String Theory based on Polchinski

Personal Study Notes
chandra.pp@alumni.iitg.ac.in

April 21, 2026

Preface

These notes are personal study materials based primarily on Joseph Polchinski's *String Theory*. While they derive insights from multiple sources, Polchinski's text served as the primary roadmap and foundational framework for this project.

The central goal of these notes is to make the subject of string theory more accessible, both for my own mastery and for any other students navigating this demanding field. To achieve this, I have made every effort to **derive all calculations explicitly**. By filling in the "missing steps" and technical nuances often omitted in standard texts, I hope to ensure the material is easily parsable and the underlying mathematical narrative is transparent.

I welcome feedback, corrections, or discussions regarding the content of these notes. Please feel free to reach out to me via email.

Contact:
chandra.pp@alumni.iitg.ac.in

Disclaimer: These are pedagogical study notes. While I have strived for accuracy, any errors in the calculations or physical interpretations remain entirely my own.

Contents

Chapter 1

1 A first look at strings

1.1 Why strings?

String theory is an ambitious attempt to provide a unified "Theory of Everything," incorporating all fundamental forces of nature—gravity, electromagnetism, and the nuclear forces—into a single quantum mechanical framework. The primary motivation for this shift from standard physics lies in the inherent limitations of Quantum Field Theory (QFT) when applied to gravity.

In QFT, elementary particles are treated as zero-dimensional points. This assumption leads to severe ultraviolet (UV) divergences because interactions occur at a single point in spacetime, resulting in mathematical singularities. While theories like the Standard Model can be "renormalized" to manage these infinities, point-particle gravity is famously non-renormalizable. At the Planck scale (10^{-33} cm), the point-particle description of gravitational interactions simply breaks down.

String theory solves this problem by replacing point-like particles with one-dimensional loops and strands of string. This extension "smears out" interactions in spacetime. From a world-sheet perspective, what appears as a sharp, singular vertex in point-particle diagrams becomes a smooth, continuous topology in string theory. This smoothing effectively removes the UV divergences that plague point-particle gravity.

Furthermore, string theory is a uniquely constrained theory. It does not just allow for gravity; it necessitates it. The spectrum of string excitations naturally includes a massless spin-2 particle, which is identified as the graviton. Thus, general relativity emerges as a necessary low-energy consequence of the theory rather than an added assumption.

While the theory requires higher dimensions (e.g., $D = 26$ for the bosonic string) to maintain Lorentz invariance, it remains our most promising route toward a complete and consistent unification of the physical laws.

1.2 Action principles

The string moves in D-dimensional flat spacetime with metric $\eta_{\mu\nu} = \mathrm{diag}(-, +, +, \cdots, +)$.

First, let us review the classical mechanics of a zero-dimensional object. For a relativistic point particle, we can describe it using $D-1$ positions as functions of time $\mathbf{X}(X^0)$. However, this hides the covariance of the theory. Therefore, it is better to introduce a parameter τ along the particle's worldline and describe the motion in spacetime with D functions $X^\mu(\tau)$. The parameterization is arbitrary: different parameterizations of the same path are physically equivalent. That is, for any monotonic function $\tau'(\tau)$, the two paths are identical:

$$X'^\mu(\tau'(\tau)) = X^\mu(\tau) \tag{1.2.1}$$

The simplest Poincaré-invariant action (which does not depend on the parameterization) is proportional to the proper time along the worldline:

$$S_{\text{pp}} = -m \int \mathrm{d}\tau \, (-\dot{X}^\mu \dot{X}_\mu)^{1/2} \tag{1.2.2}$$

where $\dot{X}^\mu = \frac{\partial X^\mu}{\partial \tau}$. The variation of the action (after integration by parts) is:

$$\delta S_{\text{pp}} = +m \int \mathrm{d}\tau \, \tfrac{1}{2}(-\dot{X}^\nu \dot{X}_\nu)^{-1/2} \, 2\dot{X}_\mu \delta \dot{X}^\mu = m \int \mathrm{d}\tau \, u_\mu \delta \dot{X}^\mu \,,$$

so

$$\delta S_{\text{pp}} = -m \int \mathrm{d}\tau \, \dot{u}_\mu \delta X^\mu \tag{1.2.3}$$

where

$$u^\mu = \dot{X}^\mu (-\dot{X}^\nu \dot{X}_\nu)^{-1/2} \tag{1.2.4}$$

is the normalized D-velocity. Thus, the equation of motion $\dot{u}^\mu = 0$ describes free motion. The normalization constant m is the particle mass.

By introducing an auxiliary field on the worldline, a worldline-independent metric $\gamma_{\tau\tau}(\tau)$, this action can be transformed into another form. It is convenient to use the einbein $\eta(\tau) = (-\gamma_{\tau\tau}(\tau))^{1/2}$. This is defined to be positive-definite. We use the General Relativity term "tetrad" in any dimension (even though its root means "4"). Then:

$$S'_{\text{pp}} = \frac{1}{2} \int \mathrm{d}\tau \left(\eta^{-1} \dot{X}^\mu \dot{X}_\mu - \eta m^2 \right) \tag{1.2.5}$$

This action has the same symmetries as S_{pp}, namely Poincaré invariance and reparameterization invariance. For the latter, the transformation of $\eta(\tau)$ is:

$$\eta'(\tau')\mathrm{d}\tau' = \eta(\tau)\mathrm{d}\tau \tag{1.2.6}$$

Proof.

$$\begin{aligned}
S'_{\text{pp}} &= \frac{1}{2} \int \mathrm{d}\tau \left(\eta^{-1} \dot{X}^\mu \dot{X}_\mu - \eta m^2 \right) \\
&= \frac{1}{2} \int \mathrm{d}\tau' \left(\frac{\mathrm{d}\tau}{\mathrm{d}\tau'} \eta^{-1} \frac{\partial X^\mu}{\partial \tau} \frac{\partial X_\mu}{\partial \tau} - \frac{\mathrm{d}\tau}{\mathrm{d}\tau'} \eta m^2 \right) \\
&= \frac{1}{2} \int \mathrm{d}\tau' \left(\frac{\mathrm{d}\tau}{\mathrm{d}\tau'} \eta^{-1} \left(\frac{\mathrm{d}\tau'}{\mathrm{d}\tau} \right)^2 \frac{\partial X^\mu}{\partial \tau'} \frac{\partial X_\mu}{\partial \tau'} - \eta' m^2 \right) \\
&= \frac{1}{2} \int \mathrm{d}\tau' \left(\eta'^{-1} \frac{\partial X^\mu}{\partial \tau'} \frac{\partial X_\mu}{\partial \tau'} - \eta' m^2 \right)
\end{aligned}$$

□

Varying with respect to the einbein $\eta(\tau)$ yields the equation of motion:

$$\begin{aligned}
\delta S'_{\text{pp}} &= \frac{1}{2} \int \mathrm{d}\tau \left(-\frac{1}{\eta^{-2}} \dot{X}^\mu \dot{X}_\mu - m^2 \right) \delta\eta \\
&\implies \eta^2 = -\dot{X}^\mu \dot{X}_\mu / m^2
\end{aligned} \tag{1.2.7}$$

Using this to eliminate η from (1.2.5), S'_{pp} becomes S_{pp}. S'_{pp} can handle massless particles, whereas S_{pp} cannot. Although S'_{pp} and S_{pp} are classically equivalent, S_{pp} is difficult to path-integrate due to its complicated form. Conversely, S'_{pp} is a quadratic in derivatives, making the path integral quite easy to compute. One can infer that any attempt to define a quantum theory for S_{pp} will lead to results equivalent to the path integral of S'_{pp}. We take the latter as the starting point for the quantum theory.

A one-dimensional object sweeps out a two-dimensional worldsheet, described by two parameters, denoted $X^\mu(\tau,\sigma)$. We still insist that the action depends only on the embedding in spacetime, not on the parameterization. The simplest action is the Nambu-Goto action, which is proportional to the area of the worldsheet. Introduce the induced metric h_{ab}, where a, b range over (τ, σ):

$$h_{ab} = \partial_a X^\mu \partial_b X_\mu \tag{1.2.8}$$

Then the Nambu-Goto action is:

$$S_{\mathrm{NG}} = \int_M \mathrm{d}\tau \mathrm{d}\sigma \, \mathscr{L}_{NG} \tag{1.2.9a}$$

$$\mathscr{L}_{\mathrm{NG}} = -\frac{1}{2\pi\alpha'}(-\det h_{ab})^{1/2} \tag{1.2.9b}$$

where M represents the worldsheet, and α' has dimensions of L^2, representing the Regge slope. The relationship between the string tension T and the Regge slope is:

$$T = \frac{1}{2\pi\alpha'} \tag{1.2.10}$$

Now consider the symmetries of this action. Transformations of $X^\mu(\tau,\sigma)$ that satisfy $S_{\mathrm{NG}}(X') = S_{\mathrm{NG}}(X)$ include:

1. Isometry group of flat spacetime (D-dimensional Poincaré group):

$$X'^\mu(\tau,\sigma) = \Lambda^\mu{}_\nu X^\nu(\tau,\sigma) + a^\mu. \tag{1.2.11}$$

2. Two-dimensional coordinate invariance (called diffeomorphism invariance, abbreviated as diff):

$$X'^\mu(\tau',\sigma') = X^\mu(\tau,\sigma). \tag{1.2.12}$$

The NG action is analogous to S_{pp}. Similarly, we introduce a worldsheet-independent metric $\gamma_{ab}(\tau,\sigma)$. Henceforth, the "metric" refers to γ_{ab} (unless the induced metric is specified). Let γ_{ab} have Lorentz signature $(-,+)$. The action is:

$$S_{\mathrm{P}}[X,\gamma] = -\frac{1}{4\pi\alpha'}\int_M \mathrm{d}\tau \mathrm{d}\sigma \, (-\gamma)^{1/2}\gamma^{ab}\partial_a X^\mu \partial_b X_\mu \tag{1.2.13}$$

where $\gamma = \det \gamma^{ab}$. This is the Brink-Di Vecchia-Howe-Deser-Zumino action, or the Polyakov action. It was proposed to generalize local worldsheet supersymmetry. Polyakov emphasized its advantages, especially for path-integral quantization.

Equivalence with S_{NG}: Variation with respect to the metric:

$$\begin{aligned}\delta_\gamma S_{\mathrm{P}}[X,\gamma] &= -\frac{1}{4\pi\alpha'}\int_M \mathrm{d}\tau \mathrm{d}\sigma \left[\delta\gamma^{ab}(-\gamma)^{1/2}h_{ab} - \tfrac{1}{2}(-\gamma)^{-1/2}(-\gamma\gamma_{ab}\delta\gamma^{ab})h_{cd}\gamma^{cd}\right] \\ &= -\frac{1}{4\pi\alpha'}\int_M \mathrm{d}\tau \mathrm{d}\sigma \, (-\gamma)^{1/2}\delta\gamma^{ab}\Big(h_{ab} - \tfrac{1}{2}\gamma_{ab}\gamma^{cd}h_{cd}\Big)\end{aligned} \tag{1.2.14}$$

where we used:

$$\delta\gamma = \gamma\gamma^{ab}\delta\gamma_{ab} = -\gamma\gamma_{ab}\delta\gamma^{ab} \tag{1.2.15}$$

$\delta_\gamma S_{\mathrm{P}}[X,\gamma] = 0$ implies:

$$h_{ab} = \tfrac{1}{2}\gamma_{ab}\gamma^{cd}h_{cd} \tag{1.2.16}$$

Dividing both sides by $(-h)^{1/2}$:

$$h_{ab}(-h)^{-1/2} = \frac{\frac{1}{2}\gamma_{ab}\gamma^{cd}h_{cd}}{\sqrt{-\gamma}\frac{1}{2}\gamma^{cd}h_{cd}} = \gamma_{ab}(-\gamma)^{-1/2} \tag{1.2.17}$$

Thus γ_{ab} is proportional to h_{ab}. Using $h_{ab}\gamma^{ab}(-\gamma)^{1/2} = (-h)^{1/2}\gamma_{ab}\gamma^{ab}$ derived from (1.2.17), we eliminate γ_{ab} from S_{P}:

$$S_{\mathrm{P}}[X,\gamma] \to -\frac{1}{4\pi\alpha'}\int_M \mathrm{d}\tau\mathrm{d}\sigma\,(-\gamma)^{1/2}h_{ab}\gamma^{ab} = -\frac{1}{2\pi\alpha'}\int_M \mathrm{d}\tau\mathrm{d}\sigma\,(-h)^{1/2} = S_{\mathrm{NG}}[X] \tag{1.2.18}$$

S_{P} has the following symmetries:

1. D-dimensional Poincaré invariance:

$$\begin{aligned} X'^{\mu}(\tau,\sigma) &= \Lambda^{\mu}{}_{\nu}X^{\nu}(\tau,\sigma) + a^{\mu} \\ \gamma'_{ab}(\tau,\sigma) &= \gamma_{ab}(\tau,\sigma)\,. \end{aligned} \tag{1.2.19}$$

2. Diff invariance:

$$\begin{aligned} X'^{\mu}(\tau',\sigma') &= X^{\mu}(\tau,\sigma) \\ \frac{\partial\sigma'^{c}}{\partial\sigma^{a}}\frac{\partial\sigma'^{d}}{\partial\sigma^{b}}\gamma'_{cd}(\tau',\sigma') &= \gamma_{ab}(\tau,\sigma)\,. \end{aligned} \tag{1.2.20}$$

3. Two-dimensional Weyl invariance:

$$\begin{aligned} X'^{\mu}(\tau,\sigma) &= X^{\mu}(\tau,\sigma) \\ \gamma'_{ab}(\tau,\sigma) &= \exp[2\omega(\tau,\sigma)]\gamma_{ab}(\tau,\sigma) \end{aligned} \tag{1.2.21}$$

Proof.

$$\begin{aligned} S'_{\mathrm{P}} &= -\frac{1}{4\pi\alpha'}\int_M \mathrm{d}\tau\mathrm{d}\sigma\,(-\gamma')^{1/2}\gamma'^{ab}\partial_a X'^{\mu}\partial_b X'_{\mu} \\ &= -\frac{1}{4\pi\alpha'}\int_M \mathrm{d}\tau\mathrm{d}\sigma\,\mathrm{e}^{2\omega}(-\gamma)^{1/2}\mathrm{e}^{-2\omega}\gamma^{ab}h_{ab} \\ &= S_{\mathrm{P}} \end{aligned}$$

□

Weyl invariance, the local scaling of the worldsheet metric, has no counterpart in the NG formulation. Its appearance can be understood through the equation of motion (1.2.17) relating the Polyakov and NG actions. Eq. (1.2.17) does not uniquely determine γ_{ab}, but only a local scaling. Thus, Weyl-equivalent metrics correspond to the same embedding in spacetime. This is an additional redundancy in the Polyakov formulation. The variation of the action with respect to the metric gives the energy-momentum tensor:

$$T^{ab}(\tau,\sigma) = -4\pi(-\gamma)^{-1/2}\frac{\delta}{\delta\gamma_{ab}}S_{\mathrm{P}} = -\frac{1}{\alpha'}\left(\partial^a X^{\mu}\partial^b X_{\mu} - \tfrac{1}{2}\gamma_{ab}\partial_c X^{\mu}\partial^c X_{\mu}\right) \tag{1.2.22}$$

Proof. Using $\gamma_{ab}\delta\gamma^{ab} = -\gamma^{ab}\delta\gamma_{ab}$, we have:

$$\begin{aligned} \delta\gamma^{ab}h_{ab} - \tfrac{1}{2}\delta\gamma^{ab}\gamma_{ab}\gamma^{cd}h_{cd} &= -\gamma^{ac}\delta\gamma_{ab}\gamma^{bd}h_{cd} + \tfrac{1}{2}\gamma^{ab}\delta\gamma_{ab}\gamma^{cd}h_{cd} \\ &= -\delta\gamma_{ab}h^{ab} + \tfrac{1}{2}\gamma^{ab}\gamma^{cd}h_{cd}\delta\gamma_{ab} \\ &\implies \frac{\delta}{\delta\gamma_{ab}}S_{\mathrm{P}} = -\frac{1}{4\pi\alpha'}(-\gamma)^{1/2}(-h^{ab} + \tfrac{1}{2}\gamma^{ab}\gamma^{cd}h_{cd}) \end{aligned}$$

□

The energy-momentum tensor is conserved ($\nabla_a T^{ab} = 0$), which is a consequence of diff invariance. Weyl invariance implies:

$$\gamma_{ab}\frac{\delta}{\delta\gamma_{ab}}S_{\mathrm{P}} = 0 \Rightarrow T_a{}^a = 0 \tag{1.2.23}$$

Proof. $\delta S \sim \int \delta\gamma_{ab}T^{ab}$, and for a Weyl transformation $\delta\gamma_{ab} = 2\omega\gamma_{ab}$, so $T^{ab}\gamma_{ab} = 0$. □

The S_{NG} and S_{P} actions define 2D field theories on the string worldsheet. In string theory, we see that amplitudes for spacetime processes are given by matrix elements of the 2D quantum field theory on the worldsheet. Although the physical world is 4D spacetime, most mechanisms used in string perturbation theory are 2D. Eq. (1.2.12) defines $X^\mu(\tau,\sigma)$ as scalar fields, with μ acting as an internal index. From a 2D perspective, the Polyakov action describes massless Klein-Gordon scalars X^μ covariantly coupled to γ_{ab}. Thus, Poincaré invariance is an internal symmetry.

Variation with respect to γ_{ab} gives the equation of motion:

$$T_{ab} = 0 \tag{1.2.24}$$

Using integration by parts, the variation with respect to X^μ gives:

$$\begin{aligned} 0 &= (-\gamma)^{1/2}\gamma^{ab}(2\partial_a X^\mu \delta\partial_b X_\mu) = -2\partial_b[(-\gamma)^{1/2}\gamma^{ab}\partial_a X^\mu]\delta X_\mu \\ &\Longrightarrow \partial_a[(-\gamma)^{1/2}\gamma^{ab}\partial_b X^\mu] = (-\gamma)^{1/2}\nabla^2 X^\mu = 0 \end{aligned} \tag{1.2.25}$$

Proof.

$$\begin{aligned} 0 &= (-\gamma)^{1/2}\partial_a\gamma^{ab}\partial_b X^\mu + \gamma^{ab}\partial_b X^\mu \partial_a(-\gamma)^{1/2} \\ &= (-\gamma)^{1/2}[\partial_a(\gamma^{ab}\partial_b X^\mu) + (-\gamma)^{-1/2}\partial_a(-\gamma)^{1/2}\gamma^{ab}\partial_b X^\mu] \\ &= (-\gamma)^{1/2}[\nabla_a(\gamma^{ab}\partial_b X^\mu)] = (-\gamma)^{1/2}\nabla_a\gamma^{ab}\nabla_b X^\mu = (-\gamma)^{1/2}\nabla^2 X^\mu \end{aligned}$$

where we used $\nabla_a = \partial_a + (\frac{1}{\sqrt{-\gamma}}\partial_a\sqrt{-\gamma})$. Note that X^μ is a scalar, so $\partial_b \to \nabla_b$. □

For a worldsheet with boundaries, there are also surface terms in the variation of the action. Specifically, let the coordinate region be:

$$-\infty < \tau < \infty, \quad 0 \le \sigma \le \ell \tag{1.2.26}$$

Taking τ, σ as time and space variables, then:

$$\begin{aligned} \delta S_{\mathrm{P}} &= \frac{1}{2\pi\alpha'}\int_{-\infty}^{\infty} \mathrm{d}\tau \int_0^\ell \mathrm{d}\sigma\, (-\gamma)^{1/2}\delta X^\mu \nabla^2 X_\mu \\ &\quad - \frac{1}{2\pi\alpha'}\int_{-\infty}^{\infty}\mathrm{d}\tau\, (-\gamma)^{1/2}\delta X^\mu \partial^\sigma X_\mu\Big|_{\sigma=0}^{\sigma=\ell} \end{aligned} \tag{1.2.27}$$

The boundary term vanishes if:

1.
$$\partial^\sigma X^\mu(\tau, 0) = \partial^\sigma X^\mu(\tau,\ell) = 0 \tag{1.2.28}$$

   This is the Neumann boundary condition. A more covariant form is:

$$n^a\partial_a X_\mu = 0 \quad \text{on } \partial M \tag{1.2.29}$$

where n^a is normal to the boundary ∂M. Ends of open strings move freely in spacetime.

> **Remark.** *This boundary condition is invariant under Poincare transformation. It can be shown as:*
>
> $$\partial_a X'_\mu = \partial_a(\Lambda^\mu_\nu X_\mu + b_\nu) = \Lambda^\mu_\nu \underbrace{\partial_a X_\mu}_{=0} = 0$$

2.

$$X^\mu(\tau,\ell) = X^\mu(\tau,0), \quad \partial^\sigma X^\mu(\tau,\ell) = \partial^\sigma X^\mu(\tau,0) \tag{1.2.30a}$$

$$\gamma_{ab}(\tau,\ell) = \gamma_{ab}(\tau,0) \tag{1.2.30b}$$

This field is periodic, has no boundary, and forms a closed string.

3. The last possible boundary condition has endpoints fixed at

$$X^\mu(\tau,\ell) = X^\mu_\ell \qquad X^\mu(\tau,0) = X^\mu_0 \tag{1.2.31a}$$

The first two conditions are the only possibilities consistent with D-dimensional Poincaré invariance and the equations of motion. The NG and Polyakov actions are perhaps the simplest actions with the given symmetries. But simplicity is not the correct criterion; symmetry is key.

Now we generalize the Polyakov action, requiring that the symmetries are preserved and that the action is a polynomial in derivatives. Global Weyl invariance, i.e., $\omega(\tau,\sigma) = C$ (where C is a constant), requires that γ^{ab} appears one more time than γ_{ab} to cancel the variation of $(-\gamma)^{1/2}$. Additional upper indices can only contract with derivatives, so each term must have two derivatives. Coordinate invariance and Poincaré invariance allow:

$$\chi = \frac{1}{4\pi}\int_M \mathrm{d}\tau \mathrm{d}\,\sigma(-\gamma)^{1/2} R \tag{1.2.32}$$

where R is the Ricci scalar. Under a local Weyl scaling:

$$\left(-\gamma'\right)^{1/2} R' = (-\gamma)^{1/2}\left(R - 2\nabla^2\omega\right) \tag{1.2.33}$$

> *Proof.*
>
> $$\begin{aligned} R' &= \frac{R}{\mathrm{e}^{2\omega}} - 2g^{\rho\sigma}\frac{\nabla_\rho\partial_\sigma \mathrm{e}^\omega}{\mathrm{e}^{3\omega}} + 2g^{\rho\sigma}\frac{(\partial_\rho \mathrm{e}^\omega)(\partial_\sigma \mathrm{e}^\omega)}{\mathrm{e}^{4\omega}} \\ &= \frac{R}{\mathrm{e}^{2\omega}} - 2g^{\rho\sigma}\frac{\nabla_\rho\partial_\sigma \mathrm{e}^\omega}{\mathrm{e}^{3\omega}} - 2g^{\rho\sigma}\frac{(\partial_\sigma\partial_\rho\omega)\omega}{\mathrm{e}^{2\omega}} \\ &= \frac{R}{\mathrm{e}^{2\omega}} - 2g^{\rho\sigma}\frac{\nabla_\rho \mathrm{e}^\omega\partial_\sigma\omega}{\mathrm{e}^{3\omega}} - 2g^{\rho\sigma}\frac{(\nabla_\sigma\partial_\rho\omega)\,\omega}{\mathrm{e}^{2\omega}} \end{aligned}$$
>
> But
>
> $$\nabla_\rho \mathrm{e}^\omega\partial_\sigma\omega = \partial_\rho \mathrm{e}^\omega\partial_\sigma\omega - \Gamma^\lambda_{\rho\sigma}\mathrm{e}^\omega\partial_\lambda\omega = \left(-\nabla_\rho\partial_\sigma\omega - \Gamma^\lambda_{\rho\sigma}\partial_\lambda\omega\right)\mathrm{e}^\omega$$
>
> Then the last two terms become:
>
> $$\frac{-2g^{\rho\sigma}}{\mathrm{e}^{2\omega}}\left(-\nabla_\rho\partial_\sigma\omega - T^\lambda_{\rho\sigma}\partial_\lambda\omega + \omega\nabla_\sigma\partial_\rho\omega\right)$$
>
> □

In (1.2.33), the variation is a total derivative because for any v^a, we have $(-\gamma)^{1/2}\nabla_a v^a = \partial_a[(-\gamma)^{1/2}v^a]$. Thus (1.2.32) is invariant for a worldsheet without boundaries. With boundaries, there is an extra surface term.

