu} \nabla_\nu P^{\mu]\sigma} \nabla_\lambda P_{\rho\sigma}) + \nabla_\lambda P^{\tau[\nu} \nabla_\nu P^{\mu]\sigma} P_\tau{}^\rho \nabla_\mu P_{\rho\sigma} \\
& - \nabla_\lambda P_{\tau\sigma} \nabla^\rho P_{\rho\mu} \nabla_\nu P^{\sigma[\mu} P^{\nu]\tau} + P_\rho{}^{[\mu} P^{\nu]\tau} \nabla_\nu P^{\sigma\rho} R_{\sigma\tau\mu\lambda} \\
& + P_\rho{}^\sigma P^{\mu\tau} \nabla_\lambda P^{\rho\nu} R_{\sigma\tau\nu\mu} + P^{\rho\kappa} P^{\sigma\mu} R_{\rho\sigma\kappa\lambda;\mu} \}.
\end{aligned} \tag{2.3.3}$$

This follows from (2.3.2) however an extra Riemann tensor arises when the ∇_λ is commuted through the ∇_μ. The Bianchi identity, $R_{\rho\sigma[\nu\mu;\lambda]} \equiv 0$, was used to rewrite the last term of (2.3.3). Integrating $\delta(f)\sqrt{G}\sqrt{M}\,\delta R$ by parts gives

$$\begin{aligned}
\delta(f)\sqrt{G}\sqrt{M}\,\delta R = \\
&2\sqrt{G}\{\nabla_\mu(\delta(f)\sqrt{M})P^{\rho[\nu}\nabla_\nu P^{\mu]\sigma} f^\tau_\sigma N\,\delta f_\rho \\
&+\delta f^\tau \nabla_\tau(\delta(f)\sqrt{M}\,N(f^{[\nu}\nabla_\nu P^{\mu]\sigma} P^{\rho\tau}\nabla_\mu P_{\rho\sigma} - f_\sigma \nabla^\rho P_{\rho\mu}\nabla_\nu P^{\sigma[\mu}P^{\nu]\tau})) \\
&+\delta f^\tau \nabla_\rho(N f^\nu \delta(f)\sqrt{M}\,P^{\rho\sigma}P^{\mu\tau})R_{\sigma\tau\nu\mu} \\
&+\delta f^\tau N f^\nu \delta(f)\sqrt{M}\,P^{\rho\sigma}P^{\mu\tau}R_{\sigma\tau\nu\mu;\rho} + \text{total derivatives}\} \qquad (2.3.4)
\end{aligned}$$

Integrating the first term in (2.3.4) by parts again cancels the third term and the Riemann derivative cancels that in (2.3.3) (when substituted into (2.3.1)). Under the integral the total derivatives in (2.3.4) can be converted into surface contributions and ignored.

Further manipulations yield

$$\delta R - \delta f^\tau N f^\lambda R_{;\lambda} = 2\delta f^\tau N(f_{;\kappa\rho}P^{\nu\kappa}P^{\mu\tau}P^{\rho\sigma}R_{\sigma\tau\nu\mu} - f_\rho P^{\kappa[\nu}\nabla_\nu P^{\mu]\sigma}\nabla_\mu P_{\tau\sigma}\nabla_\kappa P^{\tau\rho}).$$

Thus the bracketed term in (2.3.1) becomes

$$\delta f^\tau N f_{;\rho\kappa}(f^\tau_{;\sigma\nu} N f_{;\tau\mu} - \tfrac{1}{2}R_{\sigma\tau\nu\mu})(2P^{\tau\rho}P^{\kappa[\nu}P^{\mu]\sigma} - P^{\rho\kappa}P^{\sigma[\mu}P^{\nu]\tau})$$

Rewriting this and substituting it into (2.3.1) gives

$$\delta S_{\Phi_0} = \frac{\Phi_0}{4\pi}\int d^D x\,\delta(f)\sqrt{G}\sqrt{M}\,\delta f^\tau N f_{;\kappa}{}^\rho(f^{\tau}_{;\sigma}{}^{\nu} N f_{;\tau}{}^\mu - \tfrac{1}{2}R_{\sigma\tau}{}^{\nu\mu})P_{[\rho}{}^\tau P_\nu{}^\kappa P_{\mu]}{}^\sigma. \quad (2.3.5)$$

Equation (2.3.5) involves triple antisymmetrisation over projector indices, however these are effectively two dimensional (since $P_{\mu\nu}$ projects into a 2-dimensional tangent space), and hence the equation vanishes. Thus it has been proven that action (2.2.8) is invariant under small changes in f.

Action (2.2.8) should also be independent of small variations of the space-time metric, $G_{\mu\nu}$ and this is now demonstrated explicitly and hence it is proven that (2.2.8) is a topological invariant. Noting the following will facilitate the proof:

$$\begin{aligned}
\delta(\sqrt{G}\sqrt{M}) &= \tfrac{1}{2}\sqrt{G}\sqrt{M}\,P^{\mu\nu}\delta G_{\mu\nu} \\
\delta P^{\rho\sigma} &= -P^{\rho\mu}P^{\sigma\nu}\delta G_{\mu\nu} \\
\delta f_{;\mu\nu} &= -\delta\Gamma^\rho_{\mu\nu} f_\rho \\
\delta R_{\mu\nu\kappa\lambda} &= \delta G_{\mu\tau}R^\tau{}_{\nu\kappa\lambda} + G_{\mu\tau}(\nabla_\lambda \delta\Gamma^\tau_{\nu\kappa} - \nabla_\kappa \delta\Gamma^\tau_{\nu\lambda})
\end{aligned}$$

where the last of these is the Palatini identity. Rewriting the curvature, (2.2.7), as

$$R = P^{\lambda[\nu} P^{\mu]\kappa} f^{\tau}_{;\mu\lambda} N f_{;\nu\kappa} + P^{\mu\kappa} P^{\nu\lambda} R_{\mu\nu\kappa\lambda}$$

and varying with respect to $G_{\mu\nu}$ gives

$$\begin{aligned}\delta R = &- 2P^{\lambda\rho} P^{\sigma[\nu} P^{\mu]\kappa} f^{\tau}_{;\mu\lambda} N f_{;\nu\kappa} \delta G_{\rho\sigma} + P^{\lambda[\nu} P^{\mu]\kappa} f^{\tau}_{;\lambda\mu} N f^{\rho} f^{\tau}_{;\nu\kappa} N f^{\sigma} \delta G_{\rho\sigma} \\ &- 2P^{\lambda[\nu} P^{\mu]\kappa} \delta\Gamma^{\tau}_{\mu\lambda} f^{\tau}_{;\nu\kappa} N f_{\tau} - 2P^{\mu\rho} P^{\kappa\sigma} P^{\nu\lambda} \delta G_{\rho\sigma} R_{\mu\nu\kappa\lambda} \\ &+ P^{\mu\kappa} P^{\nu\lambda} \delta G_{\mu\tau} R^{\tau}{}_{\nu\kappa\lambda} + \nabla_{\lambda} \delta\Gamma^{\mu}_{\nu\kappa} P^{[\kappa}_{\mu} P^{\lambda]\nu} .\end{aligned} \tag{2.3.6}$$

In the following analysis the $\int d^D x \sqrt{G}\sqrt{M}\, \delta(f)$ will be omitted for clarity but accounted for implicitly when integrating by parts. Thus integrating the last term of (2.3.6) by parts and adding the third term to it gives

$$f^{\tau}_{;\tau\lambda} N f^{(\nu} P^{\mu)[\kappa} P^{\lambda]\tau} G_{\mu\sigma} \delta\Gamma^{\sigma}_{\nu\kappa} . \tag{2.3.7}$$

It is now possible to change to geodesic coordinates and write $G_{\mu\sigma}\delta\Gamma^{\sigma}_{\nu\kappa} = \delta\{\mu, \nu\kappa\}$. From (2.3.7) one can clearly see that only the $(\mu\nu)$ symmetric part of $\delta\{\mu,\nu\kappa\}$, the Christoffel symbol, will survive and so it can be replaced by $\frac{1}{2}\delta G_{\mu\nu,\kappa}$. When this is integrated by parts again and gathered up with the remaining terms of (2.3.6) one gets

$$\delta R = -(P^{\rho[\lambda} P^{\kappa]\mu} f^{\tau}_{;\rho\kappa} N f_{;\sigma\lambda} P^{\nu\sigma} + P^{\tau\lambda} P^{\rho\nu} P^{\kappa\mu} R_{\tau\rho\lambda\kappa}) \delta G_{\mu\nu} .$$

Notice how all derivatives of the delta function, obtained when integrating by parts, are removed by projectors. It is interesting that the properties of the delta function were not explicitly used in any of the above manipulations. Therefore under the integral

$$\begin{aligned}\delta(\sqrt{G}\sqrt{M}\, \delta(f) R) = \tfrac{1}{2}\sqrt{G}\sqrt{M}\, \delta(f) \delta G_{\mu\nu} (f^{\tau}_{;\rho\kappa} N f_{;\sigma\lambda} + \tfrac{1}{2} R_{\rho\sigma\lambda\kappa}) \\ \times (P^{\mu\nu} P^{\rho[\lambda} P^{\kappa]\sigma} - 2P^{\nu\sigma} P^{\rho[\lambda} P^{\kappa]\mu})\end{aligned}$$

and finally one has

$$\delta S_{\Phi_0} = \frac{\Phi_0}{8\pi} \int d^D x \sqrt{G}\sqrt{M}\, \delta(f) \delta G_{\mu\tau} G^{\tau\nu} (f^{\tau\kappa}_{;\rho} N f_{;\sigma}{}^{\lambda} + \tfrac{1}{2} R_{\rho\sigma}{}^{\lambda\kappa}) P_{[\nu}{}^{\mu} P_{\lambda}{}^{\rho} P_{\kappa]}{}^{\sigma} \tag{2.3.8}$$

which vanishes for the same reasons as (2.3.5).

This completes the proof that the constant dilaton coupling derived in the previous section is a topological invariant i.e., that it is independent of small fluctuations in both f and $G_{\mu\nu}$. All that remains is to verify that it is indeed the Euler number.

Example: 2-sphere in 4 flat dimensions.

There will be two f^i and the worldsheet is to be described as the intersection of a 3-sphere at the origin of radius a and the plane $X^1 = b$. The solution $f = 0$ describes a 2-sphere when $|a| > |b|$; there are no solutions for $|a| < |b|$. Writing

$$f = \begin{pmatrix} X^2 - a^2 \\ X^1 - b \end{pmatrix}$$

one finds that $R = -2/Y^2$ where $Y^\mu = X^\mu|_{X^1=0}$. The X^1 integral can be done immediately and one can perform the remaining one by changing to 'hypercylindrical' coordinates, $Y^2 = r^2$;

$$\begin{aligned} \chi = -S_{\Phi_0}/\Phi_0 &= \frac{1}{\pi}\int d^4X \, \frac{|Y|}{Y^2}\delta(X^2 - a^2)\delta(X^1 - b) \\ &= 2\int_0^\infty dr^2 \, \delta(r^2 + b^2 - a^2) \\ &= 0 \quad \text{if} \quad |a| < |b| \\ &= 2 \quad \text{if} \quad |a| > |b|. \end{aligned}$$

This gives the correct value of the Euler number for the sphere. Similar analysis yields $S_{\Phi_0} = 0$ for the torus. Thus in this example the equivalence of the actions (2.2.1) and (2.2.8) has been demonstrated. Note that $\det M = 4Y^2$ is non-zero as required by (2.1.4) and (2.1.6) except when $a = b$, where the 3-sphere degenerates to a point.