We can include χ in the action:

$$\begin{aligned} S'_{\mathrm{P}} &= S_{\mathrm{P}} - \lambda\chi \\ &= -\int_M \mathrm{d}\tau\mathrm{d}\sigma\, (-\gamma)^{1/2}\left(\frac{1}{4\pi\alpha'}\gamma^{ab}\partial_a X^\mu \partial_b X_\mu + \frac{\lambda}{4\pi}R\right) \end{aligned} \tag{1.2.34}$$

This is the most general (diff $\times$ Weyl)-invariant and Poincaré-invariant action. S'_{P} looks like the Hilbert action $\int(-\gamma)^{1/2}R$ for the (gravitational) metric minimally coupled to D massless scalar fields. However, in 2D, the Hilbert action depends only on the topology of the worldsheet; its variation is $R_{ab} - \frac{1}{2}\gamma_{ab}R$. In 2D, the symmetries of the curvature tensor imply $R_{ab} = \frac{1}{2}\gamma_{ab}R$. The Hilbert action is invariant under continuous transformations of the metric, but has global effects.

Remark. *In 2D, we have* $R_{abcd} = \frac{R}{2}\left(\gamma_{ac}\gamma_{bd} - \gamma_{ad}\gamma_{bc}\right)$ *and as a result:*

$$\begin{aligned} R_{ab} = \gamma^{cd}R_{acbd} = \frac{R}{2}\gamma^{cd}\left(\gamma_{ac}\gamma_{bd} - \gamma_{ad}\gamma_{bc}\right) = \frac{R}{2}\left(\delta^d_a\gamma_{bd} - \delta^c_a\gamma_{bc}\right) \\ = \frac{R}{2}\gamma_{ab}. \end{aligned} \tag{1.2.35}$$

1.3 The open string spectrum

Introduce light-cone coordinates in spacetime:

$$x^{\pm} = 2^{-1/2}\left(x^0 \pm x^1\right), \quad x^i, \quad i = 2, \dots, D-1 \tag{1.3.1}$$

Spacetime coordinates are written as X^μ, and fields on the worldsheet are written as $X^\mu(\tau,\sigma)$. In these coordinates, the metric is:

$$a^\mu b_\mu = -a^+b^- - a^-b^+ + a^ib^i \tag{1.3.2a}$$
$$a_- = -a^+, \quad a_+ = -a^-, \quad a_i = a^i \tag{1.3.2b}$$

Let the worldsheet parameter τ be the spacetime coordinate X^+. Then X^+ plays the role of time, and p^- is the energy. The longitudinal components X^-, p^+ act like spatial coordinates and momentum, as do the transverse X^i, p^i. Starting from the point particle, using S'_{pp}: Fix the worldline parameterization via:

$$X^+(\tau) = \tau \tag{1.3.3}$$

The action becomes:

$$S'_{\mathrm{pp}} = \frac{1}{2}\int \mathrm{d}\tau\left(-2\eta^{-1}\dot{X}^- + \eta^{-1}\dot{X}^i\dot{X}^i - \eta m^2\right) \tag{1.3.4}$$

Remark. *Here we used* $\dot{X}^\mu\dot{X}_\mu = -\dot{X}^+\dot{X}^- - \dot{X}^+\dot{X}^- + \dot{X}^i\dot{X}^i = -2\dot{X}^- + \dot{X}^i\dot{X}^i$.

The canonical momentum $p_\mu = \partial L/\partial\dot{X}^\mu$ is:

$$p_- = -\eta^{-1}, \quad p_i = \eta^{-1}\dot{X}^i \tag{1.3.5}$$

Using the metric $p^i = p_i$ and $p^+ = -p_-$, the Hamiltonian is:

$$\begin{aligned} H &= p_-\dot{X}^- + p_i\dot{X}^i - L \\ &= -\eta^{-1}\dot{X}^s + \eta p_i p^i + \eta^{-1}\dot{X}^s - \tfrac{1}{2}\eta p_i p_i + \tfrac{1}{2}\eta m^2 \\ &= \frac{p^i p^i + m^2}{2\eta^{-1}} = \frac{p^i p^i + m^2}{2p^+} \end{aligned} \tag{1.3.6}$$

There is no p_+X^+ term because X^+ is not a dynamical variable. Furthermore, $\dot{\eta}$ does not appear in the action, so $p_\eta = 0$. Since $\eta = -1/p_-$, we should not treat η as an independent canonical coordinate.

For quantization, impose canonical commutators on the dynamical fields:

$$\left[p_i, X^j\right] = -i\delta_i^j, \quad \left[p_-, X^-\right] = -i \tag{1.3.7}$$

The momentum eigenstates $|k_-, k^i\rangle$ form a complete set. The remaining momentum components are determined by other quantities. The gauge choice associates τ with X^+, so $H = -p_+ = p^-$. The negative sign between H and p_+ is because the former is active and the latter is passive. Thus, (1.3.6) becomes the relativistic mass-shell condition. We have obtained the spectrum of a relativistic scalar.

Turning to the open string. The coordinate region is $-\infty < \tau < +\infty$ and $0 \leq \sigma \leq l$. Choose the gauge to remove Polyakov redundancies. Let:

$$X^+ = \tau \tag{1.3.8a}$$

$$\partial_\sigma \gamma_{\sigma\sigma} = 0 \tag{1.3.8b}$$

$$\det \gamma_{ab} = -1 \tag{1.3.8c}$$

How to obtain this gauge? First, choose τ according to (1.3.8a). Note that $f = \gamma_{\sigma\sigma}\left(-\det\gamma_{ab}\right)^{-1/2}$ transforms as:

$$f'\mathrm{d}\sigma' = f\mathrm{d}\sigma \tag{1.3.9}$$

Define the invariant length $\mathrm{d}l = f\mathrm{d}\sigma$. Define the σ coordinate of a point proportional to the invariant length from $\sigma = 0$ to that point. The proportionality constant is determined by keeping the right endpoint at $\sigma = l$. In this coordinate system, $f = \mathrm{d}l/\mathrm{d}\sigma$ is independent of σ, though it may depend on τ. Finally, perform a Weyl transformation to satisfy (1.3.8c). Since f is Weyl-invariant, $\partial_\sigma f$ remains zero. Combined with (1.3.8c), this implies $\partial_\sigma\gamma_{\sigma\sigma} = 0$, so (1.3.8b) is also satisfied.

For $\gamma_{\tau\tau}(\tau,\sigma)$, we can solve the gauge condition (1.3.8c). Since $\gamma_{\sigma\sigma}$ is independent of σ, the only independent degrees of freedom in the metric are now $\gamma_{\sigma\sigma}(\tau)$ and $\gamma_{\sigma\tau}(\tau,\sigma)$. The inverse metric is:

$$\begin{bmatrix} \gamma^{\tau\tau} & \gamma^{\tau\sigma} \\ \gamma^{\sigma\tau} & \gamma^{\sigma\sigma} \end{bmatrix} = \begin{bmatrix} -\gamma_{\sigma\sigma}(\tau) & \gamma_{\tau\sigma}(\tau,\sigma) \\ \gamma_{\tau\sigma}(\tau,\sigma) & \gamma_{\sigma\sigma}^{-1}(\tau)\left(1 - \gamma_{\tau\sigma}^2(\tau,\sigma)\right) \end{bmatrix} \tag{1.3.10}$$

Proof.

$$\begin{bmatrix} \gamma_{\tau\tau} & \gamma_{\tau\sigma} \\ \gamma_{\tau\sigma} & \gamma_{\sigma\sigma} \end{bmatrix} \begin{bmatrix} \gamma^{\tau\tau} & \gamma^{\tau\sigma} \\ \gamma^{\tau\sigma} & \gamma^{\sigma\sigma} \end{bmatrix} = \begin{bmatrix} -\gamma_{\tau\tau}\gamma_{\sigma\sigma} + \gamma_{\tau\sigma}^2 & \gamma_{\tau\tau}\gamma_{\tau\sigma} + \gamma_{\tau\sigma}\gamma_{\sigma\sigma}^{-1}\left(1 - \gamma_{\tau\sigma}^2\right) \\ -\gamma_{\sigma\sigma}\gamma_{\tau\sigma} + \gamma_{\sigma\sigma}\gamma_{\tau\sigma} & \gamma_{\tau\sigma}^2 + 1 - \gamma_{\tau\sigma}^2 \end{bmatrix}$$

Then use (1.3.8c) given by $\gamma_{\tau\tau}\gamma_{\sigma\sigma} - \gamma_{\tau\sigma}^2 = -1$. □

Thus, the Polyakov Lagrangian becomes:

$$\begin{aligned} L = -\frac{1}{4\pi\alpha'}\int_0^\ell \mathrm{d}\sigma \Big[&\gamma_{\sigma\sigma}\left(2\partial_\tau x^- - \partial_\tau X^i\partial_\tau X^i\right) \\ &- 2\gamma_{\sigma\tau}(\partial_\sigma Y^- - \partial_\tau X^i\partial_\sigma X^i) + \gamma_{\sigma\sigma}^{-1}(1 - \gamma_{\tau\sigma}^2)\partial_\sigma X^i\partial_\sigma X^i\Big] \end{aligned} \tag{1.3.11}$$

Proof. Expand the Polyakov action (1.2.13) by components:

$$L = -\frac{1}{4\pi\alpha'}\int_0^\ell \mathrm{d}\sigma\left[-\gamma_{\sigma\sigma}\underbrace{(\partial_\tau X^\mu \partial_\tau X_\mu)}_{(1)} + 2\gamma_{\tau\sigma}\underbrace{(\partial_\tau X^\mu \partial_\sigma X_\mu)}_{(2)} + \gamma_{\sigma\sigma}^{-1}(1-\gamma_{\tau\sigma}^2)\underbrace{(\partial_\sigma X^\mu \partial_\sigma X_\mu)}_{(3)}\right]$$

where

$$\begin{aligned}
(1) &= -\dot{X}^+\dot{X}^- - \dot{X}^+\dot{X}^- + \dot{X}^i\dot{X}^i = -2\partial_\tau X^- + \partial_\tau X^i \partial_\tau X^i \\
(2) &= -\partial_\tau X^\dagger \partial_\sigma X^- - \partial_\tau X^- \partial_\sigma X^+ + \partial_\tau X^i \partial_\sigma X^i \\
&= -\partial_\sigma X^- - \partial_\tau X^- \partial_\sigma \tau + \partial_\tau X^i \partial_\tau X^i = -\left(\partial_\sigma Y^- - \partial_\tau X^i \partial_\sigma X^i\right) \\
(3) &= -\partial_\sigma X^- \partial_\sigma X^+ - \partial_\sigma X^+ \partial_\sigma X^- + \partial_\sigma X^i \partial_\sigma X^i = \partial_\sigma X^i \partial_\sigma X^i
\end{aligned}$$

Q.E.D. □

Here we split $X^-(\tau,\sigma)$ into two parts: $x^-(\tau)$ and $Y^-(\tau,\sigma)$. $x^-(\tau)$ is the average value of X^- at time τ.

$$x^-(\tau) = \frac{1}{\ell}\int_0^\ell \mathrm{d}\sigma X^-(\tau,\sigma) \tag{1.3.12a}$$

$$Y^-(\tau,\sigma) = X^-(\tau,\sigma) - x^-(\tau) \tag{1.3.12b}$$

Y^- does not appear in terms containing time derivatives and is thus non-dynamical. It acts like a Lagrange multiplier, constraining $\partial_\sigma \gamma_{\tau\sigma}$ to be zero. Under gauge (1.3.8), the open string boundary condition (1.2.28) becomes:

$$\gamma_{\tau\sigma}\partial_\tau X^\mu - \gamma_{\tau\tau}\partial_\sigma X^\mu = 0 \quad \text{at } \sigma = 0, \ell \tag{1.3.13}$$

Proof. $\partial^\sigma X^\mu = \gamma^{\sigma a}\partial_a X^\mu = \gamma^{\sigma\tau}\partial_\tau X^\mu + \gamma^{\sigma\sigma}\partial_\sigma X^\mu = \gamma_{\tau\sigma}\partial_\tau X^\mu + \gamma_{\sigma\sigma}^{-1}\left(1-\gamma_{\tau\sigma}^2\right)\partial_\sigma X^\mu = \gamma_{\tau\sigma}\partial_\tau X^\mu - \gamma_{\tau\tau}\partial_\sigma X^\mu$ Q.E.D. □

For $\mu = +$:

$$\gamma_{\tau\sigma} = 0 \quad \text{at } \sigma = 0, \ell \tag{1.3.14}$$

Since $\partial_\sigma \gamma_{\tau\sigma} = 0$, $\gamma_{\tau\sigma}$ vanishes everywhere. For $\mu = i$, the boundary condition is:

$$\partial_\sigma X^i = 0 \quad \text{at } \sigma = 0, \ell \tag{1.3.15}$$

Imposing gauge conditions and including the Lagrange multiplier Y^-, the Lagrangian is:

$$L = -\frac{\ell}{2\pi\alpha'}\gamma_{\sigma\sigma}\partial_\tau x^- + \frac{1}{4\pi\alpha'}\int_0^\ell \mathrm{d}\sigma\left(\gamma_{\sigma\sigma}\partial_\tau X^i \partial_\tau X^i - \gamma_{\sigma\sigma}^{-1}\partial_\sigma X^i \partial_\sigma X^i\right) \tag{1.3.16}$$

The conjugate momentum to x^- is:

$$p_- = -p^+ = \frac{\partial L}{\partial\left(\partial_\tau x^-\right)} = -\frac{\ell}{2\pi\alpha'}\gamma_{\sigma\sigma} \tag{1.3.17}$$

As with the einbein η in the particle case, $\gamma_{\sigma\sigma}$ is momentum rather than a coordinate. The conjugate momentum density to $X^i(\tau,\sigma)$ is:

$$\Pi^i = \frac{\delta L}{\delta\left(\partial_\tau X^i\right)} = \frac{1}{2\pi\alpha'}\gamma_{\sigma\sigma}\partial_\tau X^i = \frac{p^+}{\ell}\partial_\tau X^i \tag{1.3.18}$$

Thus, the Hamiltonian is:

$$\begin{aligned}
H &= p_- \partial_\tau x^- - L + \int_0^\ell \mathrm{d}\sigma \, \Pi_i \partial_\tau X^i \\
&= p_- \partial_\tau x^- - p_- \partial_\tau x^- - \frac{1}{4\pi\alpha'} \int_0^\ell \mathrm{d}\sigma \left(\gamma_{\sigma\sigma} \partial_\tau X^i \partial_\tau X^i - \frac{\partial_\sigma X^i \partial_\sigma X^i}{\gamma_{\sigma\sigma}} \right) + \int_0^\ell \mathrm{d}\sigma \, \frac{\ell \, \Pi_i^2}{p^+} \\
&= \frac{\ell}{4\pi\alpha' p^+} \int_0^\ell \mathrm{d}\sigma \, 4\pi\alpha' \Pi_i^2 - \frac{\gamma_{\sigma\sigma}^{-1}}{4\pi\alpha'} \int_0^\ell \mathrm{d}\sigma \left(\left(2\pi\alpha' \Pi^i\right)^2 - \partial_\sigma X^i \partial_\sigma X^i \right) \\
&= \frac{\ell}{4\pi\alpha' p^+} \int_0^\ell d\sigma \left(2\pi\alpha' \Pi^i \Pi^i + \frac{1}{2\pi\alpha'} \partial_\sigma X^i \partial_\sigma X^i \right)
\end{aligned} \tag{1.3.19}$$

This is precisely the Hamiltonian for $D-2$ free fields X^i, where p^+ is a conserved quantity. The equations of motion are:

$$\partial_\tau x^- = \frac{\partial H}{\partial p_-} = -\frac{\partial H}{\partial p^+} = -\frac{H p^+ \partial (p^+)^{-1}}{\partial p^+} = \frac{H}{p^+}\,, \quad \partial_\tau p^+ = \frac{\partial H}{\partial x^-} = 0 \tag{1.3.20a}$$

$$\partial_\tau X^i = \frac{\delta H}{\delta \Pi^i} = \frac{\ell}{2p^+} 2\Pi^i = 2\pi\alpha' c \Pi^i\,, \quad \partial_\tau \Pi^i = -\frac{\delta H}{\delta X^i} = \frac{c}{2\pi\alpha'} \partial_\sigma^2 X^i\,, \tag{1.3.20b}$$

where $c = \ell/(2\pi\alpha' p^+)$. This leads to the wave equation:

$$\partial_\tau^2 X^i = c^2 \partial_\sigma^2 X^i \tag{1.3.21}$$

Choose ℓ such that $c = 1$. Thus p^+ is conserved, and the total string length ℓ remains constant.

$X^\pm$ also satisfy the wave equation. X^- requires some calculation. Under (1.3.15), the general solution to the wave equation is:

$$X^i(\tau, \sigma) = x^i + \frac{p^i}{p^+}\tau + i \left(2\alpha'\right)^{1/2} \sum_{\substack{n=-\infty \\ n \neq 0}}^{\infty} \frac{1}{n} \alpha_n^i \exp\left(-\frac{\pi i n c \tau}{\ell} \right) \cos \frac{\pi n \sigma}{\ell} \tag{1.3.22}$$

The reality of X^i requires $\alpha_{-n}^i = \left(\alpha_n^i\right)^\dagger$. Define center-of-mass variables:

$$x^i(\tau) = \frac{1}{\ell} \int_0^\ell \mathrm{d}\sigma \; X^i(\tau, \sigma) \tag{1.3.23a}$$

$$p^i(\tau) = \int_0^\ell \mathrm{d}\sigma \; \Pi^i(\tau, \sigma) = \frac{p^+}{\ell} \int_0^\ell \mathrm{d}\sigma \; \partial_\tau X^i(\tau, \sigma) \tag{1.3.23b}$$

These represent the average mass and total momentum. These are Heisenberg operators. In (1.3.22), they are Schrödinger operators $x^i \equiv x^i(0)$ and $p^i \equiv p^i(0)$. For quantization, impose commutation relations:

$$\left[x^-, p^+\right] = i\eta^{-+} = -i \tag{1.3.24a}$$

$$\left[X^i(\sigma), \Pi^j\left(\sigma'\right)\right] = i\delta^{ij} \delta\left(\sigma - \sigma'\right) \tag{1.3.24b}$$

Written in terms of Fourier components:

$$\left[x^i, p^j\right] = i\delta^{ij} \tag{1.3.25a}$$

$$\left[\alpha_m^i, \alpha_n^j\right] = m\delta^{ij} \delta_{m,-n} \tag{1.3.25b}$$

Proof. Equations (1.3.24b), (1.3.25a), and (1.3.25b) form a logical loop; knowing any two allows derivation of the third. Assume the latter two are known to prove the first. Introduce the notation:

$$\sum{}' = \sum_{\substack{n=-\infty \\ n\neq 0}}^{+\infty}$$

From (1.3.18) and (1.3.22), we have:

$$\begin{aligned} \Pi^i(\tau,\sigma) =& \frac{p^i}{\ell} + \frac{p^+}{\ell} i\left(2\alpha'\right)^{1/2} \sum{}' \frac{1}{n}\alpha_n^i \left(-\frac{\pi i n c}{\ell}\right) \exp\left(-\frac{\pi i n c\tau}{\ell}\right) \cos\frac{n\pi\sigma}{\ell} \\ =& \frac{p^i}{\ell} + \frac{1}{\left(2\alpha'\right)^{1/2}\ell} \sum{}' \alpha_n^i \exp\left(-\frac{\pi i n c\tau}{\ell}\right) \cos\frac{n\pi\sigma}{\ell} \end{aligned} \tag{1.a}$$

where we used:

$$i\left(2\alpha'\right)^{1/2} \frac{-i\pi c}{\ell} = \pi\left(2\alpha'\right)^{1/2} \frac{1}{2\alpha'\pi p^+} = \frac{1}{(2\alpha')^{1/2} p^+}$$

Then:

$$[X^i, \Pi^i] = \frac{1}{\ell}\left[x^i, p^i\right] + \frac{i}{\ell}\sum_{n,m}{}' \frac{1}{n} \exp\left(-\frac{\pi i n c\tau}{\ell}\right) \cos\frac{n\pi\sigma}{\ell} \left[\alpha_n^i, \alpha_m^i\right] \exp\left(-\frac{\pi i m c\tau}{\ell}\right) \cos\frac{m\pi\sigma'}{\ell}$$

Let $\left[\alpha_m^i, \alpha_n^i\right] = m\delta_{m,-n}$. Given:

$$2\pi\delta(\sigma' - \sigma) = \sum_{n=-\infty}^{+\infty} \mathrm{e}^{in(\sigma'-\sigma)} = 1 + \sum{}' \mathrm{e}^{in(\sigma'-\sigma)}$$

The latter term becomes:

$$\begin{aligned} & -\frac{i}{\ell}\sum{}' \frac{1}{n} \exp\left(-\frac{\pi i n c\tau}{\ell}\right) \cos\frac{n\pi\sigma}{\ell} (-n) \exp\left(+\frac{\pi i n c\tau}{\ell}\right) \cos\frac{-n\pi\sigma'}{\ell} \\ =& \frac{i}{\ell}\sum{}' \cos\frac{n\pi\sigma}{\ell} \cos\frac{n\pi\sigma'}{\ell} \\ =& \frac{i}{4\ell}\sum{}' \left[\mathrm{e}^{in\pi(\sigma+\sigma')/\ell} + \mathrm{e}^{-in\pi(\sigma+\sigma')/\ell} + \mathrm{e}^{in\pi(\sigma-\sigma')/\ell} + \mathrm{e}^{-in\pi(\sigma-\sigma')/\ell}\right] \\ =& \frac{i}{2\ell}\Big(2\pi\delta\big(\pi(\sigma+\sigma')/\ell\big) - 1 + 2\pi\delta\big(\pi(\sigma-\sigma')/\ell\big) - 1\Big) \\ =& -\frac{i}{\ell} + \frac{i\pi}{\ell}\delta\big(\pi(\sigma-\sigma')/\ell\big) = -\frac{i}{\ell} + i\delta(\sigma-\sigma') \end{aligned}$$

Note $\delta(\sigma - \sigma') = \delta(\sigma' - \sigma)$, giving a factor of 2; and $\sigma, \sigma' > 0$ makes $\delta\big(\pi(\sigma+\sigma')/\ell\big) = 0$. Since $[x^i, p^i] = i$, we have:

$$\left[X^i, \Pi^i\right] = i\delta(\sigma - \sigma')$$

Q.E.D. □

For each m and i, the modes satisfy the harmonic oscillator algebra:

$$\alpha_m^i \sim m^{1/2} a, \quad \alpha_{-m}^i \sim m^{1/2} a^\dagger, \quad m > 0 \tag{1.3.26}$$

where $[a, a^\dagger] = 1$. The state $|0;k\rangle$ (where $k = (k^+, k^i)$) is annihilated by lowering operators and is an eigenstate of the center-of-mass momentum:

$$p^+|0;k\rangle = k^+|0;k\rangle, \quad p^i|0;k\rangle = k^i|0;k\rangle \tag{1.3.27a}$$

$$\alpha_m^i|0;k\rangle = 0, \quad m > 0 \tag{1.3.27b}$$

General states are obtained by acting with raising operators on $|0;k\rangle$:

$$|N;k\rangle = \left[\prod_{i=2}^{D-1}\prod_{n=1}^{\infty}\frac{\left(\alpha_{-n}^{i}\right)^{N_{in}}}{\left(n^{N_{in}}N_{in}!\right)^{1/2}}\right]|0;k\rangle \tag{1.3.28}$$

Independent states can be labeled by center-of-mass momenta k^+, k^i and occupation numbers N_{in}. Center-of-mass momentum represents the point particle degrees of freedom, while oscillators represent infinite internal degrees of freedom. From a spacetime view, each choice of occupation numbers corresponds to a different particle state or spin. States (1.3.28) form the single open string Hilbert space $\mathscr{H}$. In particular, state $|0;0\rangle$ is the ground state of a single string with zero momentum, not the vacuum state without strings. We call the latter $|\text{vacuum}\rangle$. Operators acting on it cannot create or annihilate strings, but only act on the state space of individual strings. The n-string Hilbert space $\mathscr{H}_n$ is formed by the product of n spaces of (1.3.28); wave functions must be symmetric, as all these states have integer spin (to be proven). In the free limit, the total Hilbert space of string theory is:

$$\mathscr{H} = |\text{vacuum}\rangle \oplus \mathscr{H}_1 \oplus \mathscr{H}_2 \oplus \ldots \tag{1.3.29}$$

Substituting (1.3.22) into (1.3.19) yields:

$$H = \frac{p^i p^i}{2p^+} + \frac{1}{2p^+\alpha'}\left(\sum_{n=1}^{\infty}\alpha_{-n}^{i}\alpha_{n}^{i} + A\right) \tag{1.3.30}$$

Proof. First, split the string Hamiltonian (1.3.19) into two parts:

$$H_\Pi = \frac{\ell}{2p^+}\int_0^\ell \mathrm{d}\sigma\, \Pi^i\Pi^i\,, \quad H_X = \frac{\ell}{2p^+(2\pi\alpha')^2}\int_0^\ell \mathrm{d}\sigma\, \partial_\sigma X^i \partial_\sigma X^i\,.$$

Using (1.a), we have:

$$\begin{aligned} H_\Pi &= \frac{\ell}{2p^+}\int_0^\ell \mathrm{d}\sigma \left(\frac{p^ip^i}{\ell^2} + \frac{1}{2\alpha'\ell^2}\sum_{n,m}{}' \alpha_n^i\alpha_m^i \exp\Big(-\frac{\pi i c\tau}{\ell}(n+m)\Big)\cos\frac{n\pi\sigma}{\ell}\cos\frac{m\pi\sigma}{\ell}\right.\\ &\qquad\left. + \frac{2p^i}{(2\alpha')^{1/2}\ell^2}\sum{}'\alpha_n^i \exp\left(-\frac{\pi i n c\tau}{\ell}\right)\cos\frac{n\pi\sigma}{\ell}\right)\\ &= \frac{p^ip^i}{2p^+} + \frac{1}{4p^+\alpha'}\sum_{n,m}{}'\alpha_n^i\alpha_m^i\exp\Big(-\frac{\pi i c\tau}{\ell}(n+m)\Big)\big(\delta_{n,m}+\delta_{n,-m}\big)/2 \end{aligned}$$

Using (1.3.22), we have:

$$\partial_\sigma X^i = -\frac{i\,(2\alpha')^{1/2}\,\pi}{\ell}\sum{}'\alpha_n^i\exp\Big(-\frac{\pi i n c\tau}{\ell}\Big)\sin\frac{n\pi\sigma}{\ell}\,,$$

Then:

$$\begin{aligned} H_X &= \frac{\ell}{2p^+(2\pi\alpha')^2}\frac{-2\alpha'\pi^2}{\ell^2}\int_0^\ell \mathrm{d}\sigma\sum_{n,m}{}'\alpha_n^i\alpha_m^i\exp\Big(-\frac{\pi i c\tau}{\ell}(n+m)\Big)\sin\frac{n\pi\sigma}{l}\sin\frac{m\pi\sigma}{\ell}\\ &= -\frac{1}{4p^+\alpha'}\sum_{n,m}{}'\alpha_n^i\alpha_m^i\exp\Big(-\frac{\pi i c\tau}{\ell}(n+m)\Big)\big(\delta_{n,m}-\delta_{n,-m}\big)/2 \end{aligned}$$

So:

$$H = H_\Pi + H_X = \frac{p^ip^i}{2p^+} + \frac{1}{4p^+\alpha'}\sum_{n,m}{}'\alpha_n^i\alpha_m^i\exp\Big(-\frac{\pi i c\tau}{\ell}(n+m)\Big)\delta_{n,-m}$$

$$= \frac{p^i p^i}{2p^+} + \frac{1}{4p^+\alpha'} \sum_m{}' \alpha^i_{-m}\alpha^i_m$$

And:

$$\sum_n \alpha^i_{-n}\alpha^i_n = \sum_{n=-\infty}^{-1} \alpha^i_{-n}\alpha^i_n + \sum_{n=1}^{\infty} \alpha^i_{-n}\alpha^i_n = \sum_{n=\infty}^{1} \alpha^i_n\alpha^i_{-n} + \sum_{n=1}^{\infty} \alpha^i_{-n}\alpha^i_n = \sum_{n=1}^{\infty}(\alpha^i_n\alpha^i_{-n} + \alpha^i_{-n}\alpha^i_n)$$
$$= 2\sum_{n=1}^{\infty} \alpha^i_{-n}\alpha^i_n + \sum_{n=1}^{\infty} n\delta^{ii} = 2\sum_{n=1}^{\infty} \alpha^i_{-n}\alpha^i_n + (D-2)\sum_{n=1}^{\infty} n$$

Since $i = 2, 3..., D-1$, thus $A = (D-2)\sum_{n=1}^{\infty} n/2 = (D-2)\zeta(-1)/2$. Q.E.D. □

In this H, the operator ordering is ambiguous. We have placed lowering operators on the right and raising operators on the left. The constant A in Hamiltonian arises from the commutators. In the detailed treatment of light-cone quantization, the constant is determined as follows: the choice of light-cone gauge obstructs the theory's Lorentz invariance. By seeking the generator of Lorentz transformations $M^{\mu\nu}$ and verifying its algebra with p^μ and other operators, only one case is consistent: $A = -1$. The corresponding spacetime dimension is $D = 26$.

In light-cone quantization, we do not wish to dwell too much on this. We will obtain the values of A and D using conformal gauge methods. It will show us why incorrect values of A and D in the light-cone method lose Lorentz invariance. First, we consider the operator ordering constant in the Hamiltonian to originate from the sum of zero-point energies of each oscillator mode. For a bosonic field like X^μ, this is $\omega/2$. The expressions $\frac{1}{2}\omega\left(aa^\dagger + a^\dagger a\right)$ and $\omega\left(a^\dagger a + \frac{1}{2}\right)$ are equivalent. In H:

$$A = \frac{D-2}{2}\sum_{n=1}^{\infty} n\,, \tag{1.3.31}$$

where $D-2$ comes from the sum over transverse directions. This zero-point energy diverges. Using renormalization techniques:

$$\sum_{n=1}^{\infty} n \to -\frac{1}{12}\,. \tag{1.3.32}$$

To obtain this, insert a smooth cutoff factor:

$$\exp\Big(-\epsilon\gamma_{\sigma\sigma}^{-1/2}|k_\sigma|\Big) \tag{1.3.33}$$

where $k_\sigma = n\pi/\ell$. The factor $\gamma_{\sigma\sigma}^{-1/2}$ ensures invariance under σ reparameterization.

$$\begin{aligned} A &\to \frac{D-2}{2}\sum_{n=1}^{\infty} n\exp\left[-\epsilon n\left(\pi/2p^+\alpha'\ell\right)^{1/2}\right] \\ &= \frac{D-2}{2}\left(\frac{2\ell p^+\alpha'}{\epsilon^2\pi} - \frac{1}{12} + O(\epsilon)\right) \end{aligned} \tag{1.3.34}$$

The first term of the cutoff is proportional to the string length ℓ and can be canceled by counterterms in the action proportional to $\int d^2\sigma(-\gamma)^{1/2}$. In fact, Weyl invariance requires it to be canceled, leaving only the cutoff-independent second term:

$$A = \frac{2-D}{24} \tag{1.3.35}$$

The finite residue is an example of Casimir energy, which arises because the string length is finite. For a point particle $p^- = H$, so:

$$m^2 = 2p^+H - p^ip^i = \frac{1}{\alpha'}\left(N + \frac{2-D}{24}\right) \tag{1.3.36}$$

where N is the level:

$$N = \sum_{i=2}^{D-1} \sum_{n=1}^{\infty} n N_{in} \tag{1.3.37}$$

The mass of each state is determined by its excitation level. Consider some light string states. The lightest is:

$$|0;k\rangle, \quad m^2 = \frac{2-D}{24\alpha'} \tag{1.3.38}$$

If $D > 2$, m^2 is negative, and this state is a tachyon. In field theory, the scalar potential is $\frac{1}{2}m^2\phi^2$, so a negative mass squared implies that the "no-string" vacuum is actually unstable, similar to symmetric states in spontaneous symmetry breaking. Whether the bosonic string has any stable vacuum is a complex question, and the answer remains unclear. For superstrings, there are "tachyon-free" theories. We will continue to use the bosonic string as a model for developing string techniques, ignoring this instability.

The lowest excited state of the string is the excitation of the $n = 1$ mode once:

$$\alpha_{-1}^i|0;k\rangle, \quad m^2 = \frac{26-D}{24\alpha'} \tag{1.3.39}$$

Lorentz invariance imposes requirements on D. For massless and massive particles, the spin analysis is different. For massive particles, in the rest frame $p^\mu = (m, 0, \cdots, 0)$. Then the internal states form a representation of the spatial rotation group $SO(D-1)$. For massless particles, there is no rest frame; we choose a frame $p^\mu = (E, E, 0, \cdots, 0)$. Then $SO(D-2)$ acts on the transverse dimensions leaving p^μ invariant, and internal states form a group representation. For $D = 4$, a massive particle is labeled by spin j, an $SO(3)$ representation, so there are $2j-1$ states. A massless particle is labeled by helicity λ, an eigenstate under a single generator of $SO(2)$. Lorentz invariance requires only this one state, but CPT symmetry imposes λ and $-\lambda$. In D dimensions, a massive particle has $D-1$ spin states, while a massless particle has $D-2$ states. In the first level, we only find $D-2$ states $\alpha_{-1}^i|0;k\rangle$, so it must be massless.

$$A = -1, \quad D = 26 \tag{1.3.40}$$

This is a striking and important result. Only when the spacetime dimension $D = 26$ is the spectrum Lorentz-invariant. The classical theory is Lorentz-invariant for any D, but there is an anomaly: except for $D = 26$, quantization does not preserve Lorentz invariance.

Light-cone quantization picks out two directions, leaving $SO(D-2)$ acting on the transverse dimensions. For the transverse directions, the spin generator is:

$$S^{ij} = -i\sum_{n=1}^{\infty} \frac{1}{n}(\alpha_{-n}^i \alpha_n^j - \alpha_{-n}^j \alpha_n^i) \tag{1.3.41}$$

It is antisymmetric in i, j, and together with Lorentz invariance, forbids zero-point energy constants for massless states.

Proof. S^{ij} is defined as $S^{ij} = \int_0^\ell \mathrm{d}\sigma\, X^i\Pi^j - X^j\Pi^i$. Note that at this point $p_i = 0$, so:

$$X^i\Pi^j - X^j\Pi^i = \frac{i}{\ell}\sum_{n,m}{}' \frac{1}{n}\alpha_n^i\alpha_m^j \exp\Big(-\frac{\pi i c\tau}{\ell}(n+m)\Big)\cos\frac{n\pi\sigma}{\ell}\cos\frac{m\pi\sigma}{\ell} - (i \leftrightarrow j).$$

From this, it follows:

$$S^{ij} = i\sum_{n,m}{}' \frac{1}{n}\alpha_n^i\alpha_m^j \exp\Big(-\frac{\pi i c\tau}{\ell}(n+m)\Big)\big(\delta_{n,m} + \delta_{n,-m}\big)/2 - (i \leftrightarrow j)$$

$$= i\sum_n{}' \frac{1}{n}\alpha_n^i\alpha_{-n}^j/2 - (i \leftrightarrow j)$$

Q.E.D. □

For a massless particle, by choosing the momentum direction to be the "1" direction selected by quantization, the entire $SO(D-2)$ spin symmetry is made manifest. For massive particles, only an $SO(D-2)$ subgroup is obvious within the $SO(D-1)$ symmetry of light-cone quantization. But this is still useful. Under the action of $SO(D-2)$ on the transverse directions, the vector representation of $SO(D-1)$ breaks into an invariant and a $(D-2)$-vector:

$$\mathbf{v} = \left(v^1, 0, \ldots, 0\right) + \left(0, v^2, \ldots, v^{D-1}\right) \tag{1.3.42}$$

Thus, if a massive particle is in the vector representation of $SO(D-1)$, when we examine its transformation properties under $SO(D-2)$, we will see a scalar and a vector. Higher excited states of the string are massive, forming full representations of $SO(D-1)$. At level N, the maximum eigenvalue of the spin component S^{23} is N, obtained by acting N times with $\alpha_{-1}^2 + i\alpha_{-1}^3$. Therefore:

$$S^{23} \leq 1 + \alpha' m^2 \tag{1.3.43}$$

We start by checking the Eigenvalue of S^{23} for the state $(\alpha_{-m}^2 + i\alpha_{-m}^3)|0;k\rangle$ for a given m:

$$\begin{aligned}
S^{23}(\alpha_{-m}^2 + i\alpha_{-m}^3)|0;k\rangle &= -i\sum_{n=1}^{\infty}\frac{1}{n}(\alpha_{-n}^2\alpha_n^3 - \alpha_{-n}^3\alpha_n^2)(\alpha_{-m}^2 + i\alpha_{-m}^3)|0;k\rangle \\
&= -\frac{i}{m}(\alpha_{-m}^2\alpha_m^3 - \alpha_{-m}^3\alpha_m^2)(\alpha_{-m}^2 + i\alpha_{-m}^3)|0;k\rangle \\
&= -\frac{i}{m}(\alpha_{-m}^2\alpha_m^3 i\alpha_{-m}^3 - \alpha_{-m}^3\alpha_m^2\alpha_{-m}^2)|0;k\rangle \\
&= (\alpha_{-m}^2 + i\alpha_{-m}^3)|0;k\rangle
\end{aligned} \tag{1.53}$$

where we have used $[\alpha_m^i, \alpha_n^j] = m\delta^{ij}\delta_{m+n,0}$ and $\alpha_m^i|0\rangle = 0$. So this state has spin one. Consider now

$$\begin{aligned}
S^{23}(\alpha_{-m}^2 + i\alpha_{-m}^3)^2|0;k\rangle &= -i\sum_{n=1}^{\infty}\frac{1}{n}(\alpha_{-n}^2\alpha_n^3 - \alpha_{-n}^3\alpha_n^2)(\alpha_{-m}^2 + i\alpha_{-m}^3)^2|0;k\rangle \\
&= -\frac{i}{m}(\alpha_{-m}^2\alpha_m^3 - \alpha_{-m}^3\alpha_m^2)(\alpha_{-m}^2 + i\alpha_{-m}^3)^2|0;k\rangle \\
&= -\frac{i}{m}(\alpha_{-m}^2\alpha_m^3 - \alpha_{-m}^3\alpha_m^2)(\alpha_{-m}^2\alpha_{-m}^2 + i\alpha_{-m}^2\alpha_{-m}^3 + i\alpha_{-m}^3\alpha_{-m}^2 \\
&\quad - \alpha_{-m}^3\alpha_{-m}^3)|0;k\rangle \\
&= 2(\alpha_{-m}^2\alpha_{-m}^2 + i\alpha_{-m}^2\alpha_{-m}^3 + i\alpha_{-m}^3\alpha_{-m}^2 - \alpha_{-m}^3\alpha_{-m}^3)|0;k\rangle \\
&= 2(\alpha_{-m}^2 + i\alpha_{-m}^3)^2|0;k\rangle
\end{aligned} \tag{1.54}$$

Similarly we find

$$S^{23}(\alpha_{-m}^2 + i\alpha_{-m}^3)^N|0;k\rangle = N(\alpha_{-m}^2 + i\alpha_{-m}^3)^N|0;k\rangle \tag{1.55}$$

Any other state at level N will contain at least one less factor of $\alpha^2_{-m} + i\alpha^3_{-m}$ and as S^{23} only has non-zero Eigenvalue on that specific combination, it will give a spin lower than N. Using the mass-shell condition (1.3.36) on $D = 26$, i.e. $\alpha' m^2 = N - 1$ we thus find

$$S^{23} \leq N = \alpha' m^2 + 1 \tag{1.56}$$

QED.

The slope in the inequality is called the Regge slope. Meson resonances follow this form of linear relationship, i.e., $\alpha' \sim (1\text{GeV})^{-2}$. For this reason, string theory was originally a theory of strong interactions in the 1970s. Now, as a theory of quantum gravity, α' is on the order of M_P^{-2}, and the mass of massive particles is on the order of M_P. It is so large that at current experimental levels, these particles only appear as virtual states. Therefore, we are particularly interested in the massless string spectrum. Most known particles have mass, but it is so small compared to M_P that it is zero to first approximation, becoming non-zero due to small symmetry breaking.

1.4 Closed and unoriented strings

Light-cone quantization of the closed string is very similar to the open string. Again, impose gauge conditions (1.3.8a)-(1.3.8c). For the open string, these completely determine the gauge. For the closed string, some additional coordinate freedom remains:

$$\sigma' = \sigma + s(\tau) \bmod \ell \tag{1.4.1}$$

Since $\sigma = 0$ can be chosen as any point on the string, most of this residual freedom can be fixed by an extra gauge condition:

$$\gamma_{\tau\sigma}(\tau, 0) = 0 \tag{1.4.2}$$

i.e., the line $\sigma = 0$ is perpendicular to the lines $\tau = C$ (where C is a constant). Except for a global translation, this completely determines the line $\sigma = 0$. Thus, (1.3.8) and (1.4.2), except for a τ-independent translation of σ, fix all gauge freedom.

$$\sigma' = \sigma + s \bmod \ell \tag{1.4.3}$$

The analysis now parallels the open string: Lagrangian, canonical momentum, Hamiltonian, equations of motion, etc. General periodic solutions to the equations of motion:

$$\begin{aligned} X^i(\tau, \sigma) = {} & x^i + \frac{p^i}{p^+}\tau + i\left(\frac{\alpha'}{2}\right)^{1/2} \\ & \times \sum_{\substack{n=-\infty \\ n\neq 0}}^{\infty} \left\{ \frac{\alpha_n^i}{n} \exp\left[-\frac{2\pi i n(\sigma + c\tau)}{\ell}\right] + \frac{\tilde{\alpha}_n^i}{n} \exp\left[\frac{2\pi i n(\sigma - c\tau)}{\ell}\right] \right\} \end{aligned} \tag{1.4.4}$$

In the closed string, there are two sets of independent oscillatory solutions $\alpha_n^i, \tilde{\alpha}_n^i$, corresponding to waves moving left and right along the string. In the open string, boundary conditions at the ends mix these. The independent degrees of freedom are again transverse oscillations and center-of-mass variables for both transverse and longitudinal directions:

$$\alpha_n^i, \tilde{\alpha}_n^i, x^i, p^i, x^-, p^+ \tag{1.4.5}$$

with canonical commutators:

$$\left[x^-, p^+\right] = -i \tag{1.4.6a}$$

$$\left[x^i, p^j\right] = i\delta^{ij} \tag{1.4.6b}$$
$$\left[\alpha_m^i, \alpha_n^j\right] = m\delta^{ij}\delta_{m,-n} \tag{1.4.6c}$$
$$\left[\tilde{\alpha}_m^i, \tilde{\alpha}_n^j\right] = m\delta^{ij}\delta_{m,-n} \tag{1.4.6d}$$

Starting from the state $|0,0;k\rangle$, which has center-of-mass momentum k^μ and is annihilated by $\alpha_n^i, \tilde{\alpha}_n^i$ for $m > 0$. General states are:

$$|N, \tilde{N}; k\rangle = \left[\prod_{i=2}^{D-1}\prod_{n=1}^{\infty} \frac{(\alpha_{-n}^i)^{N_{in}}(\tilde{\alpha}_{-n}^i)^{\tilde{N}_{in}}}{(n^{N_{in}} N_{in}! n^{\tilde{N}_{in}} \tilde{N}_{in}!)^{1/2}}\right] |0,0;k\rangle \tag{1.4.7}$$

The mass formula is:

$$\begin{aligned} m^2 &= 2p^+ H - p^i p^i \\ &= \frac{2}{\alpha'}\left[\sum_{n=1}^{\infty}\left(\alpha_{-n}^i \alpha_n^i + \tilde{\alpha}_{-n}^i \tilde{\alpha}_n^i\right) + A + \tilde{A}\right] \\ &= \frac{2}{\alpha'}(N + \tilde{N} + A + \tilde{A}) \end{aligned} \tag{1.4.8}$$

We split levels and zero-point energies into two parts: left-moving and right-moving. Summing zero-point energies gives:

$$A = \tilde{A} = \frac{2-D}{24} \tag{1.4.9}$$

Due to the residual gauge freedom, i.e., the σ translation in (1.4.3), there is an additional constraint. The operator generating σ translations is:

$$\begin{aligned} P &= -\int_0^\ell \mathrm{d}\sigma\, \Pi^i \partial_\sigma X^i \\ &= -\frac{2\pi}{\ell}\left[\sum_{n=1}^{\infty}\left(\alpha_{-n}^i \alpha_n^i - \tilde{\alpha}_{-n}^i \tilde{\alpha}_n^i\right) + A - \tilde{A}\right] \\ &= -\frac{2\pi}{\ell}(N - \tilde{N}) \end{aligned} \tag{1.4.10}$$

States must satisfy:

$$N = \tilde{N} \tag{1.4.11}$$

The lightest closed string state:

$$|0,0;k\rangle, \quad m^2 = \frac{2-D}{6\alpha'} \tag{1.4.12}$$

is also a tachyon. The first excited state:

$$\alpha_{-1}^i \tilde{\alpha}_{-1}^j |0,0;k\rangle, \quad m^2 = \frac{26-D}{6\alpha'} \tag{1.4.13}$$

As with the open string, these states do not add up to a full representation of $SO(D-1)$, so the level must be massless:

$$A = \tilde{A} = -1, \quad D = 26 \tag{1.4.14}$$

State (1.4.13) transforms as a 2nd-rank tensor under $SO(D-2)$. This is a reducible representation. It can be decomposed into a symmetric traceless tensor, an antisymmetric tensor, and a scalar. That is, any tensor e^{ij} can be split into:

$$e^{ij} = \frac{1}{2}\left(e^{ij} + e^{ji} - \frac{2}{D-2}\delta^{ij}e^{kk}\right) + \frac{1}{2}\left(e^{ij} - e^{ji}\right) + \frac{1}{D-2}\delta^{ij}e^{kk} \tag{1.4.15}$$

These three separate terms do not mix under rotation.