2.4 More Backgrounds

In this section couplings for a general metric $G_{\mu\nu}(x)$ and massless antisymmetric tensor $B_{\mu\nu}(x)$ are given[1] along with the couplings for the dilaton and fields at the first mass level. The coupling to the metric is:

$$S_G = \frac{1}{2\pi\alpha'}\int d^D x\ \delta(f)\sqrt{G}\sqrt{M}. \tag{2.4.1}$$

This is equivalent to a sum of Nambu actions in a general background metric (one for each worldsheet):

$$S_G = \frac{1}{2\pi\alpha'}\int d^2\sigma\sqrt{\det(\partial_a x^\mu G_{\mu\nu}\partial_b x^\nu)} \tag{2.4.2}$$

which in turn is equivalent to the usual kinetic term in the Polyakov action. The coupling to $B_{\mu\nu}$ is

$$S_B = \frac{1}{4\pi\alpha'}\int d^D x\ \delta(f) df^1 \wedge \cdots \wedge df^{D-2} \wedge B. \tag{2.4.3}$$

It is clear from integration by parts that this has the same linear gauge invariance $\delta B = d\Lambda$ as the worldsheet coupling. (2.4.3) may conveniently be rewritten

$$S_B = \frac{1}{4\pi\alpha'}\int d^D x\ \delta(f)\sqrt{G}\sqrt{M}\,\varepsilon^{\mu\nu}B_{\mu\nu} \tag{2.4.4}$$

where

$$\varepsilon^{\mu\nu} = (GM)^{-\frac{1}{2}}\varepsilon^{\mu\nu\mu_1\cdots\mu_{D-2}}\partial_{\mu_1}f^1\ldots\partial_{\mu_{D-2}}f^{D-2}.$$

Taking into account the fact that $\varepsilon^{\mu\nu\mu_1\cdots\mu_{D-2}}$ transforms as a density it is readily seen that $\varepsilon^{\mu\nu}$ is a pseudotensor (i.e. axial). This implies, for consistency of (2.4.3 and 4), that $B_{\mu\nu}$ is also axial. $\varepsilon^{\mu\nu}$ is a pseudoscalar under the internal $GL(D-2,\mathbb{R})$ in the presence of $\delta(f)$. Under the change of coordinates described in section 1, in proving equivalence of (2.4.1) to the Nambu action, (2.4.3) becomes

$$S_B = \frac{1}{4\pi\alpha'}\int d^2\sigma\,\varepsilon^{ab}\partial_a x^\mu\partial_b x^\nu B_{\mu\nu}\left|\frac{\partial x^\gamma}{\partial f^j}\frac{\partial x^\gamma}{\partial\sigma^a}\right|\det\left(\frac{\partial f^j}{\partial x^\gamma}\frac{\partial\sigma^a}{\partial x^\gamma}\right) \tag{2.4.5}$$

which is the usual $B_{\mu\nu}$ coupling up to a sign depending on the relative orientation of x^μ versus (σ^a, f^i). Comparison of (2.4.5) with (2.4.1,2 and 4) shows that, up to a possible sign, $\varepsilon^{\mu\nu}$ reduces to $\frac{1}{\sqrt{g}}\varepsilon^{ab}\partial_a x^\mu\partial_b x^\nu$ on the worldsheet (where g is the determinant of the induced metric). Note that, by the analysis in the appendix, $i\varepsilon^{\mu\nu}$ is a 'square root' of the worldsheet projector:

$$P^{\mu\nu} = \varepsilon^\mu{}_\lambda\varepsilon^{\nu\lambda}$$

which is consistent with the fact that $P^{\mu\nu}$ reduces to $\partial_a x^\mu \partial_b x^\nu g^{ab}$ on the worldsheet.

For the sake of completeness the $B_{\mu\nu}$ contribution to the f equations of motion will be derived. Perturbing by δf and integrating by parts terms involving derivatives of δf, (2.4.3) becomes

$$\begin{aligned}\delta S_B = \frac{1}{4\pi\alpha'}\int d^Dx\, \{ &\delta_i(f)\{\delta f^i df^1 \wedge \cdots \wedge df^{D-2} \wedge B - \delta f^1 df^i \wedge df^2 \wedge \cdots \wedge df^{D-2} \wedge B \\ &- \cdots - \delta f^{D-2} df^1 \wedge \ldots \wedge df^i \wedge B\} \\ &+ (-1)^D \delta(f)\{\delta f^1 df^2 \wedge \cdots \wedge df^{D-2} \wedge dB \\ &- \cdots - (-1)^D \delta f^{D-2} df^1 \wedge \ldots \wedge df^{D-3} \wedge dB\}\}\end{aligned}$$

where $\delta_i(f) = \partial\delta(f)/\partial f^i$. The first two lines vanish while the last two provide the corrections to the equations of motion for each component of f. Note that no properties of $\delta(f)$ were used in deriving these equations.

The $B_{\mu\nu}$ and $G_{\mu\nu}$ couplings, together with the constant dilaton coupling, obtained in section 2, are the only ones obtainable from the Polyakov description; any other couplings would not allow a consistent (classical) elimination of the auxiliary worldsheet metric. Nevertheless it is tempting to speculate on the form that such couplings would take in the worldsheet description. Taking into account the form of the constant dilaton coupling in section 2 (2.2.8) one might expect in general:

$$S_\Phi = \frac{1}{4\pi}\int d^Dx\ \delta(f)\sqrt{G}\sqrt{M}\,(\nabla_\mu P^{\kappa[\nu}\nabla_\nu P^{\mu]\lambda} P^*_{\kappa\lambda} + P^{\mu\kappa}P^{\nu\lambda}R_{\mu\nu\kappa\lambda})\Phi(x). \qquad (2.4.6)$$

It is certainly this plus possible terms containing derivatives of Φ. (Problems of mixing between Φ and the metric will be ignored). Note that the worldsheet projectors and $\delta(f)$ ensure that this more general coupling still has local $GL(D-2,\mathbb{R})$ invariance. (2.4.6) reduces to the worldsheet formula

$$S_\Phi = \frac{1}{4\pi}\int d^2\sigma\sqrt{g}R\Phi(x(\sigma))$$

where g_{ab} is the induced metric. This is just the general Polyakov coupling of the dilaton with the auxiliary metric replaced by the induced metric. Using the same prescription on the other Polyakov couplings one may generate couplings for all the massive fields. For example using the fact that $\nabla_a\partial_b x^\mu \partial^a x^\nu \partial^b x^\lambda$ is the worldsheet reduction of $-P^{\sigma\nu}P^{\rho\lambda}f^{\top}_{;\sigma\rho}N f^\mu$ (which may be proved by manipulations similar to those in section 2) together with the worldsheet reductions of $P^{\mu\nu}$, $\varepsilon^{\mu\nu}$ and (2.2.7), one finds that the first massive mode couplings[4]

$$S' = \frac{1}{4\pi\alpha'}\int d^2\sigma\sqrt{g}\{\partial_a x^\mu \partial^a x^\nu \partial_b x^\lambda \partial^b x^\rho B_{\mu\nu\lambda\rho}(x) + \tfrac{1}{2}\sqrt{\alpha'}\nabla_a\partial_b x^\mu \partial^a x^\nu \partial^b x^\lambda \bar{B}_{[\mu\nu]\lambda}$$
$$+ \sqrt{\alpha'}\nabla^a\partial_a x^\mu \partial_b x^\nu \partial^b x^\lambda B_{\mu\nu\lambda}(x)$$
$$+ \alpha'\nabla^a\partial_a x^\mu \nabla^b\partial_b x^\nu \tilde{B}_{\mu\nu}(x) + \alpha' R\partial_a x^\mu \partial^a x^\nu C_{\mu\nu}(x)$$
$$+ \alpha'^{\frac{3}{2}} R\nabla^a\partial_a x^\mu C_\mu(x) + \alpha'^2 R^2 C(x) + \cdots\}$$

are the worldsheet equivalents of

$$S' = \frac{1}{4\pi\alpha'}\int d^D x\ \delta(f)\sqrt{G}\sqrt{M}\,\{P^{\mu\nu}P^{\lambda\rho}B_{\mu\nu\lambda\rho} + \sqrt{\alpha'}\nabla_\sigma P^{\sigma\rho}P^{*\mu}_{\rho}P^{\nu\lambda}B_{\mu\nu\lambda}$$
$$+ \tfrac{1}{2}\sqrt{\alpha'}P^{\sigma\nu}\nabla_\sigma P^{\rho\lambda}P^{*\mu}_{\rho}\bar{B}_{[\mu\nu]\lambda}$$
$$+ \alpha'\nabla_\sigma P^{\sigma\rho}\nabla_\kappa P^{\kappa\tau}P^{*\mu}_{\rho}P^{*\nu}_{\tau}\tilde{B}_{\mu\nu} + \alpha' R P^{\mu\nu}C_{\mu\nu}$$
$$+ \alpha'^{\frac{3}{2}} R\nabla_\kappa P^{\kappa\tau}P^{*\mu}_{\tau}C_\mu + \alpha'^2 R^2 C + \cdots\} \qquad (2.4.7)$$

where "$+\cdots$" stands for terms containing an odd number of ε^{ab} or $\varepsilon^{\mu\nu}$ factors as appropriate. One may readily verify that (2.4.7) has local $GL(D-2,\mathbb{R})$ invariance.

A consideration of the perturbative arguments (in σ^a) presented in section 2 makes clear that all higher mode couplings can be similarly built from terms in N, $G_{\mu\nu}$ and *polynomials* of higher derivatives of f. Since these couplings refer only to a surface expression they will automatically have $GL(D-2,\mathbb{R})$ invariance: the equivalence structure discussed in section 1 must be preserved. This requires appropriate contractions of $\varepsilon^{\mu\nu}$ and $P^{\mu\nu}$ with multiple derivatives of f and thus always allows integration by parts of all but one of the derivatives of f onto $\varepsilon^{\mu\nu}$ or $P^{\mu\nu}$. Thus, as illustrated in (2.4.7) the higher mode couplings may be always given as $GL(D-2,\mathbb{R})$ invariant contractions of the component fields with polynomials of $P^{\mu\nu}$, $\varepsilon^{\mu\nu}$ and their derivatives. The above is a prescription for turning local worldsheet expressions of $x^\mu(\sigma)$ and the induced metric g_{ab} (polynomial in their derivatives) into expressions in this formalism. In fact there is a one to one correspondence between gauge invariant local expressions in $\varepsilon^{\mu\nu}$ and f and gauge invariant local expressions on the worldsheet. The implications of this will be discussed in section 3 of the following chapter in connection with the measure. The action (equations (2.4.1,4,6 and 7) and all higher order terms) inherit an additional local symmetry:

$$\varepsilon^{\mu\nu} \to \varepsilon^{\mu\nu} + f^\tau e^{\mu\nu} \qquad (2.4.8)$$

where $e^{\mu\nu}$ is an arbitrary smooth antisymmetric tensor field transforming contragrediently under $GL(D-2,\mathbb{R})$. $\varepsilon^{\mu\nu}$ can be defined to be invariant under local $GL(D-2,\mathbb{R})$

in the presence of $\delta(f)$. Some of the ways of writing the background couplings (particularly R) are *not* invariant under (2.4.8). The reason is that the different ways of writing the couplings use identities (given in the appendix) which do not hold when $\varepsilon^{\mu\nu}$ is only the usual expression up to terms of order f. The couplings given here are those obtained directly from the worldsheet expressions by the methods of section 2 however they have been integrated by parts just the necessary number of derivatives to turn multi-derivative terms in f into singly derivated f's and hence into $P^*_{\mu\nu}$ or other undifferentiated expressions in $\varepsilon_{\mu\nu}$. It is conjectured that all such couplings have invariance (2.4.8). (This appears to require that couplings when written explicitly in terms of $\varepsilon_{\mu\nu}$ have differentials of $\varepsilon_{\mu\nu}$ only in the form $\varepsilon^{\rho\sigma}\nabla_\sigma\varepsilon_{\mu\nu}$ as is indeed the case with those couplings already mentioned). A proper understanding of whether these are in fact couplings (perhaps with some mixing) of the massive modes requires an understanding of gauge fixing and quantisation. This is the subject of the next chapter. There are perhaps two reasons to suspect that the couplings given are not correct for the massive modes: According to the prescription given above, the coupling of the tachyon would be

$$S_t = \frac{1}{4\pi\alpha'}\int d^Dx\ \delta(f)\sqrt{G}\sqrt{M}\,t(x)$$

which is worryingly close to the metric coupling (2.4.1), for example a non-zero tachyon background would be equivalent to a zero tachyon background by a rescaling of the metric (at least in the case where all the other massive backgrounds vanish). There is no analogue of this in other formulations of string theory and it suggests that the prescription given for deriving the massive field couplings is too limited for a one-one correspondence between these couplings and the component fields. In all other formulations there are conformal anomalies or their analogues which are responsible for example for forcing $D = 26$ in flat space-time. (In string field theory this is an anomaly in its gauge invariance). The anomalies, if non-zero, destroy the equivalence to the Nambu worldsheet description. It is natural to expect that this rôle will be played here by the local $GL(D-2,\mathbf{R})$ invariance, especially since it is directly responsible for equivalence to the Nambu description. Presumably, in analogy, the couplings of non-constant dilaton and massive fields will be through explicitly non-$GL(D-2,\mathbf{R})$ invariant counterterms, or by direct coupling to the Fadeev-Popov ghosts.