Remark.

(1.4.15) =
$$\frac{1}{2}\left(\alpha^i_{-1}\tilde{\alpha}^j_{-1}+\tilde{\alpha}^j_{-1}\tilde{\alpha}^i_{-1}-\frac{2}{D-2}\delta^{ij}\alpha^k_{-1}\tilde{\alpha}^k_{-1}\right)+\frac{1}{2}\left(\alpha^i_{-1}\tilde{\alpha}^j_{-1}-\alpha^j_{-1}\tilde{\alpha}^i_{-1}\right)+\frac{1}{D-2}\delta^{ij}\alpha^k_{-1}\tilde{\alpha}^k_{-1}$$

At any energy level, N_{in} and $\tilde{N}_{in}$ are independent except for the $N=\tilde{N}$ constraint. Therefore, the closed string spectrum at $m^2 = 4(N-1)/\alpha'$ is the product of two pairs of open string levels $m^2 = (N-1)/\alpha'$.

Light cone quantization was gauge fixed version, even thought $\Pi^0 \neq 0$, but $T^{ab}=0$ which implies Π^0 and Π^i are not independent. Since X^0 comes with wrong sign, we need to eliminate Π^0. 2D diff invariance removes two classes of normal modes from the spectrum. If we attempted to build a covariant theory without this invariance, we would have to generalize the transverse commutators (1.3.25) to:
$$[\alpha^\mu_m, \alpha^\nu_n] = m\eta^{\mu\nu}\delta_{m,-n} \tag{1.4.16}$$
Lorentz invariance forces the timelike oscillators to have a wrong-sign commutator.

Remark. *$[\alpha^\mu_m, \alpha^\nu_n] = m\eta^{\mu\nu}\delta_{m,-n}$, so $\left[\alpha^0_m, \alpha^0_{-m}\right] = -m$, and $\alpha^0_{-m} = {\alpha^0_m}^\dagger$. Because $\langle 0|\alpha^0_m\alpha^0_{-m}|0\rangle = -m\langle 0|0\rangle < 0$, $\alpha^0_m|0\rangle$ is a negative-norm state.*

States with an odd number of timelike excitations will have a negative norm. This is inconsistent with quantum mechanics. In fact, the theory of (1.4.16) predates the string theory we describe, which was discovered by requiring that negative-norm states cannot be produced in physical processes. The commutator (1.4.16) arises from covariant quantization of string theory, where coordinate invariance appears as a constraint to eliminate negative-norm states.

The appearance of massless vectors in open string theory, and massless symmetric and antisymmetric tensors in closed string theory, is significant. General principles require a massless vector to couple to a conserved current, thus the theory has gauge invariance. In all fundamental string theories, this is our first example of a massless gauge particle. Actually, this particular gauge boson, called the photon, is not very interesting because the gauge group is only $U(1)$, and all particles in the theory are neutral. In Chapter 6, we will discuss a simple generalization of open string theory, adding Chan–Paton degrees of freedom at the string endpoints. This leads to $U(n)$, $SO(n)$, and $Sp(n)$ gauge groups. Similarly, massless symmetric tensor particles must couple to a conserved symmetric tensor source. The only such source is the energy-momentum tensor (with other possibilities in special cases; coupling consistent with it requires the theory to have spacetime coordinate invariance). Therefore, the massless symmetric tensor is the graviton, and general relativity is included as a small part of closed string theory. The massless antisymmetric tensor is called a 2-form gauge boson. In section 3.7, we will see that there is a local spacetime symmetry associated with it. In 4D, massless string states excited by 2nd and 3rd direction oscillators are identified by their helicity $\lambda = S^{23}$. The photon has $\lambda = 1$, the graviton $\lambda = 2$. Closed string scalars and antisymmetric tensors give $\lambda = 0$ states, called the dilaton and axion, respectively.

We conclude this section with unoriented string theory. Everything discussed previously was oriented string theory. we have not considered the coordinate transformation:
$$\sigma' = \ell - \sigma, \quad \tau' = \tau \tag{1.4.17}$$
This changes the orientation (chirality) of the worldsheet. This symmetry is generated by the worldsheet parity operator Ω. Performing (1.4.17) twice gives the identity, so $\Omega^2 = 1$, and eigenvalues of Ω are ± 1. From the mode expansions (1.3.22) and (1.4.4), we see that in the open string:
$$\Omega\alpha^i_n\Omega^{-1} = (-1)^n\alpha^i_n \tag{1.4.18}$$

In the closed string:

$$\Omega\alpha_n^i\Omega^{-1} = \tilde{\alpha}_n^i \tag{1.4.19a}$$

$$\Omega\tilde{\alpha}_n^i\Omega^{-1} = \alpha_n^i \tag{1.4.19b}$$

By fixing $\Omega = +1$ for the ground states $|0;k\rangle, |0,0;k\rangle$, we define the phase of Ω. Later, we will see this choice is made so that Ω is conserved under interactions. Thus:

$$\Omega|N;k\rangle = (-1)^N|N;k\rangle \tag{1.4.20a}$$

$$\Omega|N,\tilde{N};k\rangle = |\tilde{N},N;k\rangle. \tag{1.4.20b}$$

There is a consistent interacting string theory, the unoriented string theory, in which only states with $\Omega = +1$ are retained. Focusing on massless states, the open string photon has $\Omega = -1$ and is thus forbidden in the unoriented theory. Acting on massless tensor states in the closed string, the parity operator Ω changes e^{ij} into e^{ji}. Thus gravitons and dilatons appear in unoriented theory, but the antisymmetric tensor does not. Note that both open and closed string tachyons survive in unoriented theory, so we have to work harder to remove them.

Two other constraints arise when studying interactions: First, it is possible to have consistent theories with only closed strings, or both closed and open strings, but not with only open strings. Closed strings can always be produced in open string scattering. Second, oriented or unoriented open strings can only couple with the same type of closed strings. We list possible combinations and their massless spectra. Let $G_{\mu\nu}, B_{\mu\nu}, \Phi, A_\mu$ denote the graviton, antisymmetric tensor, dilaton, and photon. Then:

1. Oriented bosonic closed string: $G_{\mu\nu}, B_{\mu\nu}, \Phi$.

2. Unoriented bosonic closed string: $G_{\mu\nu}, \Phi$.

3. Oriented bosonic closed and open strings: $G_{\mu\nu}, B_{\mu\nu}, \Phi, A_\mu$.

4. Unoriented bosonic closed and open strings: $G_{\mu\nu}, \Phi$.

All of these have gravitons, dilatons, and tachyons. Massless oriented open strings will be $U(n)$ gauge bosons, while massless unoriented open strings will be $SO(n)$ or $Sp(n)$ gauge bosons.

Chapter 2

Conformal field theory

2.1 Massless scalars in two dimensions

We start from D free scalar fields $X^\mu(\sigma^1, \sigma^2)$ in two dimensions. In anticipation of application to string theory, we call this two-dimensional space the worldsheet. The action is

$$S = \frac{1}{4\pi\alpha'} \int \mathrm{d}^2\,\sigma \Big(\partial_1 X^\mu \partial_1 X_\mu + \partial_2 X^\mu \partial_2 X_\mu \Big)\,. \tag{2.1.1}$$

This is the Polyakov action (1.2.13) where the worldsheet metric γ_{ab} is taken to be the flat Euclidean metric δ_{ab}, with signature $(+,+)$. Most string theory calculations will be performed on the Euclidean worldsheet. At least for flat metrics, the relationship between Euclidean amplitudes and Minkowski amplitudes is given by standard analytic continuation. We will use the Minkowski flat metric for the index μ.

Canonical quantization of the action (2.1.1) is straightforward: finding the spectrum, vacuum expectation values, and so on. What we did in Chapter 1 was to complete these in the light-cone gauge. Here we will take a different route, first studying various local properties, such as equations of motion, operator products, Ward identities, and conformal invariance, and only then the spectrum. We will use the more effective path integral formalism. We will primarily use path integral representations to derive operator equations; these equations can also be obtained in the Hilbert space formalism. Adopting complex coordinates

$$z = \sigma^1 + i\sigma^2\,, \qquad \bar{z} = \sigma^1 - i\sigma^2 \tag{2.1.2}$$

is very convenient. We will use a bar to represent the conjugate of z and other complex variables, and for longer expressions, an asterisk. Also define

$$\partial_z = \frac{1}{2}(\partial_1 - i\partial_2)\,, \qquad \partial_{\bar{z}} = \frac{1}{2}(\partial_1 + i\partial_2)\,. \tag{2.1.3}$$

These variables satisfy

$$\partial_z z = 1\,, \quad \partial_z \bar{z} = 0\,, \quad \partial_{\bar{z}} z = 0\,, \quad \partial_{\bar{z}} \bar{z} = 1\,. \tag{2.1.4}$$

When there is no ambiguity, ∂_z is abbreviated as ∂ and $\partial_{\bar{z}}$ as $\bar{\partial}$. For a general vector v^a, define in the same way

$$v^z = v^1 + iv^2\,, \quad v^{\bar{z}} = v^1 - iv^2\,, \quad v_z = \frac{1}{2}(v^1 - iv^2)\,, \quad v_{\bar{z}} = \frac{1}{2}(v^1 + iv^2)\,. \tag{2.1.5}$$

For complex indices, raising and lowering indices is accomplished via

$$g_{z\bar{z}} = g_{\bar{z}z} = \frac{1}{2}\,, \quad g_{zz} = g_{\bar{z}\bar{z}} = 0\,, \quad g^{z\bar{z}} = g_{\bar{z}z} = 2\,, \quad g^{zz} = g^{\bar{z}\bar{z}} = 0 \tag{2.1.6}$$

In matrix form:

$$g_{\mu\nu} = \begin{array}{c c} & \begin{array}{cc} z & \bar{z} \end{array} \\ \begin{array}{c} z \\ \bar{z} \end{array} & \begin{pmatrix} 0 & 1/2 \\ 1/2 & 0 \end{pmatrix} \end{array} \tag{2.1.7}$$

Also note

$$\mathrm{d}^2 z = 2\mathrm{d}\sigma^1 \mathrm{d}\sigma^2 \tag{2.1.8}$$

The factor 2 comes from the Jacobian, $\mathrm{d}^2 z\, |\det g|^{1/2} = \mathrm{d}\sigma^1 \mathrm{d}\sigma^2$. We define

$$\int \mathrm{d}^2 z\, \delta^2(z, \bar{z}) = 1\,, \tag{2.1.9}$$

so that $\delta^2(z,\bar{z}) = \frac{1}{2}\delta(\sigma^1)\delta(\sigma^2)$. The divergence formula in complex coordinates is

$$\int_R \mathrm{d}^2 z\, (\partial_z v^z + \partial_{\bar{z}} v^{\bar{z}}) = i \oint_{\partial R} (v^z \mathrm{d}\bar{z} - v^{\bar{z}} \mathrm{d}z)\,, \tag{2.1.10}$$

where the integration contour runs counterclockwise around the boundary of the region R. In this sign convention, the action is

$$S = \frac{1}{2\pi\alpha'} \int 2\mathrm{d}^2\sigma \frac{\partial_1 X^\mu \partial_1 X_\mu + \partial_2 X^\mu \partial_2 X_\mu}{4} = \frac{1}{2\pi\alpha'} \int \mathrm{d}^2 z\, \partial X^\mu \bar{\partial} X_\mu\,, \tag{2.1.11}$$

The classical equations of motion are

$$\frac{\delta S}{\delta X_\mu(z,\bar{z})} = 0 \implies \partial\bar{\partial} X^\mu(z,\bar{z}) = 0\,. \tag{2.1.12}$$

Here $f(z)$ refers to those functions that are analytic (equivalently, holomorphic) in z. Writing the equations of motion as

$$\partial(\bar{\partial} X^\mu) = \bar{\partial}(\partial X^\mu) = 0\,, \tag{2.1.13}$$

it follows that ∂X^μ is holomorphic and $\bar{\partial} X^\mu$ is anti-holomorphic (holomorphic in $\bar{z}$), thus we have $\partial X^\mu(z)$ and $\bar{\partial} X^\mu(\bar{z})$.

Under the Minkowski continuation $\sigma^2 = i\sigma^0$, the holomorphic field becomes a function of $\sigma^0 - \sigma^1$, while the anti-holomorphic field becomes a function of $\sigma^0 + \sigma^1$. Thus there is the following correspondence:

$$\text{holomorphic} = \text{left-moving}\,, \tag{2.1.14a}$$

$$\text{anti-holomorphic} = \text{right-moving}\,. \tag{2.1.14b}$$

Vacuum expectation values are defined by path integrals,

$$\langle \mathscr{F}[X] \rangle = \int [\mathrm{d}X]\, \exp(-S) \mathscr{F}[X]\,, \tag{2.1.15}$$

where $\mathscr{F}[X]$ is any functional of X, for example a product of local operators. In order for the path integral over X^0 to be a convergent Gaussian integral, it is defined through analytic continuation $X^0 \to -iX^D$. We do not normalize $\langle \mathscr{F}[X] \rangle$ by dividing by $\langle 1 \rangle$. The path integral of a total derivative is zero. Then,

$$\begin{aligned} 0 &= \int [\mathrm{d}X]\, \frac{\delta}{\delta X_\mu(z,\bar{z})} \exp(-S) \\ &= -\int [\mathrm{d}X]\, \exp(-S) \frac{\delta S}{\delta X_\mu(z,\bar{z})} \\ &= -\left\langle \frac{\delta S}{\delta X_\mu(z,\bar{z})} \right\rangle \qquad \text{(can be simplified using (2.1.12))} \end{aligned}$$

$$= \frac{1}{\pi\alpha'}\Big\langle \partial\bar{\partial}X^\mu(z,\bar{z})\Big\rangle\,. \tag{2.1.16}$$

If we have additional arbitrary insertions "$\cdots$" in the path integral, the calculation is the same as long as those additional operators are not inserted at z. Therefore

$$\Big\langle \partial\bar{\partial}X^\mu(z,\bar{z})\cdots\Big\rangle = 0\,. \tag{2.1.17}$$

We can view the additional insertions as preparing arbitrary initial and final states (we can do the same thing for situations with boundary conditions). The path integral formula (2.1.17) thus corresponds to the operator equation

$$\partial\bar{\partial}\hat{X}^\mu(z,\bar{z}) = 0 \tag{2.1.18}$$

which holds for all matrix elements of the operator $\hat{X}^\mu(z,\bar{z})$ in the Hilbert space formalism. Therefore we call the relationship established in the sense of (2.1.17) an operator equation. The equation (2.1.18) is the Ehrenfest theorem translating the classical equations of motion into operator equations.

The notation "$\cdots$" in the path integral (2.1.17) implies that the insertions are away from z. Now consider the case where the insertion is exactly at z:

$$\begin{aligned} 0 &= \int [\mathrm{d}X]\, \frac{\delta}{\delta X_\mu(z,\bar{z})}\Big[\exp(-S)X^v(z',\bar{z}')\Big] \\ &= \int [\mathrm{d}X]\, \exp(-S)\, \left[\eta^{\mu v}\delta^2(z-z',\bar{z}-\bar{z}') + \frac{1}{\pi\alpha'}\partial_z\partial_{\bar{z}}X^\mu(z,\bar{z})X^v(z',\bar{z}')\right] \\ &= \eta^{\mu v}\left\langle \delta^2(z-z',\bar{z}-\bar{z}')\right\rangle + \frac{1}{\pi\alpha'}\partial_z\partial_{\bar{z}}\left\langle X^\mu(z,\bar{z})X^v(z',\bar{z}')\right\rangle \end{aligned} \tag{2.1.19}$$

That is, the equation of motion holds outside of the coincidence point. Inserting "$\cdots$" in the path integral once more:

$$\frac{1}{\pi\alpha'}\partial_z\partial_{\bar{z}}\left\langle X^\mu(z,\bar{z})X^v(z',\bar{z}')\cdots\right\rangle = -\eta^{\mu\nu}\left\langle \delta^2(z-z',\bar{z}-\bar{z}')\cdots\right\rangle \tag{2.1.20}$$

thus

$$\frac{1}{\pi\alpha'}\partial_z\partial_{\bar{z}}X^\mu(z,\bar{z})X^v(z',\bar{z}') = -\eta^{\mu v}\delta^2(z-z',\bar{z}-\bar{z}') \tag{2.1.21}$$

holds as an operator equation. In the Hilbert space formalism, the products in the path integral appear as time-ordered products, and the δ-function comes from derivatives acting on the time-ordered product. This connection is further developed in the appendix.

In free field theory, it is useful to introduce normal-ordered operators. The normal-ordering of an operator is denoted as $:\!\mathscr{A}\!:$, defined as follows:

$$:\!X^\mu(z,\bar{z})\!: = X^\mu(z,\bar{z})\ :, \tag{2.1.22a}$$

$$:\!X^\mu(z_1,\bar{z}_1)X^v(z_2,\bar{z}_2)\!: = X^\mu\left(z_1,\bar{z}_1\right)X^v(z_2,\bar{z}_2) + \frac{\alpha'}{2}\eta^{\mu v}\ln|z_{12}|^2\,, \tag{2.1.22b}$$

where

$$z_{ij} = z_i - z_j \tag{2.1.23}$$

The mode function for closed string (2.7.4) with $w = \sigma^1 + i\sigma^2$ (refer (2.6.3)) can be given as:

$$X^\mu(w,\overline{w}) = x^\mu - i\frac{\alpha'}{2}p^\mu(w+\overline{w}) + i\sqrt{\frac{\alpha'}{2}}\sum_{n\neq 0}\frac{1}{n}\left(\alpha_n^\mu e^{-2\pi n w/L} + \overline{\alpha}_n^\mu e^{-2\pi n\overline{w})/L}\right), \tag{2.1.24}$$

with now $w \sim w + iL$. Using a conformal transformation, we map all the operators from the cylinder to the complex plane:

$$z = e^{2\pi w/L}, \qquad \overline{z} = e^{2\pi \overline{w}/L}$$

we finally obtain the expansion by simply replacing w by z.

$$X^\mu(z,\overline{z}) = x^\mu - i\frac{\alpha'}{2}p^\mu \ln(z\overline{z}) + i\sqrt{\frac{\alpha'}{2}}\sum_{n\neq 0}\frac{1}{n}(\alpha_n^\mu z^{-n} + \overline{\alpha}_n^\mu \overline{z}^{-n})\,. \tag{2.1.25}$$

The propagator could be calculated by assuming $|z| > |z'|$ and considering the product

$$\begin{aligned} X^\mu(w,\bar{w})\,X^\mu(w',\bar{w}') \;=\; & \Big[x^\mu - i\frac{\alpha'}{2}p^\mu \ln|z|^2 + i\sqrt{\frac{\alpha'}{2}}\sum_{m\neq 0}\frac{1}{m}(\alpha_m^\mu z^{-m} + \overline{\alpha}_m^\mu \overline{z}^{-m})\Big] \\ & \times \Big[x^\mu - i\frac{\alpha'}{2}p^\mu \ln|z'|^2 + i\sqrt{\frac{\alpha'}{2}}\sum_{n\neq 0}\frac{1}{n}(\alpha_n^\mu z'^{-n} + \overline{\alpha}_n^\mu \overline{z}'^{-n})\Big]. \end{aligned}$$

We will move annihilation operators ($a_k, \overline{a}_k$ with $k > 0$) and π_0 to the right to reach normal ordering; non-vanishing c-number contributions come from $[x^\mu, p^\nu] = i\eta^{\mu\nu}$ and $[\alpha_m^\mu, \alpha_n^\nu] = m\eta^{\mu\nu}\delta_{m+n,0}$, $[\overline{\alpha}_m^\mu, \overline{\alpha}_n^\mu] = m\eta^{\mu\nu}\delta_{m+n,0}$. Let's first focus on these terms

$$\begin{aligned} &\Big(x^\mu - i\frac{\alpha'}{2}p^\mu \ln|z|^2\Big)\Big(x^\nu - i\frac{\alpha'}{2}p^\nu \ln|z'|^2\Big) \\ &\quad = x^\mu x^\nu - i\frac{\alpha'}{2}x^\mu p^\nu \ln|z'|^2 - i\frac{\alpha'}{2}p^\mu x^\nu \ln|z|^2 - \Big(\frac{\alpha'}{2}\Big)^2 p^\mu p^\nu \ln|z|^2 \ln|z'|^2 \\ &\quad = x^\mu x^\nu - i\frac{\alpha'}{2}x^\mu p^\nu \ln|z'|^2 - i\frac{\alpha'}{2}\big(x^\nu p^\mu - i\eta^{\mu\nu}\big) \ln|z|^2 - \Big(\frac{\alpha'}{2}\Big)^2 p^\mu p^\nu \ln|z|^2 \ln|z'|^2 \\ &\quad = x^\mu x^\nu - i\frac{\alpha'}{2}x^\mu p^\nu \ln|z'|^2 - i\frac{\alpha'}{2}x^\nu p^\mu \ln|z|^2 - \Big(\frac{\alpha'}{2}\Big)^2 p^\mu p^\nu \ln|z|^2 \ln|z'|^2 - \frac{\alpha'}{2}\eta^{\mu\nu}\ln|z|^2 \\ &\quad = \underbrace{x^\mu x^\nu - i\frac{\alpha'}{2}x^\mu p^\nu \ln|z'|^2 - i\frac{\alpha'}{2}x^\nu p^\mu \ln|z|^2 - \Big(\frac{\alpha'}{2}\Big)^2 p^\mu p^\nu \ln|z|^2 \ln|z'|^2}_{\text{normal ordered}} - \frac{\alpha'}{2}\eta^{\mu\nu}\ln|z|^2 \\ &\quad = :\Big(x^\mu - i\frac{\alpha'}{2}p^\mu \ln|z|^2\Big)\Big(x^\nu - i\frac{\alpha'}{2}p^\nu \ln|z'|^2\Big): \; - \frac{\alpha'}{2}\eta^{\mu\nu}\ln|z|^2. \end{aligned}$$

The zero-mode reordering produces the c-number $-\frac{\alpha'}{2}\eta^{\mu\nu}\ln|z|^2$. Next, we'll consider holomorphic oscillator pieces

$$i\sqrt{\frac{\alpha'}{2}}\sum_{m\neq 0}\frac{1}{m}\alpha_m^\mu z^{-m} \quad \text{and} \quad i\sqrt{\frac{\alpha'}{2}}\sum_{n\neq 0}\frac{1}{n}\alpha_n^\mu z'^{-n}.$$

We can split the sum in two parts,

$$\sum_{m\neq 0}\frac{1}{m}\alpha_m^\mu z^{-m} = \sum_{m=-\infty}^{-1}\frac{1}{m}\alpha_m^\mu z^{-m} + \sum_{m=1}^{\infty}\frac{1}{m}\alpha_m^\mu z^{-m}.$$

The only non-normal-ordered combination producing a c-number is the product of the $m > 0$ part from the first bracket with the $n < 0$ part from the second bracket. That term is

$$\left(i\sqrt{\frac{\alpha'}{2}}\right)^2\Big(\sum_{m=1}^{\infty}\frac{1}{m}\alpha_m^\mu z^{-m}\Big)\Big(\sum_{n=-\infty}^{-1}\frac{1}{n}\alpha_n^\nu z'^{-n}\Big) = -\frac{\alpha'}{2}\sum_{m=1}^{\infty}\sum_{n=-\infty}^{-1}\frac{1}{mn}\,\alpha_m^\mu\alpha_n^\nu\,z^{-m}z'^{-n}.$$

Using $\alpha_m^\mu\alpha_n^\nu =: \alpha_n^\mu\alpha_m^\nu : +[\alpha_m^\mu,\alpha_n^\nu] =: \alpha_n^\mu\alpha_m^\nu : +m\eta^{\mu\nu}\delta_{m+n,0}$. The only contributing commutator in the sum over n occurs when $n=-m$. Isolating that, we have:

$$-\frac{\alpha'}{2}\sum_{m=1}^{\infty}\frac{1}{m(-m)}\,m\,z^{-m}z'^{m} = -\frac{\alpha'}{2}\sum_{m=1}^{\infty}\frac{-1}{m}\Big(\frac{z'}{z}\Big)^m = \frac{\alpha'}{2}\sum_{m=1}^{\infty}\frac{1}{m}\Big(\frac{z'}{z}\Big)^m.$$

Thus the holomorphic oscillators give the c-number

$$\boxed{\frac{\alpha'}{2}\sum_{m=1}^{\infty}\frac{1}{m}\Big(\frac{z'}{z}\Big)^m}.$$

By identical steps for $\overline{a}_m$ modes we obtain

$$\boxed{\frac{\alpha'}{2}\sum_{m=1}^{\infty}\frac{1}{m}\Big(\frac{\bar{z}'}{\bar{z}}\Big)^m}.$$

Collecting the zero-mode, holomorphic and anti-holomorphic c-number contributions:

$$-\frac{\alpha'}{2}\ln|z|^2 + \frac{\alpha'}{2}\sum_{m=1}^{\infty}\frac{1}{m}\Big(\frac{z'}{z}\Big)^m + \frac{\alpha'}{2}\sum_{m=1}^{\infty}\frac{1}{m}\Big(\frac{\bar{z}'}{\bar{z}}\Big)^m.$$

We can use $-\ln(1-x)=\sum_{m=1}^{\infty}x^m/m$ to simplify the series expansion (valid for $|z'|<|z|$):

$$\sum_{m=1}^{\infty}\frac{1}{m}\Big(\frac{z'}{z}\Big)^m = -\ln\Big(1-\frac{z'}{z}\Big),\qquad \sum_{m=1}^{\infty}\frac{1}{m}\Big(\frac{\bar{z}'}{\bar{z}}\Big)^m = -\ln\Big(1-\frac{\bar{z}'}{\bar{z}}\Big).$$

So the total c-number becomes

$$-\frac{\alpha'}{2}\ln|z|^2 - \frac{\alpha'}{2}\ln\Big|1-\frac{z'}{z}\Big|^2 = -\frac{\alpha'}{2}\ln\Big|z\Big(1-\frac{z'}{z}\Big)\Big|^2 = -\frac{\alpha'}{2}\ln|z-z'|^2.$$

Therefore

$$\boxed{X^\mu(w,\bar{w})X^\nu(w',\bar{w}') =: X^\mu(w,\bar{w})X^\nu(w',\bar{w}') : -\frac{\alpha'}{2}\eta^{\mu\nu}\,\ln|z-z'|^2}$$

Readers may be familiar with the normal ordering defined by raising and lowering operators. These two definitions will be connected later. The key point of this definition is the property:

$$\partial_1\bar{\partial}_1{:}X^\mu(z_1,\bar{z}_1)X^v(z_2,\bar{z}_2){:}=0 \tag{2.1.26}$$

The above comes from the operator equation (2.1.21) and the following differential equation:

$$\partial\bar{\partial}\ln|z|^2 = 2\pi\delta^2(z,\bar{z}) \tag{2.1.27}$$

Since $\ln|z|^2=\ln z+\ln\bar{z}$, the above equation is obvious for $z\neq 0$. The normalization of the δ-function can be easily verified by integrating both sides using (2.1.10).