2.5 Summary of Chapter 2

In this chapter an approach to second quantised string theory using implicit functions to describe worldsheets, originally proposed in ref. [1] has been reviewed and extended to more general backgrounds. In section 2 the constant dilaton action

$$S_{\Phi_0} = \frac{\Phi_0}{4\pi} \int d^D x \delta(f) \sqrt{G} \sqrt{M}\, P^{\mu\nu} P^{\lambda\sigma} \{ f^{\top}_{;\mu[\lambda} \boldsymbol{N} f_{;\nu]\sigma} + R_{\mu\lambda\nu\sigma} \} \tag{2.5.1}$$

was derived where $\boldsymbol{M}$, $\boldsymbol{N}$ and $P^{\mu\nu}$ (the worldsheet projector) are expressions constructed from f and discussed in sections 1, 4 and the appendix. It is tempting to comment on the similarity between S_{Φ_0} and the Einstein action and it is not inconceivable that quantum corrections to this action will yield the Einstein equations. The constant dilaton action is Φ_0 multiplied by minus the sum of Euler numbers for each worldsheet described by f. This is the same as including a string coupling constant $g = \exp(-\Phi_0)$. Thus, once the quantisation of this system is properly understood, the above action will play an important rôle in comparing this non-perturbative formulation with first quantised string theory. In particular for large Φ_0 it should be possible to derive a perturbative expansion in g equivalent to the first quantised results.

In order to make explicit such a comparison it is necessary to compute S matrix elements. One can generate these by coupling the system to the appropriate background fields which then act as sources. In sections 1 and 4 the couplings of the background metric $G_{\mu\nu}$, and $B_{\mu\nu}$ were derived and it was discussed how the non-constant dilaton and massive component fields may couple. As a by-product this enabled it to be shown that all local worldsheet expressions in $x^\mu(\sigma)$ and the induced metric, which are polynomial in their derivatives, can be turned into local expressions involving $P_{\mu\nu}$, $\varepsilon_{\mu\nu}$ and the background fields, polynomial in their derivatives. This may be important for identifying those requirements on the functional measure which ensure the preservation of unitarity, a point which is addressed in the following chapter. An interesting point uncovered in this chapter, which will also prove useful in chapter 3, is that the variations of S_G, S_B and S_D do not contain derivatives of $\delta(f)$. In fact the same results would be obtained if $\delta(f)$ were replaced by any other function of f (provided that the constraint (2.1.4) hold wherever the function is non-vanishing). It is expected that this phenomenon holds for all the other couplings given in section 4, and is presumably yet another reflection of the fact that classically only the zeroes of f have physical significance.

The constant dilaton coupling (2.5.1) is topological: In section 3, it was directly proven that it is independent of small variations in f and $G_{\mu\nu}$. It has already been noted that the familiar string perturbation theory should be regained as $\Phi_0 \to \infty$. From the point of view of this formulation this would most naturally be achieved by perturbing about a theory described purely by (2.5.1). Such an action would yield a topological field theory. This may have relevance to recent speculations on the true degrees of freedom for string theory. This conjecture is considered in chapter 4.

Appendix

In the appendix the projector $P_{\mu\nu}$ is investigated in more depth and the equivalence of the equations of motion (2.1.8) and (2.1.9) is demonstrated. Firstly, however, the relationship used to factorise the determinant in section 1 is derived. The starting point is the definition of determinant as the integral over (not necessarily conjugate) Grassman odd vectors η and $\bar{\eta}$:

$$\det \boldsymbol{G} = \int d\bar{\eta} d\eta \exp(-\bar{\eta}\boldsymbol{G}\eta). \qquad (A1)$$

Writing

$$\boldsymbol{G} = \begin{pmatrix} \boldsymbol{A} & \boldsymbol{B} \\ \boldsymbol{C} & \boldsymbol{D} \end{pmatrix}, \qquad \eta = \begin{pmatrix} c \\ b \end{pmatrix} \quad \text{and} \quad \bar{\eta} = (\bar{c}, \bar{b}),$$

where the dimension of the Grassman parameters c and $\bar{c}$ is the same as the dimension of the square matrix $\boldsymbol{A}$ (and similarly for b, $\bar{b}$ and $\boldsymbol{B}$), allows the exponent in $(A1)$ to be written

$$\begin{aligned} \bar{\eta}\boldsymbol{G}\eta &= \bar{c}\boldsymbol{A}c + \bar{c}\boldsymbol{B}b + \bar{b}\boldsymbol{C}c + \bar{b}\boldsymbol{D}b \\ &= \bar{b}\boldsymbol{D}b + \bar{c}(\boldsymbol{A} - \boldsymbol{B}\boldsymbol{D}^{-1}\boldsymbol{C})c. \end{aligned} \qquad (A2)$$

In the last line the replacements

$$\begin{aligned} b &\to b - \boldsymbol{D}^{-1}\boldsymbol{C}c \\ \bar{b} &\to \bar{b} - \bar{c}\boldsymbol{B}\boldsymbol{D}^{-1} \end{aligned}$$

have been made. Inserting $(A2)$ into $(A1)$ shows that

$$\det \boldsymbol{G} = \det \boldsymbol{D} \det(\boldsymbol{A} - \boldsymbol{B}\boldsymbol{D}^{-1}\boldsymbol{C}).$$

Returning now to the projector, it is assumed that $df^1 \wedge \cdots \wedge df^{D-p} \neq 0$ so that $\boldsymbol{N} = \boldsymbol{M}^{-1}$ is well defined. $P_{\mu\nu} = G_{\mu\nu} - f_\mu^{\top} \boldsymbol{N} f_\nu$ is symmetric and is a projector because

$$P^{\mu}{}_{\nu} P^{\nu}{}_{\rho} = P^{\mu}{}_{\rho} \qquad (A3)$$

It projects onto the maximal subspace perpendicular to the df^i's as evidenced by

$$P^{\mu\nu} f^i_\nu = 0 \qquad \text{for} \quad i = 1, \ldots, D-p \qquad (A4)$$

$$\text{and} \qquad P_\mu{}^\mu = p \qquad \text{(for a } p\text{-brane)}. \qquad (A5)$$

Some useful properties of the $P_{\mu\nu}$, all of which follow straightforwardly from the above, are listed below;

$$
\begin{aligned}
P^{\mu\nu} f_{;\nu\rho} &= -\nabla_\rho P^{\mu\nu} f_\nu \\
\text{also} \qquad P^{\mu\nu} f_{,\nu\rho} &= -\partial_\rho P^{\mu\nu} f_\nu \\
P^{\mu\nu} f^{\top}_{;\nu\rho} N f_\sigma &= -P^{\mu\nu} \nabla_\rho P_{\nu\sigma} \\
\nabla_\rho P_{\mu\nu} &= -G^{\kappa\lambda} P_{\kappa(\mu} f^{\top}_{\nu)} N f_{;\lambda\rho} \\
\nabla_\mu P_{\rho\sigma} \nabla_\nu P^{\sigma\lambda} P_{\lambda\kappa} &= P_{\rho\sigma} \nabla_\mu P^{\sigma\lambda} \nabla_\nu P_{\lambda\kappa} \\
P_{\mu\nu} \nabla_\rho P^{\nu\sigma} P_{\sigma\kappa} &= 0 \\
\nabla_\rho P_{\mu\nu} P^{\nu\sigma} &= P^*_{\mu\nu} \nabla_\rho P^{\nu\sigma}
\end{aligned}
$$

where $P^*_{\mu\nu} = G_{\mu\nu} - P_{\mu\nu}$ is a projector onto the $D-p$ dimensional subspace spanned by the f_μ. It is expressions such as the last one above which invalidate symmetry under $\varepsilon^{\mu\nu} \to \varepsilon^{\mu\nu} + f^{\top} e^{\mu\nu}$.

Finally a proof that $P'_{\mu\nu} = \frac{1}{GM}(*df^1 \wedge \cdots \wedge df^{D-2})_\mu{}^\sigma (*df^1 \wedge \cdots \wedge df^{D-2})_{\nu\sigma}$ introduced in section 1 equals $P_{\mu\nu}$ is presented. In order to do this it need only be shown that $P'_{\mu\nu}$ possesses properties $(A3-5)$ since these uniquely determine the projector. It is convenient to start with $(A4)$. Now

$$
(*df^1 \wedge \cdots \wedge df^{D-2})^{\mu\nu} = \varepsilon^{\mu_1 \cdots \mu_{D-2}\mu\nu} f^1_{\mu_1} \cdots f^{D-2}_{\mu_{D-2}}
$$

where $\varepsilon^{\mu_1\cdots\mu_D}$ is the totally antisymmetric tensor density in D dimensions. When this is contracted with an f^i_μ it vanishes since there is a combination $f^j_{\mu_j} f^i_\mu$ for some $j = i$ which gets antisymmetrised over by the epsilon tensor. Hence it follows that

$$
P'^{\mu\nu} f^i_\nu = 0 \qquad \text{for} \qquad i = 1, \ldots, D-2
$$

which is property $(A4)$ for $p = 2$. Noticing that

$$
\varepsilon^{\mu_1\cdots\mu_D} \varepsilon_{\nu_1\cdots\nu_D} = G \delta^{\mu_1\cdots\mu_D}_{\nu_1\cdots\nu_D}
$$

implies that

$$
\begin{aligned}
P'_{\mu\nu} P'^{\nu}{}_\rho = \frac{1}{GM^2} &(*df^1 \wedge \cdots \wedge df^{D-2})_\mu{}^\sigma \, \delta^{\rho_1\cdots\rho_{D-2}\nu\kappa}_{\nu_1\cdots\nu_{D-2}\nu\sigma} \\
&\times f^{1,\nu_1} \cdots f^{D-2,\nu_{D-2}} f^1_{\rho_1} \cdots f^{D-2}_{\rho_{D-2}} (*df^1 \wedge \cdots \wedge df^{D-2})_{\rho\kappa}.
\end{aligned}
$$

All but one of the terms generated by the contraction of the epsilon tensors can be ignored by a similar argument to that used in verifying $(A4)$, being killed against one of the two bracketed terms, i.e.,

$$
\delta^{\rho_1\cdots\rho_{D-2}\kappa}_{\nu_1\cdots\nu_{D-2}\sigma} = \delta^\kappa_\sigma \delta^{\rho_1\cdots\rho_{D-2}}_{\nu_1\cdots\nu_{D-2}} - (\delta^\kappa_{\nu_1} \delta^{\rho_1\cdots\cdots\rho_{D-2}}_{\sigma\nu_2\cdots\nu_{D-2}} + \text{ other discardable terms}).
$$

Therefore

$$P'_{\mu\nu}P'^{\nu}{}_{\rho} = \frac{1}{M}P'_{\mu\rho}\,\delta^{\rho_1\cdots\rho_{D-2}}_{\nu_1\cdots\nu_{D-2}}\,f_1^{\nu_1}\cdots f_{D-2}^{\nu_{D-2}}\,f^1_{\rho_1}\cdots f^{D-2}_{\rho_{D-2}} = P'_{\mu\rho}$$

using (2.1.5) and ($A3$) is verified. A more straightforward analysis verifies (A5),

$$P'_{\mu}{}^{\mu} = 2.$$

Hence $P'_{\mu\nu} = P_{\mu\nu}$ and the equations of motion (2.1.8) and (2.1.9) are the same.

References

[1] T.R. Morris, *Phys. Lett.* **202B** (1988) 222.

[2] A.M. Polyakov, *Phys. Lett.* **103B** (1981) 207.

[3] H. Whitney, *Trans. Ams.* **36** (1934) 63.

[4] A. Tseytlin, *Phys. Lett.* **185B** (1987) 59.

CHAPTER 3

GAUGE FIXING AND QUANTISATION

This chapter addresses the gauge fixing and quantisation of the candidate second quantised string theory introduced and developed in chapter 2. A discussion of the need for a second quantised string theory and the motivation behind this particular formulation was given in chapter 1. The candidate theory has action (2.1.1) (with $p = 2$) and at first sight quantisation of the system with this action looks hopeless, some of the major problems being:

(a) It does not have a kinetic term at most quadratic in derivatives.

(b) The action contains the distribution $\delta(f)$ which of course cannot be expanded perturbatively for small fields f.

(c) The local $GL(D-p,\mathbb{R})$ invariance must be gauge fixed but there does not appear to be a sensible gauge fixing constraint which can be applied to the functions f.

(d) Some way of incorporating the constraint (2.1.4) is needed.

(e) The naïve functional measure $(\prod_x df(x))$ in the Feynman path integral is not local $GL(D-p,\mathbb{R})$ invariant. Some other definition of the measure is required which is invariant, and leads to a local unitary quantum theory.