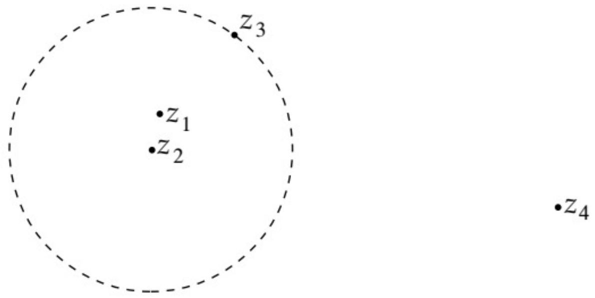


Figure 2.1: The figure above shows the product of four local operators. The OPE represents the asymptotic behavior as $z_1 \to z_2$ as a series. In this process, a pair of operators at z_1, z_2 is replaced by a single operator at z_2. The radius of convergence is the distance to other nearest operators, represented by the dashed lines in the figure.

2.2 The operator product expansion

The basic objects of study in string perturbation theory are path integral expectation values of products of local operators:

$$\langle \mathscr{A}_{i_1}(z_1,\bar{z}_1)\mathscr{A}_{i_2}(z_2,\bar{z}_2)\cdots\mathscr{A}_{i_n}(z_n,\bar{z}_n)\rangle \tag{2.2.1}$$

where $\mathscr{A}_i$ is some basis of the set of local operators. It is very important to understand the behavior of this expectation value in the limit where two of the operators approach each other. A systematic description of this limit is given by the operator product expansion (OPE). As shown in Figure 2.1. The OPE states that the product of two local operators close to each other can be approximated to any precision by a sum of local operators:

$$\mathscr{A}_i(\sigma_1)\mathscr{A}_j(\sigma_2) = \sum_k c^k{}_{ij}(\sigma_1-\sigma_2)\mathscr{A}_k(\sigma_2)\,. \tag{2.2.2}$$

This is an operator statement, meaning it holds in a general expectation value:

$$\langle \mathscr{A}_i(\sigma_1)\mathscr{A}_j(\sigma_2)\cdots\rangle = \sum_k c^k{}_{ij}(\sigma_1-\sigma_2)\,\langle \mathscr{A}_k\,(\sigma_2)\cdots\rangle \tag{2.2.3}$$

where the distance between σ_1, σ_2 is smaller than the distance to all other operators. The coefficient function $c^k{}_{ij}(\sigma_1-\sigma_2)$ determines the dependence on the separation, and also depends on i, j, k, but is independent of other operators in the expectation value; the dependence on the latter only appears in the expectation value on the right side of (2.2.3). These terms are typically arranged in decreasing order of magnitude in the limit $\sigma_1 \to \sigma_2$. Except that the coefficient functions are not simple power series and can be singular as $\sigma_1 \to \sigma_2$, this is analogous to an ordinary Taylor series.

We will now use the special properties of free field theory to give a derivation of the OPE for X^μ theory. In Section 2.9, we will give a derivation for any conformal invariant field theory.

We have seen that normal products satisfy the equations of motion (2.1.26). This indicates that the operator product is a harmonic function of $(z_1, \bar{z}_1)$. A simple result from complex analysis is that a harmonic function is locally the sum of a holomorphic function and an anti-holomorphic function. In particular, it means it is not singular as $z_1 \to z_2$, and can be freely expanded in a Taylor series in z_{12} and $\bar{z}_{12}$. Therefore,

$$X^\mu(z_1,\bar{z}_1)X^v(z_2,\bar{z}_2) = -\frac{\alpha'}{2}\eta^{\mu\nu}\ln|z_{12}|^2 + :X^vX^\mu(z_2,\bar{z}_2):$$

$$+\sum_{k=1}^{\infty}\frac{1}{k!}\Big[(z_{12})^k :X^v\partial^k X^\mu(z_2,\bar z_2): +(\bar z_{12})^k :X^v\bar\partial^k X^\mu(z_2,\bar z_2):\Big] \tag{2.2.4}$$

Remark. *Here (2.1.22a) and*

$$\begin{aligned}&:X^\mu(z_1,\bar z_1)X^\nu(z_2,\bar z_2):\\ =&:X^\nu X^\mu(z_2,\bar z_2): +\sum_{k=1}^{\infty}\frac{1}{k!}\Big[(z_{12})^k :X^\nu\partial^k X^\mu(z_2,\bar z_2): +(\bar z_{12})^k :X^\nu\bar\partial^k X^\mu(z_2,\bar z_2):\Big]\end{aligned}$$

are used. Because of the equations of motion, the terms with mixed derivatives $\partial\bar\partial$ are zero.

Equation (2.2.4) has the form of an OPE. Like the equation of motion (2.1.26) that derived it, this is a statement about operators. For any vacuum expectation value of the product of $X^\mu(z_1,\bar z_1)X^v(z_2,\bar z_2)$ with fields at other points, its behavior as $z_1 \to z_2$ is an infinite series, where each term is a known function of z_{12} and (or) $\bar z_{12}$ multiplied by the expectation value of the local operator that replaces this pair of fields.

The OPE is usually used as an asymptotic expansion, where the first few terms give the dominant behavior when the separation is very small. Most of our applications will be of this type, and we usually write the OPE as clearly singular terms plus unspecified non-singular remainder terms. Instead of using $=$, we will use "$\sim$" to mean "equal in the sense of differing only by non-singular terms." In fact, the OPE converges in conformal invariant field theories. In certain applications, this is very important to us: it makes it possible to reconstruct the entire theory from the coefficient functions. As an example, the radius of convergence of the free field OPE (2.2.4) in any given expectation value is equal to the distance from it to the nearest of the other insertions in the path integral, as shown in Figure 2.1, and the convergence can be proven by standard complex analysis theory.

Various operators on the right side of the OPE (2.2.4) include the product of fields at the same point. In quantum field theory, such products are usually divergent, and thus need to be appropriately truncated and renormalized, but the normal-ordering here makes it well-defined. In free field theory, normal-ordering is a convenient way to define composite operators. In interacting field theory, normal-ordering is of little use because composite operators have additional divergences caused by interaction vertices connected to them. However, the field theories we are interested in are mostly free field theories, or theories related to free field theories, so more details on normal-ordering are given here. Normal-ordering of any number of fields can be defined recursively as

$$\begin{aligned}&:X^{\mu_1}(z_1,\bar z_1)\cdots X^{\mu_n}(z_n,\bar z_n):\\ &\qquad = X^{\mu_1}(z_1,\bar z_1)\cdots X^{\mu_n}(z_n,\bar z_n)+\sum \text{subtractions},\end{aligned} \tag{2.2.5}$$

where the subtractions refer to picking one or more pairs of fields from the product of fields and replacing them with $\frac{1}{2}\alpha'\eta^{\mu_i\mu_j}\ln|z_{ij}|^2$. For example,

$$\begin{aligned}&:X^{\mu_1}(z_1,\bar z_1)X^{\mu_2}(z_2,\bar z_2)X^{\mu_3}(z_3,\bar z_3):= X^{\mu_1}(z_1,\bar z_1)X^{\mu_2}(z_2,\bar z_2)X^{\mu_3}(z_3,\bar z_3)\\ &\quad+\left(\frac{\alpha'}{2}\eta^{\mu_1\mu_2}\ln|z_{12}|^2 X^{\mu_3}(z_3,\bar z_3)+2\text{ permutations}\right).\end{aligned} \tag{2.2.6}$$

This definition can be compactly summarized as

$$:\mathscr{F}:=\exp\left(\frac{\alpha'}{4}\int \mathrm{d}^2z_1\mathrm{d}^2z_2\ln|z_{12}|^2\frac{\delta}{\delta X^\mu(z_1,\bar z_1)}\frac{\delta}{\delta X_\mu(z_2,\bar z_2)}\right)\mathscr{F}, \tag{2.2.7}$$

where $\mathscr{F}$ is any functional of X. This is equivalent to equation (2.2.5): the double functional derivative in the exponent contracts each pair of fields, and summing the exponent of multiple pairings cancels out the fractions in the derivative action. An application of this formal expression is to apply the inverse of the exponent on both sides, which gives

$$\begin{aligned}\mathscr{F} &= \exp\left(-\frac{\alpha'}{4}\int \mathrm{d}^2 z_1 \mathrm{d}^2 z_2 \ln|z_{12}|^2 \frac{\delta}{\delta X^\mu(z_1,\bar{z}_1)}\frac{\delta}{\delta X_\mu(z_2,\bar{z}_2)}\right) {:}\mathscr{F}{:} \\ &= {:}\mathscr{F}{:} + \sum \text{contractions}\,,\end{aligned} \tag{2.2.8}$$

where the contractions are the inverse of subtractions: picking one or more pairs of fields from ${:}\mathscr{F}{:}$ and replacing them with $-\frac{1}{2}\alpha'\eta^{\mu_i\mu_j}\ln|z_{ij}|^2$. For any functionals $\mathscr{F}$ and $\mathscr{G}$ of X, their OPE is generated by

$$ {:}\mathscr{F}{:}{:}\mathscr{G}{:} = {:}\mathscr{F}\mathscr{G}{:} + \sum \text{cross contractions} \tag{2.2.9}$$

The cross contractions now refer to contracting a field in $\mathscr{F}$ and a field in $\mathscr{G}$ and summing over them. This can also be written as

$$ {:}\mathscr{F}{:}{:}\mathscr{G}{:} = \exp\left(-\frac{\alpha'}{2}\int \mathrm{d}^2 z_1 \mathrm{d}^2 z_2\ \ln|z_{12}|^2 \frac{\delta}{\delta X^\mu_F(z_1,\bar{z}_1)}\frac{\delta}{\delta X_{G\mu}(z_2,\bar{z}_2)}\right) {:}\mathscr{F}\mathscr{G}{:} \tag{2.2.10}$$

where the functional derivatives only act on fields in $\mathscr{F}$ or $\mathscr{G}$ respectively. As an example,

$$\begin{aligned}&{:}\partial X^\mu(z)\partial X_\mu(z){:}{:}\partial' X^v(z')\partial' X_v(z'){:} \\ &\quad= {:}\partial X^\mu(z)\partial X_\mu(z)\partial' X^v(z')\partial' X_v(z'){:} \\ &\qquad\quad -4\cdot\frac{\alpha'}{2}(\partial\partial'\ln|z-z'|^2)\ {:}\partial X^\mu(z)\partial' X_\mu(z'){:} \\ &\qquad\quad +2\cdot\eta^\mu{}_\mu\left(-\frac{\alpha'}{2}\partial\partial'\ln|z-z'|^2\right)^2 \\ &\quad\sim \frac{D\alpha'^2}{2(z-z')^4} - \frac{2\alpha'}{(z-z')^2}\ {:}\partial' X^\mu(z')\partial' X_\mu(z'){:} \\ &\qquad\quad -\frac{2\alpha'}{z-z'}\ {:}\partial'^2 X^\mu\left(z'\right)\partial' X_\mu\left(z'\right){:}\ .\end{aligned} \tag{2.2.11}$$

The second term after the equal sign comes from the four ways to contract a single pair of fields, while the third term comes from the two ways to contract two pairs of fields. In the last line, we write the OPE in standard form by performing a Taylor expansion inside the normal ordering: each operator in each term is at z', and arranged in order of singularity.

Remark. *The second term after the equal sign comes from*

$$
\begin{aligned}
: \partial X^\mu(z) \partial X_\mu(z) :: \partial' X^\nu(z') \partial' X_\nu(z') : &= \partial\partial' \left(-\frac{\alpha'}{2} \eta^{\mu\nu} \ln|z-z'|^2 \right) :\partial X_\mu(z) \partial X_\nu(z): \\
&= -\frac{\alpha'}{2} \left(\partial\partial' \ln|z-z'|^2 \right) :\partial X_\mu(z) \partial' X^\mu(z):
\end{aligned}
$$

while the third term comes from

$$
\begin{aligned}
&: \partial X^\mu(z) \partial X_\mu(z) :: \partial' X^\nu(z') \partial' X_\nu(z') : \\
&= \partial\partial' \left(-\frac{\alpha'}{2} \eta^{\mu\nu} \ln|z-z'|^2 \right) \partial\partial' \left(-\frac{\alpha'}{2} \eta_{\mu\nu} \ln|z-z'|^2 \right) \\
&= \eta^{\mu\nu} \eta_{\mu\nu} \left(-\frac{\alpha'}{2} \partial\partial' \ln|z-z'|^2 \right)^2
\end{aligned}
$$

Because $\eta^{\mu\nu}\eta_{\nu\rho} = \delta^\mu{}_\rho$*, so* $\eta^{\mu\nu}\eta_{\mu\nu} = \delta^\mu{}_\mu = D$*, and then using* $\partial\partial' \ln|z-z'|^2 = (z-z')^{-2}$ *and*

$$
:\partial X^\mu(z) \partial' X_\mu(z'): \sim :\partial' X^\mu(z') \partial' X_\mu(z'): + (z-z') :\partial'^2 X^\mu(z') \partial' X_\mu(z):
$$

(2.2.11) is obtained.

Another important example is

$$
\mathscr{F} = \mathrm{e}^{ik_1 \cdot X(z,\bar{z})}, \qquad \mathscr{G} = \mathrm{e}^{ik_2 \cdot X(0,0)} \tag{2.2.12}
$$

The variations $\delta/\delta X_F^\mu$ and $\delta/\delta X_G^\mu$ yield the factors $ik_{1\mu}$ and $ik_{2\mu}$ respectively, so the general result (2.2.10) becomes

$$
\begin{aligned}
:\mathrm{e}^{ik_1 \cdot X(z,\bar{z})}::\mathrm{e}^{ik_2 \cdot X(0,0)}: &= \exp\left(\frac{\alpha'}{2} k_1 \cdot k_2 \ln|z|^2 \right) :\mathrm{e}^{ik_1 \cdot X(z,\bar{z})} \mathrm{e}^{ik_2 \cdot X(0,0)}: \\
&= |z|^{\alpha' k_1 \cdot k_2} :\mathrm{e}^{ik_1 \cdot X(z,\bar{z})} \mathrm{e}^{ik_2 \cdot X(0,0)}: .
\end{aligned} \tag{2.2.13}
$$

To derive the OPE, Taylor expansion in the normal ordering gives

$$
:\mathrm{e}^{ik_1 \cdot X(z,\bar{z})}::\mathrm{e}^{ik_2 \cdot X(0,0)}:= |z|^{\alpha' k_1 \cdot k_2} :\mathrm{e}^{i(k_1+k_2) \cdot X(0,0)} [1 + O(z,\bar{z})]: . \tag{2.2.14}
$$

Remark. *The general result (2.2.10) here is*

$$
\begin{aligned}
&:\mathrm{e}^{ik_1 \cdot X(z,\bar{z})}::\mathrm{e}^{ik_2 \cdot X(0,0)}: \\
&= \left(1 - \frac{\alpha'}{2} \int \mathrm{d}^2 z_1 \mathrm{d}^2 z_2 \ \ln|z_{12}|^2 \frac{\delta}{\delta X_F^\mu(z_1,\bar{z}_1)} \frac{\delta}{\delta X_{G\mu}(z_2,\bar{z}_2)} + \cdots \right) :\mathrm{e}^{ik_1 \cdot X(z,\bar{z})} \mathrm{e}^{ik_2 \cdot X(0,0)}: \\
&=:\mathrm{e}^{ik_1 \cdot X(z,\bar{z})} \mathrm{e}^{ik_2 \cdot X(0,0)}: \\
&\quad \times \left(1 - \frac{\alpha'}{2} \int \mathrm{d}^2 z_1 \mathrm{d}^2 z_2 \ \ln|z_{12}|^2 \times ik_{1\mu} \delta^2(z-z_1, \bar{z}-\bar{z}_1) ik_2^\mu \delta^2(0-z_2, 0, -\bar{z}_2) + \cdots \right) \\
&=:\mathrm{e}^{ik_1 \cdot X(z,\bar{z})} \mathrm{e}^{ik_2 \cdot X(0,0)}: \left(1 + \frac{\alpha'}{2} k_1 \cdot k_2 \ln|z|^2 + \cdots \right) = |z|^{\alpha' k_1 \cdot k_2} :\mathrm{e}^{ik_1 \cdot X(z,\bar{z})} \mathrm{e}^{ik_2 \cdot X(0,0)}: .
\end{aligned}
$$

Note that OPEs such as (2.2.2), (2.2.4) are non-symmetric with respect to σ_1 and σ_2. These OPEs can be reshaped into a symmetric form by expanding around $(\sigma_1 + \sigma_2)/2$. The coefficient functions in the symmetric form satisfy, under the exchange of two operators:

$$
c^k{}_{ij}(\sigma_1 - \sigma_2)_{\mathrm{sym}} = \pm c^k{}_{ji}(\sigma_2 - \sigma_1)_{\mathrm{sym}} , \tag{2.2.15}
$$

where the sign is negative when $\mathscr{A}_i$ and $\mathscr{A}_j$ are both anti-commuting.

2.3 Ward identities and Noether theorem

Worldsheet symmetry plays an important role in string theory. In this section, we first derive some general results for symmetry in field theory.