In section 1 it is shown how problem (a) is readily solved by introducing auxiliary fields and also that an action at most quadratic in derivatives may be derived from a partial conventional gauge fixing. A byproduct of this approach is a satisfying solution to problem (d). In section 2 it is shown that replacing $\delta(f)$ with a general function of f leads to an action with gauge fixed equations of motion. Hence this procedure solves both problems (b) and (c), however it is not as yet known how to include the requisite ghosts. The procedure may be applied to the action including auxiliary fields and to actions including the background field couplings introduced in chapter 2. In section 3 the definition of the functional measure is considered. It is shown how to define the measure so that it is local $GL(D-p,\mathbb{R})$ invariant up to possible anomalies. This is done by obtaining a simple fully gauge fixed partition function and then proving that it is independent of gauge fixing parameters. Thus a complete solution to problem (c) is demonstrated. Some arguments are presented for the case of strings which suggest that the full action is renormalisable and unitary if coupled to all the background fields, provided a certain auxiliary field is introduced. The functional measure can only be completely defined by comparison with the hamiltonian approach[1] and therefore this formalism is considered in section 4. Problem (b) is more acute here because $\delta(f)$ enters in the expression for the conjugate momentum which then appears squared in the hamiltonian. However it turns out that, at least in the simplest case of $p = 1$, $D = 2$, the $\delta(f)$ may be completely excised from the formulation while still describing the full gauge invariant equations of motion. The system is subject to only first class constraints which, with a simple ordering prescription, remain first class at the quantum level. The functional form of the physical states is derived and discussed. The case of general p and general D is much more complicated and is not completely understood. The problems are briefly described and some suggestions are made for their solution. Section 5 is a summary of the results together with an evaluation.

3.1 Auxiliary Fields and Conventional Gauge Fixing

In the previous chapter it was shown that equation (2.1.1) may be rewritten

$$S = \lambda \int d^D x \, \delta(f) \sqrt{G} \sqrt{M} \tag{3.1.1}$$

where $M = \det \boldsymbol{M}$ and $\boldsymbol{M} = \partial_\mu f G^{\mu\nu} \partial_\nu f^\top = \partial_\mu f \partial^\mu f^\top$. Condition (2.1.4) corresponds to $\det \boldsymbol{M} \neq 0$ at $f = 0$. General p is being considered. λ is a parameter of dimensions $(\text{mass})^p$ and is the mass for a massive particle ($p = 1$) or the string tension $T = 1/2\pi\alpha'$ for $p = 2$. An auxiliary $(D-p) \times (D-p)$ symmetric matrix field $\boldsymbol{g}$ may be introduced and the action written

$$S = \frac{1}{2} \int d^D x \, \delta(f) \sqrt{G} \, (\sqrt{g})^{-1} (\partial^\mu f^\top \boldsymbol{g} \, \partial_\mu f - \Lambda) \tag{3.1.2}$$

where $\Lambda = (D - p - 2)\lambda^{-2/(D-p-2)}$, $g = \det \boldsymbol{g}$ and it is assumed for the moment that $D \neq p + 2$. Some restriction on $\boldsymbol{g}$ is required in order for equation (3.1.2) to be well defined. A sufficient condition is that $\boldsymbol{g}$ be positive definite (and hence invertible) when $f = 0$. Action (3.1.2) is classically equivalent to action (3.1.1) by elimination of $\boldsymbol{g}$ via its equation of motion. Its equation of motion implies

$$\delta(f)(\boldsymbol{g}^{-1} - \frac{D-p-2}{\Lambda} \partial^\mu f \partial_\mu f^\top) = 0. \tag{3.1.3}$$

Thus classically the requirement that $\boldsymbol{g}$ be invertible at $f = 0$ incorporates the constraint (2.1.4) since this constraint is equivalent to requiring that $\partial^\mu f \partial_\mu f^\top$ be invertible at $f = 0$ (i.e., that $\det \boldsymbol{M} \neq 0$). Action (3.1.2) still has local $GL(D-p, \mathbb{R})$ invariance which may be incorporated as ($\Omega(x) \in GL(D-p, \mathbb{R})$):

$$\begin{aligned} f(x) &\to \Omega(x) f(x) \\ \boldsymbol{g}(x) &\to \Omega^{-\top}(x) \boldsymbol{g}(x) \Omega^{-1}(x) \end{aligned} \tag{3.1.4}$$

where $\Omega^{-\top}$ means the transpose of the inverse of Ω. In addition, the formulation inherits a new gauge transformation

$$g_{ij}(x) \to g_{ij}(x) + h_{ijk}(x) f^k(x) \tag{3.1.5}$$

where $h_{ijk}(x)$ is any smooth 3^{rd} rank $GL(D-p, \mathbb{R})$ tensor field (with the restriction that it be symmetric over its first two indices).

Some alterations are required if $D = p + 2$. Action (3.1.2) is replaced by

$$S = \frac{\lambda}{2} \int d^D x \, \delta(f) \sqrt{G} \, (\sqrt{g})^{-1} \partial^\mu f^\top \boldsymbol{g} \, \partial_\mu f. \tag{3.1.6}$$

This system has, in addition to gauge invariances (3.1.4 and 5), a further local invariance:

$$g(x) \rightarrow w(x) g(x) \tag{3.1.7}$$

where $w(x)$ is any scalar field restricted to be positive when $f = 0$. (This invariance is not independent of (3.1.5) when $f \neq 0$). Once again this system is equivalent to equations (2.1.1 and 2.1.4).

This completes the solution of problem (a). Many similarities will have been noticed in this discussion to the introduction of worldsheet auxiliary metrics in strings and membranes however they are only similarities. Following the analysis given in chapter 2, actions (3.1.2 and 6) have been reduced to worldsheet actions. They are not found to reduce to natural worldsheet actions since the field $g(x)$ is not related to a worldsheet auxiliary metric.

It is interesting to note that the restriction on g and the gauge invariance (3.1.4) are sufficient to gauge fix $g = \mathit{1}$ when $f = 0$ i.e.,

$$\delta(f)(g - \mathit{1}) = 0.$$

This fixes the gauge group only at $f = 0$ and fixes it down to a local $SO(D-p)$ there. Essentially the same effect can be gained from the original action (3.1.1) by gauge fixing to

$$\delta(f)(M - \mathit{1}) = 0 \tag{3.1.8}$$

This neatly incorporates the restriction (2.1.4) on the original formalism. Standard methods of BRST[2] will be used to simplify the action. In this approach the first step is to replace the gauge transformation by a nilpotent BRS transformation generated by a general anticommuting ghost field c. The gauge fixing constraint, $\mathcal{G} = 0$, is then introduced as the equation of motion of a Lagrange multiplier β via the addition of the lagrangian

$$\mathcal{L}_{\mathrm{gf}} = \mathrm{tr}\beta\mathcal{G}.$$

Writing this as $\frac{1}{2}\{\delta\mathrm{tr}\bar{c}\mathcal{G} + \mathrm{tr}\bar{c}\delta\mathcal{G}\}$ implies that in order to keep the total lagrangian BRS invariant, an additional (ghost) term

$$\mathcal{L}_{\mathrm{gh}} = -\tfrac{1}{2}\mathrm{tr}\bar{c}\delta\mathcal{G}$$

must be added. $\bar{c}$ is an antighost field and $\delta\bar{c} = 2\beta$ (and hence $\delta\beta = 0$).

In the case of (3.1.1) the BRS transformations are

$$\delta f = cf, \qquad \delta c = c^2$$

where c is a Grassman odd matrix field and its BRS transformation ensures nilpotency of δ on f. Also $\mathcal{G} = \delta(f)(\boldsymbol{M} - \boldsymbol{1})$ which implies that $\bar{c}$ is a symmetric matrix antighost field. Thus (3.1.1) can be written

$$S = \int d^D x\, \delta(f)\sqrt{G}\{\lambda\sqrt{M} + \mathrm{tr}[\beta(\boldsymbol{M} - \boldsymbol{1}) + \tfrac{1}{2}\bar{c}(\boldsymbol{M} - \boldsymbol{1})\mathrm{tr}c - \tfrac{1}{2}\bar{c}(c\boldsymbol{M} + \boldsymbol{M}c^{\top})]\}.$$

Integrating out β (amounting to setting $\boldsymbol{M} = \boldsymbol{1}$) gives the much simplified action

$$S = \int d^D x\, \delta(f)\sqrt{G}(\lambda - \mathrm{tr}\bar{c}c). \tag{3.1.9}$$

The constraint (3.1.8) will appear as a delta function in the functional integral. Now that β has been eliminated the BRS transform of $\bar{c}$ will no longer be $\delta\bar{c} = 2\beta$ but can be found by insisting that $\delta S = 0$. The remaining local invariance of $GL(D-p, \mathbb{R})$ acts on c and $\bar{c}$ in the same way as on g in equation (3.1.4). c has a linear gauge invariance of the form $\delta c = \omega$ where ω is antisymmetric expressing the fact that (3.1.9) is independent of any antisymmetric component of c. In addition c and $\bar{c}$ inherit gauge invariances of the form of equation (3.1.5).

In order to calculate the gauge fixed equations of motion it is helpful to include the gauge fixing term explicitly in (3.1.9) and write

$$S = \int d^D x\, \delta(f)\sqrt{G}\{\lambda + \mathrm{tr}[\beta(\boldsymbol{M} - \boldsymbol{1}) - \bar{c}c]\}.$$

With the c, $\bar{c}$ and β equations of motion enforced, the BRS transform of the antighost $\bar{c}$ now changes to

$$\delta\bar{c} = -\lambda + 2\beta.$$

Condition (3.1.8) implies that

$$\boldsymbol{M} = \boldsymbol{1} + Af + O(f^2)$$

where A is a rank 3 tensor symmetric on its first two indices. Differentiating this at $f = 0$ enables A to be determined and hence that

$$\boldsymbol{M} = \boldsymbol{1} + (f_{;}^{\mu\nu} f_{\mu}^{\top} + f_{\mu} f_{;}^{\top\mu\nu}) f_{\nu}^{\top} f + O(f^2). \tag{3.1.10}$$

The classical gauge fixed equations of motion of (3.1.9) can now be calculated. Enforcing the (anti)ghost equations of motion leaves the f equation

$$\delta'(f)\{\lambda + \mathrm{tr}[\beta(M - 1)]\} - 2\partial_\mu(\delta(f)\beta f^\mu) = 0.$$

Substituting in the expression for M given by (3.1.10) yields

$$\delta'(f)(\lambda - 2\beta) - 2\delta(f)\{\partial_\mu(\beta f^\mu) - f_\mu^T f_\nu \beta f_;^{\mu\nu}\}.$$

The antighost equation of motion implies $\delta\bar{c} = 0$ which in turn implies that $\lambda = 2\beta$. Hence the above gauge fixed equations of motion take the correct form:

$$\delta(f)\{\Box f - (f_\mu^T f_\nu) f_;^{\mu\nu}\} = 0 \tag{3.1.11}$$

where $\Box f = \nabla^\mu \partial_\mu f$. It is not clear how to make further progress with this action however. It is necessary to understand how to deal with the $\delta(f)$ in performing the functional integral and one is still left with a large gauge invariance which will require further gauge fixing. It does not seem possible to find a sensible *further* gauge fixing which will remove all the gauge invariance.

3.2 Unconventional Gauge Fixing

An intriguing direction for avoiding the problem with $\delta(f)$ is to formulate an action with functions f and auxiliary fields, but without the $\delta(f)$, in such a way as to respect the $GL(D-p,\mathbf{R})$ invariance on the functions f (at least on-shell). Only the zeroes of f will then have physical significance and the action will be classically equivalent to a worldsurface action. Unfortunately attempts in this direction have only led to actions with inconsistent or over constrained equations of motion.