Consider a field theory in d-dimensional spacetime with action S_ϕ, where $\phi_\alpha(\sigma)$ labels a general field. Suppose it has the symmetry

$$\phi'_\alpha(\sigma) = \phi_\alpha(\sigma) + \delta\phi_\alpha(\sigma)\,, \tag{2.3.1}$$

where $\delta\phi_\alpha$ is proportional to an infinitesimal parameter ϵ. The product of the path integral measure and the weight $\exp(-S)$ is invariant:

$$[\mathrm{d}\phi']\ \exp\big(-S[\phi']\big) = [\mathrm{d}\phi]\exp(-S[\phi])\,. \tag{2.3.2}$$

In field theory, a continuous symmetry implies the existence of a conserved current (Noether's theorem) and Ward identities (which constrain operator products of the current). To derive these results, consider the field variation

$$\phi'_\alpha(\sigma) = \phi_\alpha(\sigma) + \rho(\sigma)\delta\phi_\alpha(\sigma) \tag{2.3.3}$$

This is not a symmetry, but $[\mathrm{d}\phi]\exp(-S)$ is invariant when ρ is a constant, so its variation must be proportional to $\partial_a\rho$:

$$\begin{aligned}&[\mathrm{d}\phi']\ \exp\big(-S[\phi']\big)\\ &\quad= [\mathrm{d}\phi]\ \exp(-S[\phi])\left[1 + \frac{i\epsilon}{2\pi}\int \mathrm{d}^d\sigma\, g^{1/2} j^a(\sigma)\partial_a\rho(\sigma) + O\left(\epsilon^2\right)\right]\end{aligned} \tag{2.3.4}$$

The unknown coefficient $j^a(\sigma)$ comes from the variation of the measure and the action. Take ρ to be non-zero only in a small region. Outside this region, consider a path integral containing a general insertion "$\cdots$". This insertion is invariant under (2.3.3):

$$\begin{aligned}0 &= \int[\mathrm{d}\phi']\ \exp\big(-S[\phi']\big)\cdots - \int[\mathrm{d}\phi]\exp(-S[\phi])\cdots\\ &= \frac{i\epsilon}{2\pi}\int \mathrm{d}^d\sigma\, g^{1/2} j^a(\sigma)\partial_a\rho(\sigma) + O(\epsilon^2) = \frac{\epsilon}{2\pi i}\int \mathrm{d}^d\sigma\, g^{1/2}\rho(\sigma)\left\langle\nabla_a j^a(\sigma)\cdots\right\rangle\end{aligned} \tag{2.3.5}$$

thus we have

$$\nabla_a j^a = 0\,. \tag{2.3.6}$$

Proof. We use the fact that $\partial_a(g^{1/2}v^a(c)) = g^{1/2}\nabla_a v^a(c)$ as shown here

$$\begin{aligned}\mathrm{LHS} = \partial_a(\sqrt{g}v^a) &= (\partial_a\sqrt{g})v^a + \sqrt{g}\partial_a v^a\\ &= \frac{1}{2\sqrt{g}} g g^{bc}\partial_a g_{bc} v^a + \sqrt{g}\partial_a v^a\\ &= \sqrt{g}\left(\frac{1}{2}g^{bc}\partial_a g_{bc}v^a + \partial_a v^a\right)\end{aligned}$$

For the RHS we find

$$\begin{aligned}
\text{RHS} = \sqrt{g}\nabla_a v^a &= \sqrt{g}\left(\partial_a v^a + \Gamma^a_{ab} v^b\right) \\
&= \sqrt{g}\left[\partial_a v^a + \frac{1}{2}(g^{ac}\partial_a g_{cb} + g^{ac}\partial_b g_{ca} - g^{ac}\partial_c g_{ab})v^b\right] \\
&= \sqrt{g}\left(\partial_a v^a + \frac{1}{2} g^{ac}\partial_b g_{ca} v^b\right)
\end{aligned}$$

The second and last term of the connection cancel after interchanging the dummy indices a and c. This is now equal to the LHS.
To derive (2.3.5) we start from (2.3.4), which implies

$$\begin{aligned}
0 &= \frac{i\epsilon}{2\pi}\int d^d\sigma\, g^{1/2} j^a(\sigma)\partial_a\rho(\sigma) \\
&= -\frac{i\epsilon}{2\pi}\int d^d\sigma\, \partial_a(g^{1/2} j^a(\sigma))\rho(\sigma) \\
&= \frac{\epsilon}{2\pi i}\int d^d\sigma\, g^{1/2}\nabla_a j^a(\sigma)\rho(\sigma)
\end{aligned}$$

□

To derive Ward identities, let $\rho(\sigma)$ be 1 in region R and 0 outside R. Additionally, introduce a general local operator $\mathscr{A}(\sigma_0)$ at a point σ_0 within the path integral region R. Inserting arbitrary operators "$\cdots$" outside R gives

$$\delta\mathscr{A}(\sigma_0) + \frac{\epsilon}{2\pi i}\int_R \mathrm{d}^d\sigma\, g^{1/2}\nabla_a j^a(\sigma)\mathscr{A}(\sigma_0) = 0\,. \tag{2.3.7}$$

Proof. Similar to (2.3.5), we now have

$$0 = \int [\mathrm{d}\phi]\, \mathrm{e}^{-S[\phi]}\left(\delta\mathscr{A}(\sigma_0) + \frac{i\epsilon}{2\pi}\int \mathrm{d}^d\sigma\, j^a(\sigma)\partial_a\rho(\sigma)\mathscr{A}(\sigma_0)\right)\cdots\,,$$

then using integration by parts, (2.3.7) can be obtained. □

Equivalently,

$$\nabla_a j^a(\sigma)\mathscr{A}(\sigma_0) = g^{-1/2}\delta^d(\sigma-\sigma_0)\frac{2\pi}{i\epsilon}\delta\mathscr{A}(\sigma_0) + \sigma \text{ total derivative}\,. \tag{2.3.8}$$

Divergence theorem gives

$$\int_{\partial R} \mathrm{d}A\, n_a j^a \mathscr{A}(\sigma_0) = \frac{2\pi}{i\epsilon}\delta\mathscr{A}(\sigma_0)\,, \tag{2.3.9}$$

where $\mathrm{d}A$ is the area element, and n_a is the outward normal vector. In 2-dimensional flat space, this is

$$\oint_{\partial R}(j\mathrm{d}z - \tilde{\jmath}\mathrm{d}\bar{z})\mathscr{A}(z_0,\bar{z}_0) = \frac{2\pi}{\epsilon}\delta\mathscr{A}(z_0,\bar{z}_0) \tag{2.3.10}$$

Here we have dropped the indices $j \equiv j_z$, $\tilde{\jmath} \equiv j_{\bar{z}}$. We use $\tilde{\jmath}$ instead of $\bar{\jmath}$ because $\tilde{\jmath}$ is not the conjugate of j. The Minkowski density j_0 is generally Hermitian, $(j_z)^\dagger = \frac{1}{2}(j_1 - ij_2)^\dagger = j_z$.

Proof. The proof of (2.3.9) used

$$\int_R \mathrm{d}^d\sigma\, g^{1/2}\nabla_a j^a = \int_R \mathrm{d}^d\sigma\, \partial_a\big(g^{1/2}j^a\big) = \int_{\partial R} \mathrm{d}A\, n_\alpha j^a\,,$$

Using the two-dimensional Stokes' theorem $\int_R \mathrm{d}^2 z\,\left(\partial_z v^z + \partial_{\bar z} v^{\bar z}\right) = i\oint_{\partial R}(v^z \mathrm{d}\bar z - v^{\bar z}\mathrm{d}\bar z)$ (2.3.10) is obtained. □

Importantly, the Noether theorem and Ward identities are local properties, and do not depend on distant boundary conditions, nor on whether they are invariant under the symmetry. In particular, because $\rho(\sigma)$ is only non-zero in R, only symmetry transformations in R need to be defined.

In conformal invariant theories, usually j_z is holomorphic and $j_{\bar z}$ is anti-holomorphic. In this case, the currents $(j_z, 0)$ and $(0, j_{\bar z})$ are separately conserved. Thus the integral yields the residue in the OPE:

$$\operatorname{Res}_{z\to z_0} j(z)\mathscr{A}(z_0,\bar z_0) + \overline{\operatorname{Res}}_{\bar z\to\bar z_0}\tilde\jmath(\bar z)\mathscr{A}(z_0,\bar z_0) = \frac{1}{i\epsilon}\delta\mathscr{A}(z_0,\bar z_0) \tag{2.3.11}$$

Here $\operatorname{Res}, \overline{\operatorname{Res}}$ refer to taking the coefficients of $(z-z_0)^{-1}$ and $(\bar z-\bar z_0)^{-1}$ respectively.

An example: Returning to the massless free scalar field, and consider *spacetime* translations $\delta X^\mu = \epsilon a^\mu$. Under $\delta X^\mu(\sigma) = \epsilon\rho(\sigma)a^\mu$,

$$\delta S = \frac{1}{4\pi\alpha'}\int \mathrm{d}^2\sigma\, 2(\delta\partial^a X^\mu)\partial_a X_\mu = \frac{\epsilon a_\mu}{2\pi\alpha'}\int \mathrm{d}^2\sigma\,\partial^a X^\mu \partial_a\rho\,. \tag{2.3.12}$$

This is equivalent to taking the Noether current in (2.3.5) as $j_a(\sigma) = a_\mu j^\mu_a$, where

$$j^\mu_a = \frac{i}{\alpha'}\partial_a X^\mu\,. \tag{2.3.13}$$

The OPE of the current and the exponential operator gives

$$j^\mu(z)\,{:}\mathrm{e}^{ik\cdot X(0,0)}{:} \sim \frac{k^\mu}{2z}\,{:}\mathrm{e}^{ik\cdot X(0,0)}{:}\,, \tag{2.3.14a}$$
$$\tilde\jmath^\mu(\bar z)\,{:}\mathrm{e}^{ik\cdot X(0,0)}{:} \sim \frac{k^\mu}{2\bar z}\,{:}\mathrm{e}^{ik\cdot X(0,0)}{:}\,. \tag{2.3.14b}$$

Proof. Explanation for (2.3.14a): from (2.3.13) it can be obtained that $j^\mu = j^\mu_z = \frac{i}{\alpha'}\partial_z X^\mu$, so

$$\begin{aligned}
\frac{i}{\alpha'}\partial X^\mu(z) : \sum_{n=0}^\infty \frac{i^n}{n!}(k\cdot X(0))^n : &= \frac{i}{\alpha'} : \sum_{n=0}^\infty \frac{i^n}{n!}\partial X^\mu(z)(k\cdot X(0))^n : \\
&\sim \frac{i}{\alpha'} : \sum_{n=1}^\infty \frac{i^n}{n!}\sum_{i=1}^n \partial X^\mu(z)(k\cdot X(0))_i \prod_{j\neq i}(k\cdot X(0))_j : \\
&= \frac{i}{\alpha'} : \sum_{n=1}^\infty \frac{i^n}{n!}\, n\,\partial X^\mu(z)(k\cdot X(0))(k\cdot X(0))^{n-1} : \\
&= \frac{i}{\alpha'}\left(-\frac{\alpha'}{2}\eta^{\mu\nu}\partial\ln(z\bar z)\right)k_\nu : \sum_{n=1}^\infty \frac{i^n}{(n-1)!}(k\cdot X(0))^{n-1} :
\end{aligned}$$

$$
\begin{aligned}
&= \frac{i}{\alpha'}\left(-\frac{\alpha'}{2}\frac{k^\mu}{z}\right) : \sum_{n=1}^{\infty} \frac{i^n}{(n-1)!}(k \cdot X(0))^{n-1} : \\
&= \frac{k^\mu}{2z} : e^{ik\cdot X(0)} : .
\end{aligned}
$$

□

Another example: worldsheet translations $\delta\sigma^a = \epsilon v^a$, then $\delta X^\mu = -\epsilon v^a \partial_a X^\mu$, the Noether current is

$$
j_a = i v^b T_{ab} \,, \tag{2.3.15a}
$$

$$
T_{ab} = -\frac{1}{\alpha'} :\left(\partial_a X^\mu \partial_b X_\mu - \frac{1}{2}\delta_{ab}\partial_c X^\mu \partial^c X_\mu\right): \,. \tag{2.3.15b}
$$

Proof. Proof of (2.3.15a):

$$
\delta\sigma^u = \epsilon v^a, \quad \delta X^\mu = -\partial_a X^\mu \delta\sigma^a = -\epsilon v^a \partial_a X^\mu
$$

$$
\begin{aligned}
\delta S &= \frac{1}{4\pi\alpha'}\int d^2\sigma 2\delta\partial^\alpha X^\mu \partial_a X_\mu \\
&= -\frac{1}{2\pi\alpha'}\int d^2\sigma \delta X^\mu \partial^a \partial_a X_\mu \\
&= -\frac{1}{2\pi\alpha'}\int d^2\sigma \left(-\epsilon v^a \partial_\alpha X^\mu\right)\rho(\sigma)\partial^\alpha \partial_a X_\mu \\
&= \frac{\epsilon v^\alpha}{2\pi\alpha'}\int d^2\sigma \partial_a X^\mu \rho(\sigma)\partial^b \partial_b X_u
\end{aligned}
$$

The last integral can now be evaluated as:

$$
\begin{aligned}
\int d^2\sigma \partial_b\left(\partial_a X^\mu \rho(\sigma)\right)\partial^b X_u &= \int d^2\sigma \partial_a X^\mu \partial^b X_\mu \partial_b \rho(\sigma) + \rho(\sigma)\partial_b \partial_a X^\mu \partial^b X_\mu \\
&= \int d^2\sigma \partial_a X^\mu \partial^b X_\mu \partial_b \rho(\sigma) - \partial_a \rho(\sigma)\partial_b X^\mu \partial^b X_\mu \\
&\quad - \rho(\sigma)\partial_a X^\mu \partial_b \partial^b X^\mu \\
&= \int d^2\sigma \partial_a X^\mu \rho(\sigma)\partial^b \partial_b X^\mu \\
&= \frac{1}{2}\int d^2\sigma \partial_a X^\mu \partial^b X_\mu \partial_b \rho(\sigma) - \partial_a \rho(\sigma)\partial_b X^\mu \partial^b X_u
\end{aligned}
$$

Hence

$$
\frac{\epsilon v^a}{2\pi\alpha'}\int d^2\sigma \frac{1}{2}\left[\partial_a X^\mu \partial^b X_\mu \partial_b \rho(\sigma) - \partial_a \rho \partial_b X^\mu \partial^b X_\mu\right] = \frac{i\epsilon}{2\pi}\int d^2 z j^a(\sigma)\partial_a(\rho)
$$

□

2.4 Conformal invariance

The energy-momentum tensor (2.3.15b) is traceless: $T_a{}^a = 0$. In complex coordinates, this is

$$
g^{ab}T_{ab} = \underbrace{g^{zz}}_{=0} T_{zz} + 2g^{z\bar{z}}T_{z\bar{z}} + \underbrace{g^{\bar{z}\bar{z}}}_{=0} T_{\bar{z}\bar{z}} \implies T_{z\bar{z}} = 0\,. \tag{2.4.1}
$$

The conservation law $\partial^a T_{ab} = 0$ implies that in any theory with $T_a{}^a = 0$

$$\bar{\partial} T_{zz} = \partial T_{\bar{z}\bar{z}} = 0\,. \tag{2.4.2}$$

Therefore

$$T(z) \equiv T_{zz}(z)\,, \qquad \tilde{T}(\bar{z}) \equiv T_{\bar{z}\bar{z}}(\bar{z}) \tag{2.4.3}$$

are holomorphic and anti-holomorphic respectively. For massless free scalars

$$T(z) = -\frac{1}{\alpha'} :\partial X^\mu \partial X_\mu:\,, \qquad \tilde{T}(\bar{z}) = -\frac{1}{\alpha'} :\bar{\partial} X^\mu \bar{\partial} X_\mu:\,, \tag{2.4.4}$$

according to the equations of motion, they are indeed holomorphic and anti-holomorphic.

The tracelessness of T_{ab} implies a much larger symmetry. The currents

$$j(z) = iv(z)T(z)\,, \qquad \tilde{\jmath}(\bar{z}) = iv(z)^* \tilde{T}(\bar{z}) \tag{2.4.5}$$

are conserved for holomorphic $v(z)$.

Proof. Denote $j_z = iv^z T_{zz}$, $j_{\bar{z}} = iv^{\bar{z}} T_{\bar{z}\bar{z}}$. So $\nabla_a j^a = \bar{\partial} j_z + \partial j_{\bar{z}} = 0$. □

For free scalar theory, the OPE is obtained

$$T(z)X^\mu(0) \sim \frac{1}{z}\partial X^\mu(0), \quad \tilde{T}(\bar{z})X^\mu(0) \sim \frac{1}{\bar{z}}\bar{\partial} X^\mu(0) \tag{2.4.6}$$

then the Ward identity gives the transformation

$$\delta X^\mu = -\epsilon v(z)\partial X^\mu - \epsilon v(z)^* \bar{\partial} X^\mu \tag{2.4.7}$$

This is the infinitesimal coordinate transformation $z' = z + \epsilon v(z)$. The finite transformation is

$$X'^\mu\left(z', \bar{z}'\right) = X^\mu(z, \bar{z}), \quad \text{where } z' = f(z)\,. \tag{2.4.8}$$

(2.4.8) is called a *conformal transformation*.

Proof for (2.4.7): from (2.3.11)

$$\frac{1}{i\epsilon}\delta X^\mu = \operatorname{Res}_{z\to z_0} iv^z T(z) X^\mu(z_0, \bar{z}_0) + \overline{\operatorname{Res}}_{\bar{z}\to\bar{z}_0} iv^{\bar{z}} \bar{T}(\bar{z}) X^\mu(z_0, \bar{z}_0)$$

Using (2.4.6), (2.4.7) is obtained.

Conformal invariance should not be confused with diff invariance. We are in flat space, with no independent metric field used for variation. So the transformation $z \to z'$ really changes the distance between points. We don't just happen to have this invariance; it is a non-trivial statement about the dynamics. For the scalar action (2.1.11), the conformal transformations of ∂ and $\bar{\partial}$ exactly cancel out the conformal transformation of $\mathrm{d}^2 z$. The mass term $m^2 X^\mu X_\mu$ is not conformal invariant. Ultimately, there must be a close connection with the diff× Weyl symmetry of the Polyakov string.

Consider the special case

$$z' = \zeta z\,, \tag{2.4.9}$$

ζ is any complex number. The phase of ζ is the rotation of the system, while its magnitude is the rescaling of the size of the system. This scale invariance is often viewed as an approximate symmetry in particle physics, and statistical systems near critical points are also described by scale-invariant field theories.

For generalized conformal transformations, examine its effect on $\mathrm{d}s^2 = g_{ab}\mathrm{d}\sigma^a\mathrm{d}\sigma^b = \mathrm{d}z\mathrm{d}\bar{z}$

$$\mathrm{d}s'^2 = \mathrm{d}z'\mathrm{d}\bar{z}' = \frac{\partial z'}{\partial z}\frac{\partial \bar{z}'}{\partial \bar{z}}\mathrm{d}z\mathrm{d}\bar{z}\,. \tag{2.4.10}$$

As shown in Figure 2.2, we see that the conformal transformation turns an infinitesimal square into an infinitesimal square but adjusts its size. An anti-holomorphic function $z' = f(z)^*$ has the same property but changes the orientation. Most systems invariant under (2.4.9) are usually invariant under the much larger conformal invariance. A theory with this invariance is called a *conformal field theory* (CFT).

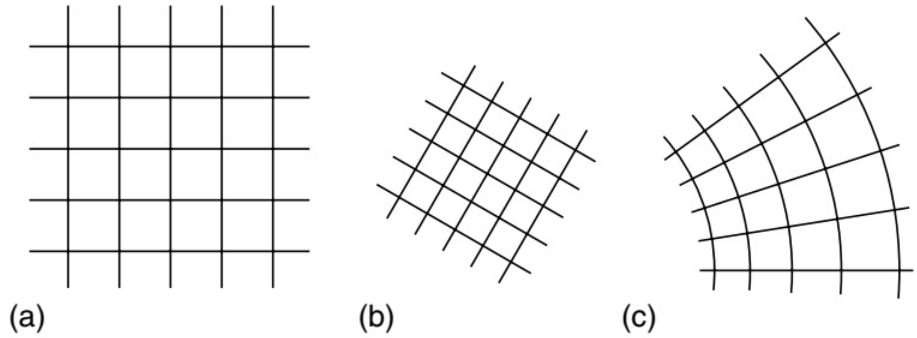


Figure 2.2: (a) Two-dimensional region. (b) The effect of the special conformal transformation (2.4.9). (c) The effect of a more general conformal transformation.

Conformal Invariance and the OPE

Conformal invariance imposes a strong constraint on the form of the OPE, especially the OPE of the energy-momentum tensor. Examine the OPE of T and a general operator $\mathscr{A}$. Because $T(z)$ and $\tilde{T}(\bar{z})$ are (anti-)holomorphic except at coincidence points, the corresponding coefficient functions must also have this property. The OPE of T and $\mathscr{A}$ is thus a Laurent expansion, where the terms are integer powers of z, though the power may be negative. Furthermore, all singular terms are determined by the conformal transformation of $\mathscr{A}$. To see this, we write a general expansion for the singularity

$$T(z)\mathscr{A}(0,0) \sim \sum_{n=0}^{\infty} \frac{1}{z^{n+1}}\mathscr{A}^{(n)}(0,0) \tag{2.4.11}$$

similar for $\tilde{T}$; the operator coefficient $\mathscr{A}^{(n)}$ is left to be determined later. Under an infinitesimal conformal transformation $z' = z + \epsilon v(z)$, when the z^{-n-1} term in the $T\mathscr{A}$ OPE is multiplied by the z^n term in $v(z)$, a simple pole appears in $v(z)T(z)\mathscr{A}(0,0)$. Therefore the (2.3.11) form of the Ward identity implies

$$\delta\mathscr{A}(z,\bar{z}) = -\epsilon\sum_{n=0}^{\infty}\frac{1}{n!}\left[\partial^n v(z)\mathscr{A}^{(n)}(z,\bar{z}) + \bar{\partial}^n v(z)^*\tilde{\mathscr{A}}^{(n)}(z,\bar{z})\right] \tag{2.4.12}$$

Thus the operator $\mathscr{A}^{(n)}$ is determined by the conformal transformation of $\mathscr{A}$. It is convenient to take operators that are eigenstates under the rigid transformation (2.4.9) as a basis

$$\mathscr{A}'(z',\bar{z}') = \zeta^{-h}\bar{\zeta}^{-\tilde{h}}\mathscr{A}(z,\bar{z})\,. \tag{2.4.13}$$

$(h, \tilde{h})$ are called weights. $h + \tilde{h}$ is the dimension of $\mathscr{A}$, which determines its behavior under rescaling, and $h - \tilde{h}$ is the spin, which determines its behavior under rotation. The derivative ∂_z raises h by one, $\partial_{\bar{z}}$ raises $\tilde{h}$ by one. The Ward identity for the transformation (2.4.13) and the Ward identity for translation $\delta\mathscr{A} = -\epsilon v^a \partial_a \mathscr{A}$ determine the part of the OPE

$$T(z)\mathscr{A}(0,0) = \cdots + \frac{h}{z^2}\mathscr{A}(0,0) + \frac{1}{z}\partial\mathscr{A}(0,0) + \cdots \tag{2.4.14}$$

Similarly for $\tilde{T}$.

Remark. *The variation caused by the transformation (2.4.13) is*

$$\delta\mathscr{A} = \mathscr{A}'(z,\bar{z}) - \mathscr{A}(z,\bar{z}) = \mathscr{A}'(z'/\zeta, \bar{z}'/\zeta) - \mathscr{A}(z,\bar{z})$$

Let $\zeta = 1 + \epsilon v$, *then* $z'/(1+\epsilon v) = z'(1-\epsilon v)$ *Then*

$$\begin{aligned}\delta\mathscr{A} &= \mathscr{A}'(z',\bar{z}') - \epsilon v z'\partial\mathscr{A}'\left(z'\right) - \epsilon\bar{v}\bar{z}'\bar{\partial}\mathscr{A}'(\bar{z}') - \mathscr{A}(z,\bar{z}) \\ &= [(1+\epsilon v)^h(1+\epsilon\bar{v})^{-h} - 1]\mathscr{A} - \epsilon v z'\partial\mathscr{A}'\left(z'\right) - \epsilon\bar{v}\bar{z}'\bar{\partial}\mathscr{A}'(\bar{z}')\end{aligned}$$

Particularly important are tensor operators or primary fields $\mathcal{O}$. The general conformal transformation acting on them gives

$$\mathcal{O}'(z',\bar{z}') = (\partial_z z')^{-h}(\partial_{\bar{z}}\bar{z}')^{-\tilde{h}}\mathcal{O}(z,\bar{z})\,. \tag{2.4.15}$$

The OPE (2.4.11) reduces to

$$T(z)\mathcal{O}(0,0) = \frac{h}{z^2}\mathcal{O}(0,0) + \frac{1}{z}\partial\mathcal{O}(0,0) + \cdots\,, \tag{2.4.16}$$

The more singular terms in the general OPE (2.4.14) are gone.