It is however possible to formulate actions that do not contain $\delta(f)$, whose equations of motion are gauge fixed versions of the equations that follow from action (2.1.1). This possibility arises because, as noted in chapter 2, the derivation of the equations of motion from equations (2.1.1) or (3.1.1);

$$\frac{\delta S}{\delta f} = 0 \quad \Rightarrow \quad \delta(f)\sqrt{G}\,\sqrt{M}\,\boldsymbol{N}\,P^{\mu\nu} f_{;\mu\nu} = 0, \tag{3.2.1}$$

where $\boldsymbol{N} = \boldsymbol{M}^{-1}$ and $P_{\mu\nu} = G_{\mu\nu} - f_\mu^{\scriptscriptstyle T} \boldsymbol{N} f_\nu$, do not require the use of any of the properties of $\delta(f)$. Thus

$$S = \lambda \int d^D x\, \Delta(f)\sqrt{G}\,\sqrt{M}\,, \tag{3.2.2}$$

where $\Delta(f)$ is *any* function of f, gives

$$\Delta(f)\sqrt{G}\,\sqrt{M}\,\boldsymbol{N}\,P^{\mu\nu} f_{;\mu\nu} = 0. \tag{3.2.3}$$

However the equations of motion (3.2.3) are more specialised than equations (3.2.1) which, local to $f = 0$, imply

$$\sqrt{G}\,\sqrt{M}\,\boldsymbol{N}\,P^{\mu\nu} f_{;\mu\nu} = \boldsymbol{A} f \tag{3.2.4}$$

where $\boldsymbol{A}$ is some matrix field. Globally this equation is ill defined because $\boldsymbol{M}$ cannot in general be invertible for all x. (M vanishes for example whenever one of the f^i has a maximum, which will always occur for bounded functions on a closed space-time manifold. Even on open manifolds, M will vanish somewhere if the surface described by $f = 0$ has topology other than $\mathbf{R}^p$). For the same reason some constraint on $\Delta(f)$ is required so that equation (3.2.3) has global solutions. It is clear that this requires $\Delta(f)$ to vanish for some value of f. $\Delta(f)$ will be chosen to be any function with support in some bounded region $\mathcal{R}$ in $\mathbf{R}^{D-p}$ enclosing the origin. This will turn out to be a sufficient condition for equation (3.2.3) to have globally well defined solutions which are gauge fixed solutions of equation (3.2.1). Equation (3.2.3) implies equation (3.2.4) with $\boldsymbol{A}$ vanishing whenever $\Delta(f) \neq 0$. Under a general infinitessimal $GL(D-p, \mathbf{R})$ gauge transformation, with $\Omega(x) = e^{\phi}$, equation (3.2.4) generates terms of the form:

$$\sqrt{G}\sqrt{M}\,(\mathrm{tr}\phi - \phi^{\mathsf{T}})\boldsymbol{N}\, P^{\mu\nu} f_{;\mu\nu} + \boldsymbol{B}(\phi) f.$$

Since there are no local restrictions on $\boldsymbol{B}$ and since it has the same number of degrees of freedom as $\boldsymbol{A}$, $(D-p)^2$, it follows that ϕ can be chosen such that the extra terms generated can be absorbed into the left or right sides of (3.2.4) in such a way as to (locally) gauge fix $\boldsymbol{A}$ to zero. Again, global restrictions arise from points where $M = 0$ and these can be moved about but not in general all removed by a gauge transformation. To remove them all in general requires a singular gauge transformation which instead causes f to diverge at some points. However, with the above definition of $\Delta(f)$ and restriction (2.1.4), it is always possible to ensure that $M = 0$ only when $\Delta(f) = 0$ (e.g., by a global scaling of the f's) while keeping $\boldsymbol{A}$ zero when $f \in \mathcal{R}$. Thus equations of motion (3.2.3) are partially gauge fixed versions of the gauge invariant equations of motion (3.2.1). They are only partially gauge fixed because equations (3.2.2 and 3) are independent of $f(x)$ when $f(x) \notin \mathcal{R}$. It may be possible to give definitions of Δ which much further constrain the remaining (non-linear) gauge invariance (for example if f is allowed to diverge at some points and Δ is required to vanish sufficiently fast as $f \to \infty$, equation (3.2.3) will be a globally well defined completely gauge fixed version of equation (3.2.1)). Also, for any definition of Δ there is some on shell non-linear gauge symmetry. It can be shown that equation (3.2.3) is also solved by an f transformed in the following manner:

$$f \to \Omega(f) f \qquad \Omega(f) \in GL(D-p, \mathbf{R})$$

where Ω is any smooth invertible matrix function of the f's. This is because, under a similar analysis given above for (3.2.4), the terms in $\boldsymbol{B}$ not proportional to (3.2.3) are removed, i.e., $\delta(\Delta\sqrt{G}\sqrt{M}\,\boldsymbol{N}\,P^{\mu\nu} f_{;\mu\nu}) \propto \sqrt{G}\sqrt{M}\,\boldsymbol{N}\,P^{\mu\nu} f_{;\mu\nu}$. The fact that a Δ' term is generated does not matter since it is non-zero in the same region as Δ. Thus the content of (3.2.3) is unaltered by the above transformation however it is not a symmetry of the gauge fixed action (3.2.2).

It has been shown that action (3.2.2) is classically equivalent to a gauge fixed action (3.1.1) and it is therefore suspected that action (3.1.1) (possibly reformulated with auxiliary fields) may be gauge fixed to an action proportional to (3.2.2). Indeed formally one can display the general action (3.2.2) as the gauge invariant action plus an infinite number of gauge constraints, each of which gauge fixes only partially (gauge transformations of the form of a field contracted into a high enough number of f's remain un-gauge fixed by any finite number of terms), by writing

$$\Delta(f) = \delta(f) + \frac{1}{2!}a^{ij}\delta_{ij}(f) + \frac{1}{4!}a^{ijkl}\delta_{ijkl}(f) + \cdots$$

where the coefficients $a^{ij\cdots}$ are real numbers and $\delta_{ij}(f) = \partial^2\delta(f)/\partial f^i \partial f^j$ etc. This may be shown by Fourier transforming with respect to f and noting the invariance $\Delta(f) = \Delta(-f)$ which Δ inherits from action (3.2.2). It is also required that $\int d^{D-p} f \Delta(f) = 1$. It has not yet been figured out how to add the required BRST ghosts to action (3.2.2) however the above equation may help in solving this problem.

One way of thinking of action (3.2.2) is as an average over actions S_k:

$$\begin{aligned} S &= \int d^{D-p}k \Delta(k) S_k \\ S_k &= \lambda \int d^D x\, \delta(f-k)\sqrt{G}\,\sqrt{M} \end{aligned} \tag{3.2.5}$$

where k is a constant $(D-p)$-vector. S_k gives the Nambu action for surfaces $f = k$. In order for these actions S_k to be well defined constraint (2.1.4) needs to be extended to hold also at $f = k \in \mathcal{R}$. It was also seen that this is required for well defined equations of motion (3.2.3), and that it may be ensured that the additional constraints hold by appropriate gauge transformations. Therefore (2.1.4) is extended to

$$df^1 \wedge \cdots \wedge df^{D-p} \neq 0 \qquad \text{when} \quad \Delta(f) \neq 0 \tag{3.2.6}$$

as part of the gauge fixing constraints.

Similarly the actions (3.1.2 and 6) may be gauge fixed by including the auxiliary field $\boldsymbol{g}$. A gauge choice on invariances (3.1.4 and 5) ensures that equations of motion following from these actions depends only $\delta(f)$ multiplicatively (cf. equation (3.2.1)). Remaining gauge invariances may be used to locally fix terms of order f (as in equation (3.2.4)) to zero. Thus gauge fixed equations of motion follow from $(D \neq p+2)$

$$S = \frac{1}{2}\int d^D x\, \Delta(f)\sqrt{G}\,(\sqrt{g})^{-1}(\partial^\mu f^\top g \partial_\mu f - \Lambda). \tag{3.2.7}$$

In particular equation (3.1.3) is replaced by

$$\Delta(f)(g^{-1} - \frac{D-p-2}{\Lambda}\partial^\mu f \partial_\mu f^\top) = 0.$$

The gauge invariance (3.1.4 and 5) has been used to fix $\boldsymbol{g}$ to be positive definite when $\Delta(f) \neq 0$. Similar arguments apply for equations (3.1.6 and 7).

In fact for $p = 2$ (strings) the analysis may be extended to all background field couplings. By comparison with the Polyakov action a massless antisymmetric tensor field $B_{\mu\nu}(x)$ and dilaton $\Phi(x)$ can be coupled in as in chapter 2, section 4:

$$S = \frac{1}{4\pi\alpha'}\int d^D x\,\sqrt{G}\,\sqrt{M}\,\delta(f)\{2+\varepsilon^{\mu\nu}B_{\mu\nu}+\alpha'\Phi(x)(\nabla_\mu P^{\kappa[\nu}\nabla_\nu P^{\mu]\lambda}P^*_{\kappa\lambda}+P^{\mu\kappa}P^{\nu\lambda}R_{\mu\nu\kappa\lambda})\}. \tag{3.2.8}$$

As noted in chapter 2 the equations of motion that follow from this action again only contain $\delta(f)$ multiplicatively, and if $\delta(f)$ is replaced by $\Delta(f)$, the same equations result with $\delta(f)$ replaced by $\Delta(f)$. No specific properties of $\delta(f)$ are required in their derivation. The analysis below equation (3.2.4) may thus be used again to show that action (3.2.8) with $\delta(f)$ replaced by $\Delta(f)$ yields gauge fixed equations of motion. Presumably this gauge fixing procedure works when all the other gauge invariant couplings in section 4 of the previous chapter are included. The averaging interpretation (3.2.5) extends to these more general actions (including (3.2.7)).

3.3 The Measure

The intention is to use action (3.1.1) (or its extensions as discussed) as the exponential weighting factor in a functional integral representation of the partition function (over f's plus other fields as appropriate). No attempt has yet been made to define the functional measure which is clearly required to be local $GL(D-p,\mathbf{R})$ invariant. However the naïve functional integral, $\int \mathcal{D}f$, is not invariant. For the moment the case where the only quantum fields are f's is considered. If auxiliaries such as g of section 1 are introduced then it is straightforward to include e.g., $\prod_x g(x)$ to a suitable power to ensure invariance of the measure). Some functional of f must be included in the measure in order to generate a functional determinant which cancels that of Ω arising from equation (3.1.1). It is straightforward to show that only $\prod_x \delta(f(x))$ will do this. Intuitively such an incorporation makes sense because only the zeroes of f have physical significance; it is enough in calculating the partition function to integrate over functions whose variations away from zero are arbitrarily small. This cannot however be simply included in the measure since that would make the partition function ill defined. This is so because the action is ill defined for $f \equiv 0$ and constraint (2.1.4) will not be obeyed. What is required is to define the order in which the implied limits are taken which define the $\delta(f)$ in the action and the functional $\prod_x \delta(f(x))$ in the measure. Clearly the limit defining $\prod_x \delta(f(x))$ should be left till last. A convenient limit is that of a gaussian, and so the measure $d\mu[f]$ is taken to be given by

$$d\mu[f] \propto \lim_{\xi\to 0} \mathcal{D}f(\mathrm{Det}\xi)^{-\frac{1}{2}} \exp(-\int d^D x\, f^{\mathsf{T}} \xi^{-1} f). \tag{3.3.1}$$

Here $\xi(x)$ is a positive definite symmetric matrix field, Det is a functional determinant and the limit as $\xi \to 0$ is to be taken at the end of the calculation of the partition function. What is meant for example is a limit as $\epsilon \to 0$ with $\xi = \epsilon\tilde{\xi}$ and ϵ a small positive number. Writing $(\mathrm{Det}\xi)^{-\frac{1}{2}}$ as a gaussian integral with real bosonic ghosts b one has

$$Z = \lim_{\epsilon\to 0} Z[\xi]\big|_{\xi=\epsilon\tilde{\xi}} \tag{3.3.2}$$

$$Z[\xi] = \int \mathcal{D}f\mathcal{D}b e^{-S_{\mathrm{tot}}} \tag{3.3.3}$$

$$S_{\mathrm{tot}} = \int d^D x\, \{\lambda\delta(f)\sqrt{G}\sqrt{M} + f^{\mathsf{T}}\xi^{-1}f + b^{\mathsf{T}}\xi b\}. \tag{3.3.4}$$

The measure in equation (3.3.3) is now manifestly invariant. However the action $S_{\mathrm{tot}}[\xi]$ is no longer gauge invariant, instead it changes to $S_{\mathrm{tot}}[\xi']$. Thus the gauge invariance has been lost (even in the limit $\epsilon \to 0$ since this limit is to be taken after the calculation). ξ may be identified as the gauge fixing parameter (in fact an infinite set with $\frac{1}{2}(D-p)(D-p+1)$ parameters for each point x). This is a valid gauge fixing since by noting that under a gauge transformation $\Omega \in GL(D-p, \mathbb{R})$, f, b and ξ become

$$\begin{aligned} f' &= \Omega f \\ b' &= \Omega^{-\top} b \\ \xi' &= \Omega \xi \Omega^{\top}. \end{aligned} \tag{3.3.5}$$

The gauge invariance is sufficient to transform the field ξ to any other positive definite symmetric field and on substituting equation (3.3.5) into equation (3.3.3) one obtains $Z[\xi] = Z[\xi']$. Thus $Z[\xi]$ is actually independent of ξ and the limit in equation (3.3.2) is superfluous. It is not too hard to understand why $Z[\xi]$ should be independent of ξ. The point is that for each integration $df(x)$ in the functional integral, the $\delta(f)$ term in S_{tot} contributes only at $f(x) = 0$. So the result is the gaussian integral over f yielding a factor of $(\det \xi(x))^{\frac{1}{2}}$ times this local contribution. But the gaussian integral over $b(x)$ cancels the ξ dependence. Of course the properties above may be subject to anomalies—causing $Z[\xi]$ to depend on ξ. This is equivalent to a loss of local $GL(D-p, \mathbb{R})$ invariance at the quantum level.