Take free X^μ CFT as an example once again. Some weights of typical operators are

$$\left.\begin{array}{llll} X^\mu & (0,0)\,, & \partial X^\mu & (1,0)\,, \\ \bar{\partial}X^\mu & (0,1)\,, & \partial^2 X^\mu & (2,0)\,, \\ :\mathrm{e}^{ik\cdot X}: & \left(\frac{\alpha' k^2}{4}, \frac{\alpha' k^2}{4}\right)\,. & & \end{array}\right\} \tag{2.4.17}$$

Remark. *Note that* X^μ *has conformal weights* $h = \bar{h} = 0$, *yet its two-point function in (2.1.22b) does not take the standard form*

$$\langle\phi_1(z_1,\bar{z}_1)\phi_2(z_2,\bar{z}_2)\rangle = \frac{C_{12}}{(z_1-z_2)^h(\bar{z}_1-\bar{z}_2)^{\bar{h}}}$$

Therefore, X^μ *is not a primary field. Nevertheless, although* X^μ *itself is not primary, the operator* $:e^{ik\cdot X(z,\bar{z})}:$ *is. Indeed, its two-point function from (2.2.13) takes the expected form:*

$$\begin{aligned}\Big\langle :e^{ik\cdot X(z,\bar{z})}::e^{ip\cdot X(w,\bar{w})}: \Big\rangle &= |z-w|^{\alpha' k\cdot p}\Big\langle :e^{ik\cdot X(z,\bar{z})}e^{ip\cdot X(w,\bar{w})}: \Big\rangle \\ &= |z-w|^{\alpha' k\cdot p}\,\delta^{(D)}(\vec{k}+\vec{p}) \\ &= \frac{1}{|z-w|^{\alpha' k^2}}\end{aligned} \tag{2.4.18}$$

where in the second line we have set $\vec{p} = -\vec{k}$. *where we used*

$$\langle 0;0|\,e^{i(k_1+k_2)\cdot x}e^{\alpha'(k_1\ln|z|^2 + k_2\ln|w|^2)\cdot p}\,|0;0\rangle = \langle 0;0|\,e^{i(k_1+k_2)\cdot x}\,|0;0\rangle = \langle 0;0|0;k_1+k_2\rangle$$

$$= \delta(\vec{k}_1 + \vec{k}_2).$$

Except for $\partial^2 X^\mu$, all quantities transform as tensors. More generally, general products of exponentials and derivatives

$$:\left(\prod_i \partial^{m_i} X^{\mu_i}\right)\left(\prod_j \bar{\partial}^{n_j} X^{v_j}\right) \mathrm{e}^{ik\cdot X}: \tag{2.4.19}$$

have weights

$$\left(\frac{\alpha' k^2}{4} + \sum_i m_i, \frac{\alpha' k^2}{4} + \sum_j n_j\right). \tag{2.4.20}$$

For any pair of operators, applying rigid transformations, rescaling, and rotation on both sides of the OPE completely determines the dependence of the coefficient function on z

$$\mathscr{A}_i(z_1, \bar{z}_1)\mathscr{A}_j(z_2, \bar{z}_2) = \sum_k z_{12}^{h_k - h_i - h_j} \bar{z}_{12}^{\tilde{h}_k - \tilde{h}_i - \tilde{h}_j} c^k{}_{ij} \mathscr{A}_k(z_2, \bar{z}_2)\,, \tag{2.4.21}$$

where $c^k{}_{ij}$ are now constants. In all interesting cases, the weights appearing on the right side of the OPE (2.4.21) have a lower bound, so the singularity in the operator product is constrained. More general conformal transformations impose further constraints on the OPE: they completely determine the OPE of all fields with the OPE of primary fields.

Note that in general, the conformal transformation properties of normal products are not determined by the naive transformation of the product. For example, the transformation rule (2.4.7) for X^μ would naively imply

$$\delta \mathrm{e}^{ik\cdot X} = -\epsilon v(z)\partial \mathrm{e}^{ik\cdot X} - \epsilon v(z)^* \bar{\partial} \mathrm{e}^{ik\cdot X} \qquad \text{(naive)} \tag{2.4.22}$$

making it a (0,0) tensor. Since defining the operator product at a point requires renormalization, this modification is a quantum effect. In particular, it enters here because the subtraction of $\ln|z_{12}|^2$ in : : has a clear dependence on the coordinate system.

Conformal Properties of the Energy-Momentum Tensor

The OPE of the energy-momentum tensor with itself is obtained from (2.2.11)

$$\begin{aligned} T(z)T(0) &= \frac{\eta^\mu{}_\mu}{2z^4} - \frac{2}{\alpha' z^2} :\partial X^\mu(z)\partial X_\mu(0): + :T(z)T(0): \\ &\sim \frac{D}{2z^4} + \frac{2}{z^2}T(0) + \frac{1}{z}\partial T(0) \end{aligned} \tag{2.4.23}$$

A similar result exists for $\tilde{T}$. The OPE of $T(z)\tilde{T}(\bar{z}')$ must be non-singular; it cannot have a singularity at $(z - z')$ because it is anti-holomorphic in z' over a non-zero interval. Similarly, it cannot have a singularity at $(\bar{z} - \bar{z}')$. The same result holds for any OPE of a holomorphic operator and an anti-holomorphic operator.

Thus T is *not* a tensor. And the OPE (2.4.23) implies the transformation rule

$$\epsilon^{-1}\delta T(z) = -\frac{D}{12}\partial_z^3 v(z) - 2\partial_z v(z)T(z) - v(z)\partial_z T(z)\,. \tag{2.4.24}$$

In a general CFT, the transformation of $T(z)$ is

$$\epsilon^{-1}\delta T(z) = -\frac{c}{12}\partial_z^3 v(z) - 2\partial_z v(z)T(z) - v(z)\partial_z T(z)\,, \tag{2.4.25}$$

where c is a constant called the central charge. The central charge of a free scalar field is 1. For D free scalar fields it is D. The transformation (2.4.25) is the most general form, which is linear in v. This is consistent with the symmetry of the TT OPE, where the three z subscripts are required by rigid transformation, rescaling and rotation invariance. Scaling, rotation and translation symmetries determine the coefficients of the 2nd and 3rd terms. Furthermore, by examining the commutator of two such transformations, it can be proven that $\partial_a c = 0$. This is a general result in quantum field theory. Operators that are independent of position must be c-numbers. The corresponding TT OPE is

$$T(z)T(0) \sim \frac{c}{2z^4} + \frac{2}{z^2}T(0) + \frac{1}{z}\partial T(0) \tag{2.4.26}$$

The finite form of the transformation rule (2.4.25) is

$$(\partial_z z')^2 T'(z') = T(z) - \frac{c}{12}\{z', z\}\,, \tag{2.4.27}$$

where $\{f, z\}$ represents the Schwarzian derivative

$$\{f, z\} = \frac{2\partial_z^3 f \partial_z f - 3\partial_z^2 f \partial_z^2 f}{2\partial_z f \partial_z f} \tag{2.4.28}$$

There is a corresponding form for $\tilde{T}$. In general CFT there may be a different central charge $\tilde{c}$. The non-tensor behavior of the energy-momentum tensor has several important physical consequences. It should be emphasized that "non-tensor" refers to conformal transformations. Under coordinate transformations, the energy-momentum tensor will have the usual tensor properties.

2.5 Free CFTs

In this section, we will discuss three types of free field CFTs—linear dilaton theory, bc theory, and $\beta\gamma$ theory. bc theory will be used in the next chapter to gauge-fix the Polyakov string. These three theories have wide applications.

Linear Dilaton CFT

This type of CFT is based on the same action (2.1.11). But the energy-momentum tensor is

$$T(z) = -\frac{1}{\alpha'} :\partial X^\mu \partial X_\mu: + V_\mu \partial^2 X^\mu \tag{2.5.1a}$$

$$\tilde{T}(\bar{z}) = -\frac{1}{\alpha'} :\bar{\partial} X^\mu \bar{\partial} X_\mu: + V_\mu \bar{\partial}^2 X^\mu \tag{2.5.1b}$$

where v_μ is some fixed D-vector. Calculating the TT OPE, one finds it is exactly the standard form (2.4.26). But the central charge is

$$c = \tilde{c} = D + 6\alpha' V_\mu V_\mu \tag{2.5.2}$$

The TX^μ OPE and Ward identity (2.4.12) imply the conformal transformation

$$\delta X^\mu = -\epsilon v \partial X^\mu - \epsilon v^* \bar{\partial} X^\mu - \frac{\epsilon}{2}\alpha' V^\mu \left[\partial v + (\partial v)^*\right] \tag{2.5.3}$$

This is a different conformal symmetry for the same action. X^μ no longer transforms as a tensor; its variation now has a non-homogeneous part. Incidentally, free massless scalar fields in two dimensions have a large class of symmetries—much more than we happened to mention.

The energy-momentum tensor plays a special role in string theory. In particular, it tells us how to couple with a curved metric—different values of V^μ can be viewed as different CFTs. The vector V^μ picks out a direction in spacetime, so this CFT is not Lorentz invariant, and we have little interest in it. In Section 3.7 we will see a physical interpretation of the linear dilaton CFT, and we will encounter it in some technical applications later. A different variation of the free scalar CFT is to take some X^μ to be periodic; we will address this in Chapter 8.

bc CFT

The second type of CFT has anti-commuting fields b and c, and action

$$S = \frac{1}{2\pi}\int \mathrm{d}^2 z\, b\bar{\partial} c\,. \tag{2.5.4}$$

For any given constant λ, such that

$$h_b = \lambda, \quad h_c = 1-\lambda \tag{2.5.5}$$

If b and c transform as tensors with weights $(h_b, 0)$ and $(h_c, 0)$, (2.5.4) is conformal invariant, and thus we have another class of CFTs (which we will see in Chapter 10 are secretly the same as the linear dilaton). The operator equations of motion can be obtained by the same method as (2.1.16), (2.1.19)

$$\bar{\partial} c(z) = \bar{\partial} b(z) = 0 \tag{2.5.6a}$$

$$\bar{\partial} b(z) c(0) = 2\pi\delta^2(z,\bar{z}) \tag{2.5.6b}$$

The bb OPE and cc OPE satisfy the sourceless equations of motion. The normal product of bc is

$$:b(z_1)c(z_2):= b(z_1)c(z_2) - \frac{1}{z_{12}} \tag{2.5.7}$$

Since

$$\bar{\partial}\frac{1}{z} = \partial\frac{1}{\bar{z}} = 2\pi\delta^2(z,\bar{z}) \tag{2.5.8}$$

(2.5.7) satisfies the naive equations of motion. (2.5.7) can be obtained by integrating over a region containing the origin and using integration by parts. The normal ordering of ordinary products of fields is combinatorially the same as X^μ CFT, being the sum over contractions or subtractions. We must be careful: since b and c are anti-commuting, exchanging fields will flip the sign. The operator products are

$$b(z_1)c(z_2) \sim \frac{1}{z_{12}}, \qquad c(z_1)b(z_2) \sim \frac{1}{z_{12}} \tag{2.5.9}$$

Two sign flips were performed in the second OPE, one coming from anti-commutation and one from $z_1 \leftrightarrow z_2$. Other operator products are non-singular:

$$b(z_1)b(z_2) = O(z_{12})\,, \qquad c(z_1)c(z_2) = O(z_{12})\,. \tag{2.5.10}$$

These are not only holomorphic, but also have a zero because of the antisymmetry.

Noether's theorem gives the energy-momentum tensor

$$T(z) =:(\partial b)c: -\lambda\partial(:bc:)\,, \tag{2.5.11a}$$

$$\tilde{T}(\bar{z}) = 0\,. \tag{2.5.11b}$$

(2.5.11) can be proven by giving the OPEs of T with b and T with c, which has the standard tensor form (2.4.16) given by weights. The TT OPE is the standard form (2.4.26). where

$$c = -3(2\lambda-1)^2 + 1\,, \qquad \tilde{c} = 0\,. \tag{2.5.12}$$

This is a purely holomorphic CFT, and also an example of $c \neq \tilde{c}$. Of course there is a corresponding anti-holomorphic theory

$$S = \frac{1}{2\pi}\int \mathrm{d}^2 z\, \tilde{b}\partial\tilde{c}\,, \tag{2.5.13}$$

which is the same as above under $z \leftrightarrow \bar{z}$. The bc theory has a ghost number symmetry $\delta b = -i\epsilon b, \delta c = i\epsilon c$. Corresponding Noether current

$$j = - :bc: \tag{2.5.14}$$

The components are holomorphic and anti-holomorphic respectively, the latter being 0. When both holomorphic bc fields and anti-holomorphic bc fields are present, ghost numbers are separately conserved. The following current is not a tensor:

$$T(z)j(0) \sim \frac{1-2\lambda}{z^3} + \frac{1}{z^2}j(0) + \frac{1}{z}\partial j(0)\,. \tag{2.5.15}$$

This implies the transformation rule

$$\epsilon^{-1}\delta j = -v\partial j - j\partial v + \frac{2\lambda - 1}{2}\partial^2 v\,, \tag{2.5.16}$$

whose finite form is

$$(\partial_z z')j_{z'}(z') = j_z(z) + \frac{2\lambda-1}{2}\frac{\partial_z^2 z'}{\partial_z z'} \tag{2.5.17}$$

The case where b and c have equal weights is $h_b = h_c = \frac{1}{2}$, then central charge $c = 1$. Here we usually use the notation $b \to \psi, c \to \bar{\psi}$. For this case, bc CFT can be split into two in a conformal invariant way

$$\psi = 2^{-1/2}(\psi_1 + i\psi_2), \quad \bar{\psi} = 2^{-1/2}(\psi_1 - i\psi_2) \tag{2.5.18a}$$

$$S = \frac{1}{4\pi}\int \mathrm{d}^2 z \left(\psi_1\bar{\partial}\psi_1 + \psi_2\bar{\partial}\psi_2\right) \tag{2.5.18b}$$

$$T = -\frac{1}{2}\psi_1\partial\psi_1 - \frac{1}{2}\psi_2\partial\psi_2\,. \tag{2.5.18c}$$

Each ψ theory has central charge $1/2$.

The bc theory with $\lambda = 2$, with weights $(h_b, h_c) = (2, -1)$, will arise from gauge-fixing the Polyakov string as Faddeev-Popov ghosts in the next chapter. In more general string theories in Volume II, the ψ theory will appear more widely.

$\beta\gamma$ CFT

The third type of CFT is much like the bc theory but the fields commute, β is a $(h_\beta, 0)$ tensor, γ is a $(h_\gamma, 0)$ tensor, where

$$h_\beta = \lambda, \qquad h_\gamma = 1 - \lambda\,. \tag{2.5.19}$$

The action is

$$S = \frac{1}{2\pi}\int \mathrm{d}^2 z\, \beta\bar{\partial}\gamma\,. \tag{2.5.20}$$

Via the equations of motion

$$\bar{\partial}\gamma(z) = \bar{\partial}\beta(z) = 0 \tag{2.5.21}$$

one can see that these fields are also holomorphic. The equations of motion and operator products are derived in the standard way. Because the statistics are changed, some signs in the operator products are different

$$\beta(z_1)\gamma(z_2) \sim -\frac{1}{z_{12}}, \qquad \gamma(z_1)\beta(z_2) \sim \frac{1}{z_{12}}\,. \tag{2.5.22}$$

The energy-momentum tensor is

$$T =:(\partial\beta)\gamma: - \lambda\partial(:\beta\gamma:)\,, \tag{2.5.23a}$$

$$\tilde{T} = 0 \, . \tag{2.5.23b}$$

The central charge differs only by a sign

$$c = 3(2\lambda - 1)^2 - 1, \qquad \tilde{c} = 0 \, . \tag{2.5.24}$$

The $\beta\gamma$ theory with $\lambda = \frac{3}{2}$, weights $(h_\beta, h_\gamma) = \left(\frac{3}{2}, -\frac{1}{2}\right)$, will appear as Faddeev-Popov ghosts from gauge-fixing the superstring in Chapter 10.

2.6 The Virasoro algebra

So far, in this chapter, we have studied the local properties of two-dimensional field theories. We now examine the spectrum of this theory. Spatial coordinates may be periodic, as for closed strings; or have boundaries, as for open strings. For the periodic case, let

$$\sigma^1 \sim \sigma^1 + 2\pi \, . \tag{2.6.1}$$

Let the Euclidean time coordinate be

$$-\infty < \sigma^2 < \infty \tag{2.6.2}$$

so that these two dimensions form an infinite cylinder. Complex coordinates are again useful, and there are two natural choices, one of which is

$$w = \sigma^1 + i\sigma^2 \, , \tag{2.6.3}$$

such that $w \sim w + 2\pi$. The other is

$$z = \exp(-iw) = \exp\left(-i\sigma^1 + \sigma^2\right) \tag{2.6.4}$$

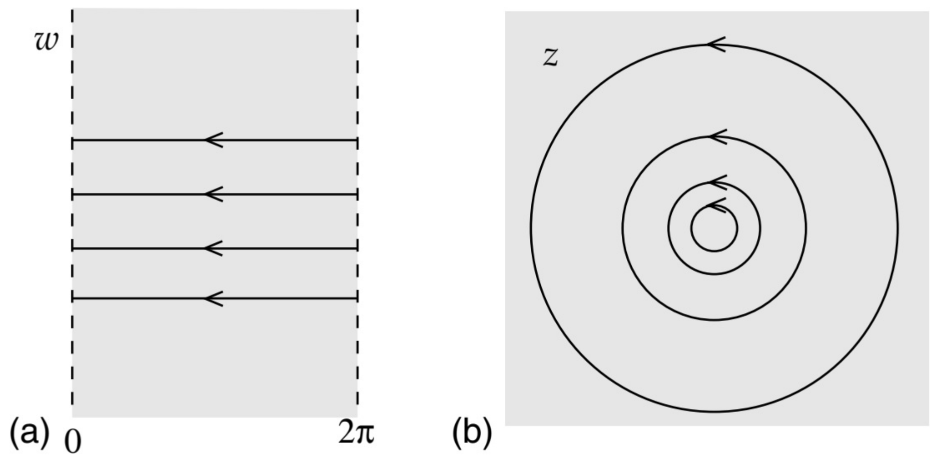


Figure 2.3: Closed String Coordinates

In w coordinates, time corresponds to moving $\sigma^2 = \operatorname{Im} w$. In z coordinates, time moves radially, and the origin is the distant past. These coordinates are related through a conformal transformation. For the canonical interpretation of the theory, w coordinates are natural, but z coordinates are also quite useful, and most expressions are written in this frame of reference.

For holomorphic or anti-holomorphic operators, we can perform a Laurent expansion

$$T_{zz}(z) = \sum_{m=-\infty}^{\infty} \frac{L_m}{z^{m+2}}, \quad \tilde{T}_{\bar{z}\bar{z}}(\bar{z}) = \sum_{m=-\infty}^{\infty} \frac{\tilde{L}_m}{\bar{z}^{m+2}} \tag{2.6.5}$$

The coefficients are called Virasoro generators. Given by contour integrals

$$L_m = \oint_C \frac{\mathrm{d}z}{2\pi i z}\, z^{m+2} T_{zz}(z) \tag{2.6.6}$$

where C is any contour around the origin counterclockwise. At $\sigma^2 = 0$, the Laurent expansion is an ordinary Fourier transform

$$T_{ww}(w) = -\sum_{m=-\infty}^{\infty} \exp\left(im\sigma^1 - m\sigma^2\right) T_m \tag{2.6.7a}$$

$$T_{\bar{w}\bar{w}}(\bar{w}) = -\sum_{m=-\infty}^{\infty} \exp\left(-im\sigma^1 - m\sigma^2\right) \tilde{T}_m \tag{2.6.7b}$$

where

$$T_m = L_m - \delta_{m,0}\frac{c}{24}, \qquad \tilde{T}_m = \tilde{L}_m - \delta_{m,0}\frac{\tilde{c}}{24}. \tag{2.6.8}$$

The additional offset in T_0 comes from the non-tensor transformation (2.4.27)

$$T_{ww} = (\partial_w z)^2 T_{zz} + \frac{c}{24}. \tag{2.6.9}$$

In the $w = \sigma^1 + i\sigma^2$ frame, the time-translation Hamiltonian is

$$H = \int_0^{2\pi} \frac{\mathrm{d}\sigma^1}{2\pi}\, T_{22} = L_0 + \tilde{L}_0 - \frac{c+\tilde{c}}{24}. \tag{2.6.10}$$

Note the $+2$ in the Laurent expansion, which is canceled by the conformal transformation of T in the Fourier transform. Similarly, the Laurent expansion of a holomorphic field with weight h will contain h in the exponent. Cut open the path integral on a circle where time $Imw = \ln|z|$ is constant. The Virasoro generators become operators in the usual sense. Because of holomorphicity, the integral (2.6.6) is independent of C, so in particular, it is invariant under time translation. That is, they are conserved charges, charges associated with conformal invariance.

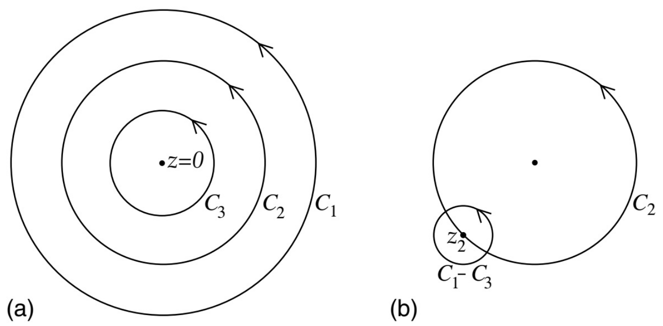


Figure 2.4: Contours

An important fact: the OPE of currents determines the algebra of the corresponding charges. Examine general charges $Q_i, i = 1, 2$, given as contour integrals of holomorphic currents

$$Q_i\{C\} = \oint_C \frac{\mathrm{d}z}{2\pi i} j_i \tag{2.6.11}$$

Examine the combination

$$Q_1\{C_1\}Q_2\{C_2\} - Q_2\{C_2\}Q_1\{C_3\} \tag{2.6.12}$$

The contour is shown in Figure 2.4a.

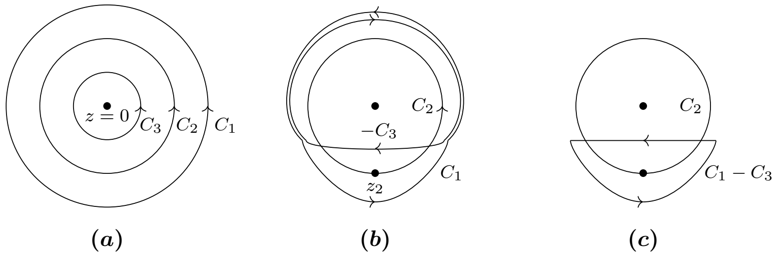


Figure 2.5: Deforming Contours. By deforming C_3 and C_1 is follows that $C_1 - C_3$ is equivalent to a contour around z_2.