It has been stated that the measure is required to be local $GL(D-p, \mathbb{R})$ invariant but it is also required that it can be defined in a way that leads to a unitary and renormalisable quantum field theory. Providing some working rules for a sufficient definition of the measure is not necessarily a straightforward application of the rules of renormalisable quantum field theory. This is because the functions f (and auxiliaries and ghosts) are not fields in the usual sense. (For example they do not asymptotically represent first quantised wavefunctions). A proper proof of these properties in this case is surely at least as difficult and subtle as it is in general for interacting quantum field theory. Some heuristic arguments for the case of strings ($p = 2$) will be presented which nevertheless indicate how to incorporate the above requirements for $p = 2$. (From first quantised string theory these properties are expected in general to be violated by anomalies).

A $GL(D-2, \mathbb{R})$ invariant measure may be reduced to one that integrates over equivalence classes of functions f identified up to multiplication by any element of local $GL(D-2, \mathbb{R})$. These classes are uniquely labelled by the worldsheets obtained from the functions' zeroes. Thus the measure depends only on worldsheet embeddings. It is generally covariant with respect to the surface (because no parameterisation of the worldsheet has been introduced). This measure may be thought of as being the Polyakov measure but with any number of generally covariant counterterms depending on the worldsheet metric and the string position. However these counterterms, providing they are local (meaning also that only polynomials of derivatives appear and for example no inverses of derivated fields), are vertex operators for various backgrounds (including higher mass modes). The Polyakov measure with all such backgrounds included is unitary and renormalisable because divergences appear as generally covariant local functionals of the worldsheet metric and position and may thus be incorporated into wavefunction renormalisation of the backgrounds. In order for our $GL(D-2,\mathbb{R})$ invariant measure to be equivalent to this measure, it is required that the divergences, which will be $GL(D-2,\mathbb{R})$ invariant local expressions in the quantum fields, be absorbable into renormalisations of local $GL(D-2,\mathbb{R})$ invariant couplings of background fields, and that these couplings correspond to the couplings for backgrounds in the Polyakov theory. In chapter 2 it was demonstrated that local invariant expressions in

$$\varepsilon^{\mu\nu} = (GM)^{-\frac{1}{2}}\varepsilon^{\mu\nu\mu_1\cdots\mu_{D-2}}\partial_{\mu_1} f^1 \cdots \partial_{\mu_{D-2}} f^{D-2}, \tag{3.3.6}$$

$P^{\mu\nu} = \varepsilon^{\mu}{}_{\sigma}\varepsilon^{\nu\sigma}$ and f are reduced to (all) generally covariant local expressions of the worldsheet metric and position. Thus the above arguments suggest that the local $GL(D-2,\mathbb{R})$ invariant measure may be assumed to be renormalisable and unitary if f *and* the antisymmetric tensor $\varepsilon^{\mu\nu}$ are taken as quantum fields, and all independent local $GL(D-2,\mathbb{R})$ invariant expressions involving these fields are coupled to background fields (as in equation (3.2.8) plus similar higher mode couplings). Some local terms depending on auxiliary fields whose equations of motion identify $\varepsilon^{\mu\nu}$ with its expression in terms of f must be added to the action:

$$S_{\mathrm{aux}} = \int d^D x\, \delta(f)\sqrt{G}\,\sqrt{M}\,\{Y_\mu^{\top}\varepsilon^{\mu\nu}\partial_\nu f + X(\varepsilon^{\mu\nu}\varepsilon_{\mu\nu} - 2)\} \tag{3.3.7}$$

where $Y_\mu(x)$ and $X(x)$ are auxiliary fields. X is invariant under gauge transformations whereas Y_μ transforms contragrediently like b: $Y_\mu \to \Omega^{-\top}Y_\mu$. At $f=0$ Y_μ forces $\varepsilon^{\mu\nu}$ to be orthogonal to the $\partial_\nu f^i$ which determines it to be proportional to the alternating tensor in the remaining two dimensions. X fixes the proportionality factor (at $f=0$). The result is (3.3.6) (up to a sign) at $f=0$. Thus (3.3.7) identifies $\varepsilon^{\mu\nu}$ with (3.3.6) up to terms of order f. This is consistent with the gauge invariance

$$\varepsilon^{\mu\nu} \to \varepsilon^{\mu\nu} + f^{\top} e^{\mu\nu} \tag{3.3.8}$$

noted in section 2.4.

Now consider the partition function

$$Z = \int d\mu[f]\mathcal{D}\varepsilon_{\mu\nu}\, e^{-S} \tag{3.3.9}$$

where S is the action including all the massless and massive backgrounds plus the extra terms (3.3.7). The functional integral over $\varepsilon_{\mu\nu}$ is trivially invariant under its gauge invariances which are just equation (3.3.8). Thus equation (3.3.9) is a gauge invariant partition function for both of the quantum fields f and $\varepsilon_{\mu\nu}$.

There now follows an indication of how, for large constant dilaton expectation value Φ_0, (3.3.9) reduces to the first quantised perturbation expansion with all the correct combinatorics. By factoring out the $GL(D-2, \mathbb{R})$ gauge invariance the functional integral becomes an integral over equivalence classes labelled uniquely by the configuration of the worldsheets; the action becomes a sum of first quantised actions, one for each worldsheet. Thus the functional integral splits into a sum of products of first quantised partition functions, one for each worldsheet. (Overlapping worldsheets are forbidden by equation (2.1.4). This would stop the integral factorising completely into products of first quantised partition functions, but in $D > 4$ dimensions the overlapping contributions can be expected to be of zero measure). In addition there is an initial term in the sum corresponding to an integral over functions with no simultaneous zeroes (no worldsheets). For this case the action vanishes (because of the $\delta(f)$) and the integral just gives unity (because all such functions define just one equivalence class). Thus Z reduces to

$$Z = 1 + \sum_{\Xi} \exp(\Phi_0 \sum_{\chi\in\Xi} n_\chi \chi) \prod_{\chi\in\Xi} \frac{1}{n_\chi!}\{Z(\chi)\}^{n_\chi}. \tag{3.3.10}$$

Here the sum runs over all disconnected topologies $\Xi = \{(\chi, A)\}$, each connected topology uniquely specified by its Euler number χ. Extra labels $A = 1, 2, \cdots, n_\chi$ distinguish, for the purposes of this first quantised expansion, disconnected pieces with the same connected topology. $g = e^{-\Phi_0}$ is the string coupling constant, the expansion parameter. $Z(\chi)$ is a first quantised integral for a string worldsheet with Euler number χ. The factorial term that appears in (3.3.10) is a symmetry factor arising from the indistinguishability of worldsheets with the same topology. The symmetry factor is required because, for n_χ worldsheets with the same topology χ, the product of first quantised integrals $\{Z(\chi)\}^{n_\chi}$ counts $n_\chi!$ times the same configuration of worldsheets i.e., the same equivalence class. Equation (3.3.10) is exactly the form expected (from quantum field theory) for the vacuum diagram perturbation series of the second quantised partition function. Indeed (3.3.10) may be rearranged so that, in analogy with quantum field theory,

$$Z = e^W$$

where

$$W = \sum_{\chi} \exp\{\chi \Phi_0\} Z(\chi)$$

i.e., W is the generator of connected string Greens functions (obtained by differentiating with respect to background fields). Z does not have to expanded in terms of first quantised integrals however. Indeed this is the motivation for the present formalism: The string theory is required to be well-defined without reference to any expansion but such that, when expanded for small g, it yields just the familiar first quantised contributions. Thus the formalism will also give further non-perturbative contributions, which are not calculable from string worldsheets, but which are necessary for the theory to be well defined.[3] (However, it is conceivable that there exists more than one non-perturbative string theory with the same perturbative contributions). To demonstrate that the theory is well defined without expanding for small g, it is necessary to show that one may completely gauge fix the partition function in some other way than the perturbative factorisation which leads to equation (3.3.10).

It has been shown that Z is independent of gauge fixing parameters (up to anomalies) and that it yields the required first quantised integrals on expansion. Thus the existence of an appropriate gauge fixed measure has been demonstrated. A gauge invariant partition function may be trivially obtained from this by integrating $Z[\xi]$ over ξ with an invariant measure $\mathcal{D}\xi(\mathrm{Det}\xi)^{\frac{p-D-1}{2}}$. This integral is infinite and proportional to the product of $GL(D-p, \mathbb{R})$ group volumes for each point x.

3.4 Hamiltonian Quantisation

Finally, a hamiltonian framework for the quantisation of the equations of motion following from action (3.1.1) will briefly be considered. For simplicity's sake flat euclidean space $G_{\mu\nu} = \delta_{\mu\nu}$ will be used. Let Greek indices α, β represent spacial indices $(\alpha, \beta = 1, \ldots, D-1)$ and indices $\mu, \nu = 0, \ldots, D-1$ represent space-time. First note that

$$
\begin{aligned}
M &= \frac{1}{n!}\varepsilon_{i_1\ldots i_n}\varepsilon_{j_1\ldots j_n} f^{i_1}_{\nu_1}\ldots f^{i_n}_{\nu_n} f^{j_1}_{\nu_1}\ldots f^{j_n}_{\nu_n} \\
&= \dot{f}^{i_1}\dot{f}^{j_1}\{\frac{1}{(n-1)!}\varepsilon_{i_1\ldots i_n}\varepsilon_{j_1\ldots j_n} f^{i_2}_{\alpha_2}\ldots f^{i_n}_{\alpha_n} f^{j_2}_{\alpha_2}\ldots f^{j_n}_{\alpha_n}\} \\
&\quad + \frac{1}{n!}\varepsilon_{i_1\ldots i_n}\varepsilon_{j_1\ldots j_n} f^{i_1}_{\alpha_1}\ldots f^{i_n}_{\alpha_n} f^{j_1}_{\alpha_1}\ldots f^{j_n}_{\alpha_n} \\
&= \dot{f}^{\top}\boldsymbol{J}\dot{f} + K
\end{aligned}
\tag{3.4.1}
$$

where $n = D - p$ and the dots above f denote differentiation with respect to time. Introducing $\boldsymbol{K} = f_\alpha f_\alpha^{\top}$ one has $K = \det(\boldsymbol{K})$ and $\boldsymbol{J}$ is the matrix of cofactors, so $\boldsymbol{J} = K\boldsymbol{K}^{-1}$ when $\boldsymbol{K}$ is invertible.

The lagrangian density $\mathcal{L} = \lambda\delta(f)\sqrt{M}$ and so from equation (3.4.1) one finds the canonical momentum

$$\pi = \lambda\delta(f)(\sqrt{M})^{-1}\boldsymbol{J}\dot{f}. \tag{3.4.2}$$

Writing

$$\pi^{\top}\boldsymbol{J}^{-1}\pi = M^{-1}\lambda^2\delta^2(f)\dot{f}^{\top}\boldsymbol{J}\dot{f}$$

implies that

$$M(\lambda^2\delta^2(f)K - \pi^{\top}\boldsymbol{K}\pi) = K^2\lambda^2\delta^2(f).$$

Hence the hamiltonian density $\mathcal{H} = \pi^{\top}\dot{f} - \mathcal{L}$ becomes

$$
\begin{aligned}
\mathcal{H} &= -\lambda\delta(f)\sqrt{M}^{-1}K \\
&= -\sqrt{\lambda^2\delta^2(f)K - \pi^{\top}\boldsymbol{K}\pi}.
\end{aligned}
\tag{3.4.3}
$$

Following the methods of Dirac it is noted that equation (3.4.2) implies the primary constraints:

$$\pi(x^\alpha)f^{\top}(x^\alpha) \approx 0. \tag{3.4.4}$$

In agreement with general principles, these constraints applied as contact transformations generate the infinitesimal $GL(D-p, \mathbb{R})$ gauge transformations, and form a representation of the local $GL(D-p, \mathbb{R})$ algebra. Specialising to the case $D = 2$, $p = 1$

(corresponding to a free particle in two dimensions) affords considerable simplification over the general case which will be commented on later. It turns out that in this simple case the hamiltonian density may be written

$$\mathcal{H} = -\sqrt{m^2 F^2(f) - \pi^2}\, f' \tag{3.4.5}$$

where $\lambda = m$, the mass of the particle, and prime denotes differentiation with respect to position. $\delta(f)$ has been changed to a general function of f, $F(f)$. It will be shown that this has no effect on the equations of motion so that a gauge invariant system of equations can still be described. The hamiltonian density (3.4.5) now appears well defined however the weak equation (3.4.4) is still required. There are no secondary constraints since

$$\begin{aligned}\frac{d}{dt}(\pi f) &= \{\pi f, H\}\\ &= -\frac{1}{\mathcal{H}}(\pi^2 f'^2 + f f'^2 m^2 F F') - \int dy \frac{\mathcal{H}}{f'} \partial\delta(x-y) f(x)\\ &= \frac{\pi}{\sqrt{(m^2F^2 - \pi^2)}}(\pi f)'\end{aligned} \tag{3.4.6}$$

where the derivative has been integrated off the delta function in the second line. This vanishes by equation (3.4.4). $\{,\}$ are Poisson brackets such that $\{f(x), \pi(y)\} = \delta(x-y)$. Thus the system only has these first class constraints. The full hamiltonian density is then

$$\mathcal{H} = -\sqrt{m^2F^2 - \pi^2}\, f' + \omega\pi f \tag{3.4.7}$$

where the constraints multiplied by an arbitrary field $\omega(x)$ have been added. Using this hamiltonian one finds:

$$\frac{\partial H}{\partial \pi} = \dot{f} \approx \frac{\pi}{\sqrt{(m^2F^2 - \pi^2)}} f' + \omega f \tag{3.4.8}$$

$$-\frac{\partial H}{\partial f} = \dot{\pi} \approx \frac{\pi\pi'}{\sqrt{(m^2F^2 - \pi^2)}} - \omega\pi. \tag{3.4.9}$$

Computing $\ddot{f} \approx \{\dot{f}, H\}$ (in order to find the classical equations of motion) produces the following:

$$\begin{aligned}\ddot{f} &= \frac{\partial \dot{f}}{\partial f}\dot{f} + \frac{\partial \dot{f}}{\partial \pi}\dot{\pi}\\ &= -\frac{\pi m^2 F F' f' \dot{f}}{\sqrt{(m^2F^2 - \pi^2)^3}} + \omega \dot{f} + \frac{f' m^2 F^2 \dot{\pi}}{\sqrt{(m^2F^2 - \pi^2)^3}} + \frac{\pi \dot{f}'}{\sqrt{(m^2F^2 - \pi^2)}}\end{aligned} \tag{3.4.10}$$

where the last term arises from integrating $\partial f'/\partial f$ by parts. Using equation (3.4.8) to define π classically and hence to find π' enables the F' term to be eliminated from (3.4.10) giving (after some rearrangement)

$$\ddot{f} = 2\frac{\dot{f}'\dot{f}}{f'} - \frac{f''\dot{f}^2}{f'^2} - \left(\frac{\dot{f}}{f'}\omega\right)' f - \frac{\omega m^2 F^2}{\sqrt{(m^2F^2 - \pi^2)}}(\pi f)'.$$

The last term vanishes by constraint (3.4.4) leaving (after multiplying through by f'^2)

$$f'^2\ddot{f} - 2\dot{f}f'\dot{f}' + f''\dot{f}^2 = -f'^2\left(\frac{\dot{f}}{f'}\omega\right)' f$$

which are the gauge invariant equations of motion (cf. equation (3.2.4)). Note that $F(f)$ does not appear in the equations of motion, nor does it affect the equations of motion in any way.

Using equation (3.4.8) to solve for π (classically) one may show that the constraint $\pi f \approx 0$ implies that $F(f) \propto \delta(f)$. However the weak constraint is only applied at the end of the calculation by which time, at least for the equations of motion, it is unnecessary. Also, once quantised, one cannot solve for π. Instead the constraints (3.4.4) vanish on physical states:

$$\pi f|\mathrm{phys}\rangle = 0 \qquad \text{and} \qquad \langle\mathrm{phys}|\pi f = 0. \tag{3.4.11}$$

Nevertheless the full gauge invariant quantised equations of motion are obtained for any function F. A particularly convenient choice would appear to be $F(f) = 1$. In this case, at least naïvely, there are no operator ordering problems. For example one may define the ordering to be such that all π terms occur to the left of all f terms: The hamiltonian density is as given in equation (3.4.5) with $F(f) = 1$. The primary constraints (3.4.4) remain first class at the quantum level and are still the only constraints because the ordering of operators in $(\dot{\pi f}) = -i[\pi f, H]$ is that of equation (3.4.6). Note that an ordering has been chosen which causes $\mathcal{H}$ and the constraints not to be hermitian. Hermitian definitions can be chosen by for example taking the symmetric combination of half the sum of two orderings, one with all π's to the left and the other with all π's to the right. However one then runs into operator ordering problems; for example $d(\pi f + f\pi)/dt$ is not weakly zero without unwanted further constraints. The constraints equation (3.4.11) may be solved to obtain the form of the physical states. Note that the constraints, being non-hermitian, have different solutions for bras and kets. Using an $f(x)$ representation $\pi(x) = -i\delta/\delta f(x)$, and imposing the requirement that the states be integrable (so that a scalar product using $\int \mathcal{D}f(x)$ exists) one finds that the physical states are determined uniquely as $\langle\text{phys}| = 1$ and $|\text{phys}\rangle = \prod_x \delta(f(x))$, up to multiplication by a constant. This is not surprising since equation (3.4.11) is equivalent to ensuring that expectation values be local $GL(D-p, \mathbb{R})$ invariant and, as has already been seen in discussing the measure in the lagrangian formulation, (naïvely) this requires an insertion of $\prod_x \delta(f(x))$ in the measure. The $\prod_x \delta(f(x))$ must again however be treated with care, otherwise with $F(f)$ continuous (at $f=0$) the theory is presumably trivial, and with $F(f) = \delta(f)$ it is not well defined. Presumably what is required is to find a hamiltonian analogue of the arguments given defining the measure in the functional integral approach.

Generalising to general space-time dimension D and general p and replacing $\delta(f)$ by $F(f)$ in equation (3.4.3) one finds

$$\dot{f} = \{f, H\} = -\frac{\boldsymbol{K}\,\pi}{\mathcal{H}}$$

and

$$\dot{\pi} = \{\pi, H\} = \lambda^2 F(f)\Big(\frac{F(f)K\boldsymbol{K}^{-1}f_\alpha}{\mathcal{H}}\Big)_\alpha - \Big(\frac{\pi^\top f_\alpha \pi}{\mathcal{H}}\Big)_\alpha .$$

Not all the differentials have been expanded for the sake of compactness. If one expands all the differentials one finds that, unlike equation (3.4.9), terms depending explicitly on $F(f)$ and dF/df remain. From the above equations one finds

$$\begin{aligned}\frac{d}{dt}(\pi f^\top) &= \{\pi f^\top, H\}\\ &\approx \lambda^2\Big(\frac{F(f)K\boldsymbol{K}^{-1}f_\alpha}{\mathcal{H}}\Big)_\alpha F(f)f^\top .\end{aligned} \tag{3.4.12}$$

So consistency with equation (3.4.4) now also requires $F(f)f \approx 0$. Now both classically and quantum mechanically this new constraint determines $F(f) \propto \delta(f)$ at the end of the calculation. However if this is substituted into the right hand side of the above (or similar) equations they become ill-defined. The need for the extra constraint $F(f)f \approx 0$ stems from the fact that dF/df appears explicitly in the expanded expression for $\dot{\pi}$. (In the special case $D-p=1$ (general D) the explicit dF/df terms cancel but consistency in this case requires $\mathcal{H}f \approx 0$ which has the same effect as $F(f)f \approx 0$).

Thus in this gauge invariant hamiltonian framework in the general case it appears that one has to deal directly with expressions involving products and ratios of $\delta(f)$ and $\delta'(f)$. Some further structure is necessary to ensure that these expressions have limits that are unambiguous. At the quantum level one must also solve problems with operator ordering. This is necessary to ensure that one retains a reasonable set of first class constraints.

One strategy for solving both of the above problems could be to consider gauge fixing the classical system (and then quantising). For example gauge fixing $\boldsymbol{M} = \boldsymbol{1}$ at $f = 0$ (cf. lagrangian system) may prove convenient. This will neatly incorporate the constraint $M > 0$ when $f = 0$, which is required to ensure that the hamiltonian expression is well defined. (In the present variables $M > 0$ if and only if the following holds:

$$\boldsymbol{J} \neq 0 \qquad \text{and if} \quad K = 0 \quad \text{then} \quad \boldsymbol{K}\pi = 0).$$

Gauge fixing $\boldsymbol{M} = \boldsymbol{1}$ is equivalent to requiring

$$\pi^{\top}\boldsymbol{K}\pi = (1 - K)K\lambda^2\delta^2(f)$$

which simplifies the hamiltonian density so that it is well defined and without operator ordering problems:

$$\mathcal{H} = -\lambda\delta(f)K.$$

This must be supplemented with ghost and constraint terms.

3.5 Summary of Chapter 3

In this chapter some progress has been made towards a tractable gauge fixing and quantisation of the model second quantised string theory outlined in chapter 2. The starting point was the action

$$S = \frac{1}{2\pi\alpha'} \int d^D x \sqrt{G}\, \delta(f) |df^1 \wedge \cdots \wedge df^{D-2}|$$

and a local $GL(D-p, \mathbb{R})$ gauge invariance. In the introductory section to this chapter the immediate problems that the quantisation of this system poses were listed. It turns out that the multi-derivative term in the action does not appear of itself to pose a problem because it can simply gauged away, leaving an action with at most two derivatives. Alternatively the full gauge invariance can be kept and auxiliary fields introduced which transform the action into one with at most two derivatives as explained in section 1.

The $\delta(f)$ is much more difficult to deal with and it is clear that standard methods of perturbation theory cannot be applied whilst the $\delta(f)$ remains in the action. On the other hand it is not clear how to make progress without perturbation theory. Connected to this is the problem of finding a complete gauge fixing of the local $GL(D-p, \mathbb{R})$ invariance. In order to follow standard procedures a constraint on the functions f (or auxiliaries) must be found that completely fixes the gauge invariance. It would appear that there is no *tractable* conventional gauge fixing procedure (i.e., proceeding from a constraint on the functions f (or auxiliaries) which completely fixes the gauge invariance). For the above reasons an unconventional gauge fixing was formulated in section 2. This procedure is particularly suited to this system because it was possible to replace the troublesome $\delta(f)$ with a general function $\Delta(f)$ which may be amenable to perturbation theory. It was shown that such an action (with or without auxiliary fields or background couplings) gave gauge fixed equations of motion. In order to complete this programme it will be necessary to find the ghost action and the corresponding BRST invariance.

Section 3 concentrated on the functional measure. At first sight there is a problem with quantisation because the measure $\mathcal{D}f = \prod_x df(x)$ is not even naïvely invariant under gauge transformations. Inserting a functional delta function in the measure cures this problem and is also physically reasonable but now the partition function is ill defined. However it was shown that, by introducing an infinite set of gauge fixing parameters of the form of a field $\xi(x)$, one can construct a well defined measure and a completely gauge fixed action. The partition function is formally independent of the gauge fixing parameters ξ and yields the required first quantised integrals upon expansion. A breakdown of this independence of ξ is equivalent to an anomaly in the gauge invariance. In the latter half of that section some arguments were provided suggesting that the string theory ($p = 2$) was renormalisable and unitary for *any* gauge invariant measure if the gauge invariance suffers no anomalies, the quantum fields are f and $\varepsilon^{\mu\nu}$, ($\varepsilon^{\mu\nu}$ identified with its expression in terms of f through equations of motion), and provided all gauge invariant background field couplings are included.

Finally in section 4 a hamiltonian formulation for the quantisation of the model was investigated. The fact that the conjugate momentum is proportional to $\delta(f)$ leads to a simple set of primary constraints which generate the local $GL(D-p,\mathbb{R})$ gauge invariance. In the special case of a particle in two dimensions ($p = 1$, $D = 2$) there are no secondary constraints and this remains true even if $\delta(f)$ is replaced by a general function $F(f)$. Surprisingly this replacement has no effect on the gauge invariant equations of motion since the expressions are all well defined with a general function $F(f)$. Choosing F to be identically unity allows one to give a simple prescription for operator ordering in the quantum theory which ensures that the constraints remain first class. The constraints appear to enforce a unique solution on the physical states which includes a functional δ-function on the functions f. This is the hamiltonian analogue of the analysis of the functional measure in section 3. The hamiltonian formalism in the case of general p would appear to require one to deal directly with expressions involving products and ratios of $\delta(f)$ and $\delta'(f)$. This is necessary at least in terms of imposing the constraint $F(f)f \approx 0$ and further structure is necessary if such expressions are to be unambiguous. At the quantum level the case of full gauge invariance and general p also has operator ordering problems. It was suggested that some partial gauge fixing might solve these problems.

In summary a number of methods have been developed to tackle various aspects of the implementation of our proposed second quantised string theory. All of the methods appear promising however some of them are not yet complete. It is not believed that there are any fundamental problems in performing the quantisation of this theory. In particular it should be possible to prove the equivalence of the theory to free particles when $p = 1$ and to string perturbation theory for $p = 2$ and large positive dilaton expectation value.

References

[1] I.A. Batalin and G.A. Vilkovisky, *Phys. Lett.* **102B** (1981) 27.

[2] T. Kugo and S. Uehara, *Nucl. Phys.* **B197** (1952) 378.

[3] D.J. Gross and V. Periwal, *Phys. Rev. Lett.* **60** (1988) 2105.

CHAPTER 4

Speculation and Summary

Although quantisation of the formalism is incomplete, there are indications that the implicit function method may be useful for describing some interesting ideas at the classical level. In section 1 it is shown how the method may yield a topological field theory, where the metric can be set to zero, when the dilaton expectation value becomes large.

Section 2 reviews the work done in chapters 1 to 3 and attempts to draw some conclusions. Some of the more pressing areas where further work is required are noted.

4.1 Speculation

It is tempting to speculate on the possible consequences of being able to perform detailed calculations in the implicit function formalism. In particular could any simplification result in the high energy regime of string theory? A number of authors have suggested that at high enough energies a new phase of string theory may take over with far fewer degrees of freedom than field theory.[1−3] It is worth noting that, at least formally, our model already has less degrees of freedom than are required to describe free particles. (The latter requires $D-1$ functions). Witten has speculated on the existence of a phase where in some sense general covariance is unbroken and the metric vanishes,[3] leaving a topological field theory. It can be shown that these ideas fit neatly within our framework. Within the renormalisation group approach our various background field couplings can be expected to depend on some energy scale μ. In particular the very soft behaviour of string scattering amplitudes at high energy implies that the string coupling constant $g = \exp\{-\Phi_0\}$ renormalises to zero as the energy increases indefinitely, i.e., the constant part of the dilaton $\Phi_0 \to \infty$. In such a regime, from the point of view of action (3.3.8), it seems natural to perturb about an action consisting purely of the Φ_0 term:

$$S_{\Phi_0} = \frac{1}{4\pi}\int d^D x\,\sqrt{G}\,\sqrt{M}\,\delta(f)\Phi_0\{\nabla_\mu P^{\kappa[\nu}\nabla_\nu P^{\mu]\lambda}P^*_{\kappa\lambda} + P^{\mu\kappa}P^{\nu\lambda}R_{\mu\nu\kappa\lambda}\} \qquad (4.1.1)$$

(plus auxiliary terms (3.3.7)). Such an action yields a topological field theory,[3] i.e., a theory with no local degrees of freedom. The action is independent of small variations in f (see section 2.3). Thus the topological theory requires the far larger gauge invariance $\delta f =$ anything (subject to constraint (2.1.4)) which identifies all string configurations with the same disconnected topology (in the first quantised expansion) providing that the configurations can be deformed into one another on the space-time manifold. The partition function can therefore be expected to be sensitive to the connectivity of the manifold and one can expect more subtle smooth topological invariants to also play a rôle.[3] The above action (4.1.1) can be obtained from the original by setting all other background fields *including* the space-time metric to zero. This seems a natural description for a fundamental string phase since all possible symmetries are restored in the process. In setting the background space-time metric to zero note that it is unnecessary in the part of the action containing the constant part of the dilaton. This is because that part is independent of variations in the metric as shown in section 2.3 (so if one likes, one can unambiguously take the limit $G_{\mu\nu} \to 0$). Note also that the remaining metric term (the first term in equation (3.3.4), or equation (3.1.1)) can be rewritten:[4]

$$\frac{1}{2\sqrt{2}\pi\alpha'}\int d^D x\,\delta(f)\sqrt{\varepsilon^{\mu_1\ldots\mu_D}\,\varepsilon^{\nu_1\ldots\nu_D}\,G_{\mu_1\nu_1}G_{\mu_2\nu_2}f^1_{\mu_3}\cdots f^{D-2}_{\mu_D}\,f^1_{\nu_3}\cdots f^{D-2}_{\nu_D}}\,.$$

Thus setting $G_{\mu\nu}$ to zero is unambiguous. This same phase also results from taking the limit entertained by Gross and Mende.[5] They took a background consisting only of flat space-time with large Φ_0 (small string coupling constant), and considered the limit $\alpha' \to \infty$ (which eventually kills the above term, leaving (4.1.1)). They showed that for all genus and all states the string scattering amplitudes vanish in this limit while simultaneously gaining new perhaps infinite dimensional symmetries.[1] This may be a consequence of entering this topological phase, with the symmetries arising through the enlarged gauge invariance. If this is the fundamental high energy phase of string theory then the lack of any local degrees of freedom would ensure and explain the finiteness properties. How might one recover the usual world with propagating degrees of freedom? The form of the action (4.1.1) suggests that quantum corrections might well contain the Einstein action (in the background field effective action). This would however signal a break-down of topological phase symmetries (e.g., $\delta G_{\mu\nu}$ = anything). It would also introduce a scale parameter (none being present in (4.1.1)) which can be identified with α'. Thus dynamical gravity and its coupling would arise spontaneously in the quantum corrections. Technically it should be possible to derive the finiteness properties and relations between different string amplitudes[1] by studying the broken topological phase Ward identities.

4.2 Summary and Conclusions

This final section reviews and draws some conclusions from the work presented in previous chapters. The work describes a formulation of second quantised string theory which attempts to overcome some of the drawbacks inherent in the perturbative framework. It was also hoped that the theory might incorporate some of the philosophical ideals that a fundamental theory of nature may be expected to encompass. Phenomenological considerations of perturbative string theory indicated that it was strongly coupled and this was reinforced by evidence that its perturbation series diverged and was not even Borel summable. However, considering that at present string theory seems to be the only theory that might yield a consistent theory of quantum gravity, it is worth trying to extend it into the non-perturbative regime. String theory also has the philosophically encouraging property of predicting the dimension of space-time. Other interesting and important evidence that a non-perturbative, second quantised string theory was required came from the implication that it must be consistent on all worldsheets simultaneously. This last point can probably be most easily addressed by having a sort of 'worldsheet field theory' and indeed in some sense the candidate theory introduced in chapter 2 can be considered as such. Thus the main criticisms of perturbative string theory are that it is so defined. It is not known for sure whether perturbative string theory encompasses the philosophical criteria mentioned in chapter 1, such as requiring that it generate the geometry and topology of space-time, however it does seem intuitively unlikely that the natural description of such effects will be in terms of the scattering of an infinite number of strings.

Some of the alternatives to perturbative string theory were briefly discussed in section 1.3 and it seems that in general they are probably not well enough understood to provide a constructive base from which to launch an investigation into non-perturbative string theory. The most popular approach, string field theory, has however enjoyed some success in areas concerning the independence of the theory on background geometry and topology. There are several reasons why a field theoretic description of string theory is inappropriate. The overcounting of vacuum graphs in field theory suggests that modular invariance and duality ought in fact to be incorporated at a fundamental level. Other reasons for searching for something other than a field theory are that string theory has too few degrees of freedom and that the high energy behaviour of its scattering amplitudes is too soft to be described by a field theory.

It would be of great æsthetic appeal to be able to derive a fundamental theory describing all of nature from a few general physical or philosophical principles. Our approach however was to make modest extensions of previous work and perhaps eventually identify some deep underlying principle. In chapter 2 an approach to second quantised string theory was presented which describes worldsheets implicitly as the solution surfaces of $D-2$ functions on space-time. This allows any number of surfaces of any topology to be described thus yielding a fully interacting second quantised string theory. It therefore appears to be some way towards being consistent on all worldsheets simultaneously. Indeed the treatment generalises to surfaces of arbitrary dimension thus yielding second quantised membrane theories. Perhaps some consistency condition will naturally select strings. Also the theory is not defined by a perturbation expansion. The worldsheets are not parameterised and so modular invariance and duality should be inherent in the formalism. Strings are not mentioned explicitly and so this formalism may be a step away from the notion of scattering infinite numbers of strings, which is intuitively unpleasant. The worldsheet description may also facilitate discussions of physics above the Hagedorn temperature where worldsheets are thought to break down. It has in fact been suggested that the formalism may yield a topological field theory, for large dilaton expectation value, and that it may even generate the dimension, topology and geometry of space-time dynamically.[6]

It was also shown in chapter 2 that the proposed action (2.1.1) for the candidate second quantised string theory reduced to a sum of Nambu actions (for $p = 2$) and was hence not entirely inconsistent with previous work. Further it was shown that the action had a local $GL(D-2, \mathbb{R})$ gauge invariance and that the massless fields the metric, the antisymmetric tensor and the dilaton could be coupled in a gauge invariant manner. $GL(D-2, \mathbb{R})$ invariant expressions corresponding to those at the first mass level in the Polyakov approach were written down in the formalism. A prescription was given for writing down the couplings of higher mass modes however it is not in fact clear whether these couplings should be $GL(D-2, \mathbb{R})$ invariant. As a check on some of the methods used to derive these couplings, the constant dilaton action was explicitly verified to be a topological invariant.

Gauge fixing and quantisation of the model were attempted in chapter 3 however progress was hampered by the delta function in the action. This was removed using a rather unconventional gauge fixing which yielded gauge fixed versions of the equations of motion. Unfortunately it is rather unclear how to introduce ghosts into such a gauge fixed action. Ghosts can easily be introduced when employing a more conventional gauge fixing however one is left with the problem of an incomplete gauge fixing. A gauge invariant measure was introduced and it was shown how the theory yielded the correct first quantised integrals on expansion. A hamiltonian quantisation was attempted in section 3.4 which yielded the correct gauge invariant equations of motion. An ordering of the conjugate variables could be defined when the delta function was replaced by unity however the case of general p and general D however led to expressions involving ratios of delta functions and derivatives of delta functions and to ordering problems. Thus although quantisation looks feasable it is as yet incomplete.

The implicit function approach to second quantised string theory appears to incorporate some fundamental principles yet still makes contact with established ideas. Although the model in theory has a lot of scope, some method of dealing with the delta function in the action must be found before there can be any hope of performing detailed calculations. Real tests of the formalism, such as reproducing the standard result $D = 26$, await a tractable quantisation. At the quantum level, consistency of the theory on all worldsheets must be verified, the measure must be shown to be free from gauge anomalies and the model must be shown to reproduce the correct perturbation theory for strings (at large Φ_0). A pressing problem is to find a supersymmetric generalisation so that fermions can be introduced. Although there is evidence that supersymmetry is no longer required to explain the vanishing of the cosmological constant,[7] it is required to cancel the tachyon and provide fermions.

A lot of work is required before it can be determined whether or not the implicit function approach is an important step towards a non-perturbative string theory. Most urgently required are a complete quantisation and supersymmetrisation of the model. None the less it is hoped that the reader has been convinced that the formalism is simple and flexible enough to encompass the exciting phenomena that lie beyond in non-perturbative string theory.

References

[1] D.J. Gross, *Phys. Rev. Lett.* **60** (1988) 1229

[2] E. Witten, *Phys. Lett.* **206B** (1988) 601 and *Commun. Math. Phys.* **117** (1988) 353.

[3] I. Klebanov and L. Susskind, *Nucl. Phys.* **B309** (1988) 175;
M. Karliner, I. Klebanov and L. Susskind, *Int. J. Mod. Phys.* **A3** (1988) 1981;
J.J. Atick and E. Witten, *Nucl. Phys.* **B310** (1988) 291.

[4] T.R. Morris, *Phys. Lett.* **202B** (1988) 222.

[5] D.J. Gross and P.Mende, *Phys. Lett.* **B197**, (1987) 129, *Nucl. Phys.* **B303** (1988) 407.

[6] D.L. Gee and T.R. Morris, "A Topological Second Quantized String Theory," Southampton preprint SHEP 88/89-12 (1989).

[7] G. Moore, *Nucl. Phys* **B293** (1987) 139 and Erratum **B299** (1988) 847.