The **order** of factors **is irrelevant** because they are just integration variables in a path integral (unless the two charges are anti-commuting. In that case there would be an extra sign and all commutators would become anti-commutators). When we split the path integral for an operator interpretation, the operator order is determined by time order, here $t_1 > t_2 > t_3$. The path integral for the combination (2.6.12) corresponds to the matrix element

$$\hat{Q}_1\hat{Q}_2 - \hat{Q}_2\hat{Q}_1 \equiv [\hat{Q}_1, \hat{Q}_2] \tag{2.6.13}$$

Now, for a given point z_2 on the contour C_2, we can deform the difference in contours C_1, C_3, as shown in Figure 2.4b, so the commutator is given by the residue of the OPE

$$[Q_1, Q_2]\{C_2\} = \oint_{C_2} \frac{\mathrm{d}z_2}{2\pi i} \operatorname{Res}_{z_1\to z_2} j_1(z_1)j_2(z_2) \tag{2.6.14}$$

Proof. We start from (2.6.12) but change the indices of the charges to letters to avoid confusion. It is the radial ordering of the contours that determines the **ordering** of the corresponding operators, therefore $Q_a(C_1)Q_b(C_2) - Q_a(C_3)Q_b(C_2)$ object as operators on the Hilbert space is same as $\hat{Q}_a\hat{Q}_b - \hat{Q}_b\hat{Q}_a = [\hat{Q}_a, \hat{Q}_b]$ as C_1 is the most outward contour, i.e. the largest time, and C_3 is the most inward contour, i.e. the smallest time. Now we also have

$$Q_a(C_1)Q_b(C_2) - Q_b(C_2)Q_a(C_3) = Q_b(C_2)[Q_a(C_1) - Q_a(C_3)] = Q_b(C_2)Q_a(C_1 - C_3) \tag{2.6.15}$$

by the contour deformation of fig. 2.5. But, on the one hand the transformation of an operator is given by the commutation relation with the corresponding charge, $\delta Q = [Q, A]$ and on the other hand we have from (2.3.11) that $\delta A(z) = \frac{1}{2\pi i}\oint dw\, j(w)A(z)$ with $j(z)$ the conserved current. Using a transformation under Q_a on an operator Q_b for a point on C_2 and the definition of the charges (2.6.11) we get

$$\begin{aligned}\delta Q_b\{C_2\} &= [Q_a, Q_b]\{C_2\} = Q_b(C_2)Q_a(C_1 - C_3)\\ &= \oint_{C_2} \frac{dz_2}{2\pi i} \oint_{C_1 - C_3} \frac{dz_1}{2\pi i} j_a(z_1)j_b(z_2)\end{aligned}$$

But the latter part just picks up the residue of the OPE $j_a(z_1)j_b(z_2)$ when $z_1 \to z_2$. Thus

$$\delta Q_b\{C_2\} = \oint_{C_2} \frac{dz_2}{2\pi i} \operatorname{Res}_{z_1\to z_2} j_a(z_1) j_b(z_2) \tag{2.6.16}$$

and so we find

$$[\hat{Q}_a, \hat{Q}_b] = \oint_{C_2} \frac{dz_2}{2\pi i} \operatorname{Res}_{z_1\to z_2} j_a(z_1) j_b(z_2) \tag{2.6.17}$$

Which is the relation that allows us to pass from OPEs to the commutation relations of conserved charges. □

Contour discussion allows us to pass back and forth between OPEs and commutation relations. Let us emphasize that for conserved currents, knowing the singularity in the OPE is equivalent to knowing the corresponding charge commutator algebra. Replacing the conserved charge $Q_2\{C_2\}$ with an arbitrary operator, the same discussion applies:

$$[Q, \mathscr{A}(z_2, \bar{z}_2)] = \operatorname{Res}_{z_1\to z_2} j(z_1)\mathscr{A}(z_2, \bar{z}_2) = \frac{1}{i\epsilon}\delta\mathscr{A}(z_2, \bar{z}_2) \tag{2.6.18}$$

This is exactly the familiar statement: charge Q generates the corresponding transformation. Similarly, for contour integrals of anti-holomorphic currents

$$\tilde{Q}\{C\} = -\oint_C \frac{\mathrm{d}\bar{z}}{2\pi i}\, \tilde{\jmath}\,, \tag{2.6.19}$$

Ward identity and contour discussion imply

$$[\tilde{Q}, \mathscr{A}(z_2, \bar{z}_2)] = \operatorname{Res}_{\bar{z}_1\to \bar{z}_2} \tilde{\jmath}(\bar{z}_1)\mathscr{A}(z_2, \bar{z}_2) = \frac{1}{i\epsilon}\delta\mathscr{A}(z_2, \bar{z}_2) \tag{2.6.20}$$

Applied to the Virasoro algebra (2.6.6),

$$\begin{aligned}
&\operatorname{Res}_{z_1\to z_2} z_1^{m+1}T(z_1)z_2^{n+1}T(z_2)\\
&= \operatorname{Res}_{z_1\to z_2} z_1^{m+1}z_2^{n+1}\left(\frac{c}{2z_{12}^4} + \frac{2}{z_{12}^2}T(z_2) + \frac{1}{z_{12}}\partial T(z_2)\right)\\
&= \frac{c}{12}(\partial^3 z_2^{m+1})z_2^{n+1} + 2(\partial z_2^{m+1})z_2^{n+1}T(z_2) + z_2^{m+n+2}\partial T(z_2)\\
&= \frac{c}{12}(m^3-m)z_2^{m+n-1} + (m-n)z_2^{m+n+1}T(z_2) + \text{total derivative}
\end{aligned} \tag{2.6.21}$$

where $j_m(z) = z^{m+1}T(z)$. Then the z_2 contour integral on the right gives the Virasoro algebra:

$$[L_m, L_n] = (m-n)L_{m+n} + \frac{c}{12}\left(m^3 - m\right)\delta_{m,-n} \tag{2.6.22}$$

$\tilde{L}_m$ satisfies the same algebra with central charge $\tilde{c}$. Thus any CFT has infinitely many conserved charges, i.e., Virasoro generators. They act in Hilbert space and satisfy algebra (2.6.22). Focus on a few simple properties first. Generally, dealing with eigenstates of L_0 and $\tilde{L}_0$, generator L_0 satisfies

$$[L_0, L_n] = -nL_n \tag{2.6.23}$$

If $|\psi\rangle$ is an eigenstate of L_0 with eigenvalue h, then

$$L_0 L_n|\psi\rangle = L_n\left(L_0 - n\right)|\psi\rangle = (h-n)L_n|\psi\rangle \tag{2.6.24}$$

So $L_n|\psi\rangle$ is an eigenstate with eigenvalue $h-n$. Generators with $n<0$ raise the eigenvalue of L_0, while $n>0$ lowers it. Three generators L_0 , $L_{\pm 1}$ form a closed algebra without central charge:

$$[L_0, L_1] = -L_1, \quad [L_0, L_{-1}] = L_{-1}, \quad [L_1, L_{-1}] = 2L_0 \tag{2.6.25}$$

This is the algebra $SL(2,\mathbb{R})$, which differs from $SU(2)$ by a sign. For a holomorphic tensor field $\mathcal{O}$ with weight $(h, 0)$, the Laurent coefficients are

$$\mathcal{O}(z) = \sum_{m=-\infty}^{\infty} \frac{m}{z^{m+h}} \tag{2.6.26}$$

From OPE (2.4.16) the commutator is obtained

$$[L_m, \mathcal{O}_n] = [(h-1)m - n]\mathcal{O}_{m+n} \,. \tag{2.6.27}$$

Again, $n > 0$ mode lowers L_0, $n < 0$ mode raises L_0.

Proof. Proof of (2.6.27): in (2.6.20) let $j_m(z) = z^{m+1}T(z)$ and $j_n(z) = z^{n+h-1}\mathcal{O}(z)$, then

$$\begin{aligned}
&\mathrm{Res}_{z_1\to z_2}\, z_1^{m+1}T(z_1)z_2^{n+h-1}\mathcal{O}(z_2) \\
&= \mathrm{Res}_{z_1\to z_2}\, z_1^{m+1}z_2^{n+h-1}\left(\frac{h}{z_{12}^2}O(z_2) + \frac{1}{z_{12}}\partial\mathcal{O}(z_2)\right) \\
&= h(\partial z_2^{m+1})z_2^{n+h-1}\mathcal{O}(z_2) + z_2^{m+n+h}\partial\mathcal{O}(z_2) \\
&= h(m+1)z_2^{m+n+h-1}\mathcal{O}(z_2) - (m+n+h)z_2^{m+n+h-1}\mathcal{O}(z_2) + \text{total derivative}
\end{aligned}$$

Using (2.6.26), the first term becomes

$$h(m+1)z_2^{m+n+h-1}\mathcal{O}(z_2) \sim h(m+1)\sum \mathcal{O}_q z_2^{m+n-q-1} \sim h(m+1)\mathcal{O}_{m+n}$$

The second term becomes

$$-(m+n+h)z_2^{m+n+h-1}\sum \frac{\mathcal{O}_q}{z_2^{q+h}} \sim -(m+n+h)\mathcal{O}_{m+n}$$

Adding the two yields (2.6.27). □

For open strings, let

$$0 \le \mathrm{Re}\, w \le \pi \quad \Leftrightarrow \quad \mathrm{Im}\, z \ge 0 \tag{2.6.28}$$

where $z = -\exp(-iw)$. The coordinate region is shown in Figure 2.6.

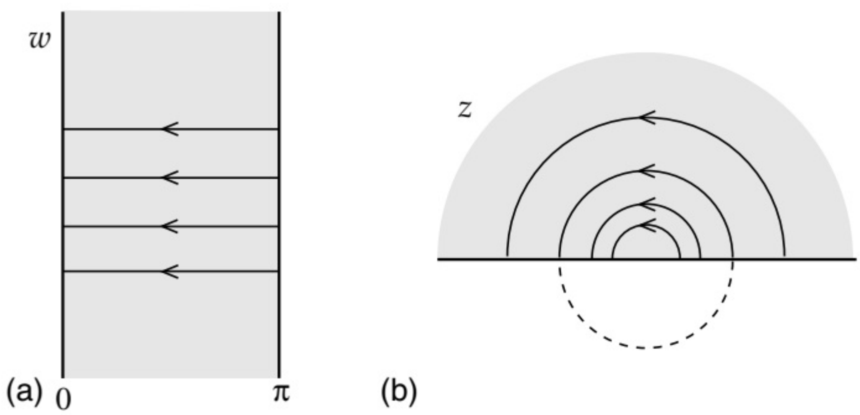


Figure 2.6: Contours

At the boundary, energy-momentum satisfies

$$T_{ab}n^a t^b = 0 \,, \tag{2.6.29}$$

where n^a and t^b are the normal vector and tangential vector respectively. To see this, examine a coordinate system where the boundary is a straight line. The appearance of the boundary breaks the normal translation invariance, but not the tangential one, making the current $T_{ab}t^b$ still conserved. Then the boundary condition (2.6.29) is a statement that the current out of the boundary is zero. In the current case, this becomes

$$T_{ww} = T_{\bar{w}\bar{w}},\ \mathrm{Re}\, w = 0\,,\pi \qquad \Leftrightarrow \qquad T_{zz} = T_{\bar{z}\bar{z}}\,, \mathrm{Im}\, z = 0\,. \tag{2.6.30}$$

It is convenient to use the "doubling trick". Define T_{zz} in the lower half z-plane as the value of $T_{\bar{z}\bar{z}}$ at the image point $z' = \bar{z}$ in the upper half plane:

$$T_{zz}(z) \equiv T_{\bar{z}\bar{z}}\left(\bar{z}'\right), \quad \mathrm{Im}\, z < 0\,. \tag{2.6.31}$$

Equations of motion and boundary conditions are summarized as: T_{zz} is holomorphic throughout the complex plane. Because the boundary condition couples T and $\tilde{T}$, there is only one set of Virasoro algebra:

$$L_m = \frac{1}{2\pi i}\int_C \left(\mathrm{d}z\, z^{m+1}T_{zz} - \mathrm{d}\bar{z}\, \bar{z}^{m+1}T_{\bar{z}\bar{z}}\right) \qquad\qquad = \frac{1}{2\pi i}\oint \mathrm{d}z\, z^{m+1}T_{zz}(z)$$

In the first line, the contour is a semicircle centered at the origin; in the second line, we use the "doubling trick" to write L_m as a closed contour. Once again, these L_m satisfy the Virasoro algebra

$$[L_m, L_n] = (m-n)L_{m+n} + \frac{c}{12}\left(m^3 - m\right)\delta_{m,-n}\,. \tag{2.6.32}$$

2.7 Mode expansions

Free Scalar

In free field theory, the field decomposes into harmonic oscillators, and the spectrum and energy-momentum tensor can be given in mode form. We start with closed strings. In the theory of X^μ, ∂X and $\bar{\partial}X$ are anti-holomorphic. So there is a similar Laurent expansion:

$$\partial X^\mu(z) = -i\left(\frac{\alpha'}{2}\right)^{1/2}\sum_{m=-\infty}^{\infty}\frac{\alpha_m^\mu}{z^{m+1}}\,, \qquad \bar{\partial}X^\mu(\bar{z}) = -i\left(\frac{\alpha'}{2}\right)^{1/2}\sum_{m=-\infty}^{\infty}\frac{\tilde{\alpha}_m^\mu}{\bar{z}^{m+1}}\,. \tag{2.7.1}$$

Equivalently,

$$\alpha_m^\mu = \left(\frac{2}{\alpha'}\right)^{1/2}\oint\frac{\mathrm{d}z}{2\pi}\, z^m\partial X^\mu(z)\,, \tag{2.7.2a}$$

$$\tilde{\alpha}_m^\mu = -\left(\frac{2}{\alpha'}\right)^{1/2}\oint\frac{\mathrm{d}\bar{z}}{2\pi}\, \bar{z}^m\bar{\partial}X^\mu(\bar{z})\,. \tag{2.7.2b}$$

The single-valuedness of X^μ implies $\alpha_0^\mu = \tilde{\alpha}_0^\mu$, furthermore, the Noether current of spacetime translation is $i\partial_a X^\mu/\alpha'$, so spacetime momentum

$$p^\mu = \frac{1}{2\pi i}\oint_C(\mathrm{d}z\, j^\mu - \mathrm{d}\bar{z}\,\tilde{\jmath}^\mu) = \left(\frac{2}{\alpha'}\right)^{1/2}\alpha_0^\mu = \left(\frac{2}{\alpha'}\right)^{1/2}\tilde{\alpha}_0^\mu \tag{2.7.3}$$

Integrating expression (2.7.1) gives

$$X^\mu(z,\bar{z}) = x^\mu - i\frac{\alpha'}{2}p^\mu\ln|z|^2 + i\left(\frac{\alpha'}{2}\right)^{1/2}\sum_{\substack{m=-\infty\\ m\neq 0}}^{\infty}\frac{1}{m}\left(\frac{\alpha_m^\mu}{z^m} + \frac{\tilde{\alpha}_m^\mu}{\bar{z}^m}\right)\,. \tag{2.7.4}$$

Whether starting from standard canonical commutators, or from contour discussions or the XX OPE, one can give

$$[\alpha_m^\mu, \alpha_n^\nu] = [\tilde{\alpha}_m^\mu, \tilde{\alpha}_n^\nu] = m\delta_{m,-n}\eta^{\mu\nu} , \tag{2.7.5a}$$

$$[x^\mu, p^\nu] = i\eta^{\mu\nu} , \tag{2.7.5b}$$

Other commutators are zero. The spectrum starts from the state $|0;k\rangle$ with momentum k^μ, which is annihilated by all $n > 0$ lowering modes α_n^μ, while other states in the spectrum are given by the action of raising modes $(n < 0)$ in all possible ways on $|0;k\rangle$.

We now wish to expand Virasoro generators in mode operator form. Substituting the Laurent expansion of X^μ into the energy-momentum tensor (2.4.4) and collecting terms of a given order in z, gives

$$L_m \sim \frac{1}{2} \sum_{n=-\infty}^{\infty} \alpha_{m-n}^\mu \alpha_{\mu n} . \tag{2.7.6}$$

"$\sim$" means neglecting the order of operators. When $m \neq 0$, the expansion (2.7.6) is well-defined and correct—the mode operator in each term commutes so the sequence is not important. For $m = 0$, place the lowering operators on the right and introduce a normal-ordering constant:

$$L_0 = \frac{\alpha' p^2}{4} + \sum_{n=1}^{\infty} \left(\alpha_{-n}^\mu \alpha_{\mu n}\right) + a^X . \tag{2.7.7}$$

We encountered the same problem in Section 1.3, where we dealt with it in a heuristic way. Now there is a finite and explicit definition on the left, defined by Laurent coefficients in the $: :$ ordered energy-momentum tensor, so the normal-ordering constant is finite and calculable. There are several ways to determine it, the simplest is using the Virasoro algebra,

$$2L_0|0;0\rangle = (L_1 L_{-1} - L_{-1} L_1)|0;0\rangle = 0 , \tag{2.7.8}$$

thus

$$a^X = 0 . \tag{2.7.9}$$

Here we have used the known forms of L_1, L_{-1}: each term either contains a lowering operator or p^μ, thus annihilating $|0;0\rangle$. From the OPE, we determined the central charge of the Virasoro algebra. It can also be directly derived from the mode operator expression of the Virasoro algebra.

Introduce a new notation. Symbol $: :$ represents creation-annihilation normal ordering: place all lowering operators on the right of all creation operators. Note there is a negative sign when exchanging anti-commuting operators. For this reason, we include p^μ among the lowering operators and x^μ among the raising operators. In this notation, we can write

$$L_m = \frac{1}{2} \sum_{n=-\infty}^{\infty} : \alpha_{m-n}^\mu \alpha_{\mu n} : \tag{2.7.10}$$

We have now introduced two forms of normal ordering, namely conformal normal ordering $: :$ and creation-annihilation normal ordering $: :$. The advantage of the former is that the operators it produces have simpler OPEs and conformal transformation properties; the latter may be more familiar to readers, and it is more convenient for dealing with matrix elements of operators. We now present the relationship between the two. Start by comparing time-ordered products and creation-annihilation normal products. For the product $X^\mu(z,\bar{z})X^\nu(z',\bar{z}')$ with $|z| > |z'|$, substitute in the mode expansion and move the lowering operator in $X^\mu(z,\bar{z})$ to the right:

$$X^\mu(z,\bar{z})X^\nu(z',\bar{z}') =: X^\mu(z,\bar{z})X^\nu(z',\bar{z}') :$$

$$
\begin{aligned}
&+\frac{\alpha'}{2}\eta^{\mu\nu}\left[-\ln|z|^2+\sum_{m=1}^{\infty}\frac{1}{m}\left(\frac{z'^m}{z^m}+\frac{\bar z'^m}{\bar z^m}\right)\right]\\
&=: X^\mu(z,\bar z)X^\nu(z',\bar z') : -\frac{\alpha'}{2}\eta^{\mu\nu}\ln|z-z'|^2
\end{aligned}
\tag{2.7.11}
$$

Since $|z|>|z'|$, the left side is time-ordered and the sum converges. Define the relationship between the time-ordered product and the conformal normal product given by (2.1.22) (definition (2.1.22) uses path integral language, so the products on the right are time-ordered under the operator system). This is the same as (2.7.11), making

$$
: X^\mu(z,\bar z)X^\nu(z',\bar z') :=: X^\mu(z,\bar z)X^\nu(z',\bar z'): \tag{2.7.12}
$$

This relationship does not hold in a general CFT, and in fact, some combination of p^μ and lowering operators gives a simpler result. This also does not hold for the w form of conformal normal sequence of operators. From (2.7.12) the mode expansion (2.7.10) can be written immediately, and the second derivation of $a^X=0$ is given.

The creation-annihilation normal product of more than two X^μ has the same combinatorial factors as the $::$ ordering. That is, they are both obtained from time-ordered products by summing over all contractions—Wick's theorem in field theory. To convert the normal ordering of an operator from one form to another, use the difference of two-point functions for contractions and sum over them. That is, if we have two orderings,

$$
[X^\mu(z,\bar z)X^v(z',\bar z')]_1=[X^\mu(z,\bar z)X^v(z',\bar z')]_2+\eta^{\mu v}\Delta(z,\bar z,z',\bar z')\,, \tag{2.7.13}
$$

then for a general operator $\mathscr{F}$

$$
[\mathscr{F}]_1=\exp\left(\frac{1}{2}\int \mathrm{d}^2z\,\mathrm{d}^2z'\,\Delta(z,\bar z,z',\bar z')\frac{\delta}{\delta X^\mu(z,\bar z)}\frac{\delta}{\delta X_\mu(z',\bar z')}\right)[\mathscr{F}]_2\,. \tag{2.7.14}
$$

For a linear dilaton CFT, the Laurent expansion and commutator are invariant, while the Virasoro generator contains an extra term

$$
L_m=\frac{1}{2}\sum_{n=-\infty}^{\infty}:\alpha^\mu_{m-n}\alpha_{\mu n}:+i\left(\frac{\alpha'}{2}\right)^{1/2}(m+1)V^\mu\alpha_{\mu m}\,. \tag{2.7.15}
$$

bc CFT

Fields b and c have Laurent expansion

$$
b(z)=\sum_{m=-\infty}^{\infty}\frac{b_m}{z^{m+\lambda}}\,,\qquad c(z)=\sum_{m=-\infty}^{\infty}\frac{c_m}{z^{m+1-\lambda}}\,. \tag{2.7.16}
$$

Precisely, if λ is an integer, they are just Laurent expansions. The half-integer case is also interesting, and we will deal with it carefully in Chapter 10. The OPE gives the anti-commutator

$$
\{b_m,c_n\}=\delta_{m,-n}\,. \tag{2.7.17}
$$

First examine states annihilated by all $n>0$ operators. The oscillator algebra of b_0 and c_0 generates two such ground states $|{\downarrow}\rangle$ and $|{\uparrow}\rangle$. They have properties

$$
b_0|{\downarrow}\rangle=0\,,\quad b_0|{\uparrow}\rangle=|{\downarrow}\rangle\,, \tag{2.7.18a}
$$

$$
c_0|{\downarrow}\rangle=|{\uparrow}\rangle\,,\quad c_0|{\uparrow}\rangle=0\,, \tag{2.7.18b}
$$

$$
b_n|{\downarrow}\rangle=b_n|{\uparrow}\rangle=c_n|{\downarrow}\rangle=c_n|{\uparrow}\rangle=0\,,\quad n>0\,. \tag{2.7.18c}
$$

Ordinary states are obtained by acting on these states with $n < 0$ modes, but due to anti-commutation, they can act at most once. For some later reasons, it will be convenient to group b_0 with lowering operators and c_0 with raising operators, so we pick $|\downarrow\rangle$ as the ghost vacuum $|0\rangle$. In string theory we will have a holomorphic bc theory and an anti-holomorphic $\tilde{b}\tilde{c}$ theory. Each has $\lambda = 2$. And the closed string spectrum contains the product of two pairs of theories. The Virasoro generators are

$$L_m = \sum_{n=-\infty}^{\infty} (m\lambda - n) : b_n c_{m-n} : + \delta_{m,0} a^{\mathrm{g}} . \tag{2.7.19}$$

The ordering constant can be determined like (2.7.8), which gives

$$\begin{aligned} 2L_0|\downarrow\rangle &= (L_1 L_{-1} - L_{-1} L_1)|\downarrow\rangle \\ &= (\lambda b_0 c_1)[(1-\lambda) b_{-1} c_0]|\downarrow\rangle = \lambda(1-\lambda)|\downarrow\rangle . \end{aligned} \tag{2.7.20}$$

Thus $a^{\mathrm{g}} = \frac{1}{2}\lambda(1-\lambda)$ and

$$L_m = \sum_{n=-\infty}^{\infty} (m\lambda - n) : b_n c_{m-n} : + \frac{\lambda(1-\lambda)}{2} \delta_{m,0} . \tag{2.7.21}$$

The constant can also be obtained by solving for the relationship between $: :$ and $: :$. For the ghost number current (2.5.14), $j = -\,:bc:$, the charge is

$$\begin{aligned} N^{\mathrm{g}} &= -\frac{1}{2\pi i} \int_0^{2\pi} \mathrm{d}w \, j_w \\ &= \sum_{n=1}^{\infty} (c_{-n} b_n - b_{-n} c_n) + c_0 b_0 - \frac{1}{2} . \end{aligned} \tag{2.7.22}$$

It satisfies

$$[N^{\mathrm{g}}, b_m] = -b_m , \qquad [N^{\mathrm{g}}, c_m] = c_m , \tag{2.7.23}$$

thus c minus b number is excitation. The ground states have ghost number $\pm\frac{1}{2}$:

$$N^{\mathrm{g}}|\downarrow\rangle = -\frac{1}{2}|\downarrow\rangle , \qquad N^{\mathrm{g}}|\uparrow\rangle = \frac{1}{2}|\uparrow\rangle . \tag{2.7.24}$$

This depends on the value of the ordering constant, but can be guessed: the average ghost number of the ground state should be zero and the ghost number changes sign under $b \leftrightarrow c$.

Open String

For open strings, Neumann boundary condition becomes $\partial_z X^\mu = \partial_{\bar{z}} X^\mu$ on the real axis. There is only one set of modes, the boundary condition requires $\alpha_m^\mu = \tilde{\alpha}_m^\mu$ in the expansion (2.7.1). The spacetime momentum integral (2.7.3) only integrates a semicircle, so the current normalization is

$$\alpha_0^\mu = (2\alpha')^{1/2} p^\mu . \tag{2.7.25}$$

Then the expansion for X^μ is

$$X^\mu(z, \bar{z}) = x^\mu - i\alpha' p^\mu \ln |z|^2 + i \left(\frac{\alpha'}{2}\right)^{1/2} \sum_{\substack{m=-\infty \\ m\neq 0}}^{\infty} \frac{\alpha_m^\mu}{m} (z^{-m} + \bar{z}^{-m}) . \tag{2.7.26}$$

Also

$$L_0 = \alpha' p^2 + \sum_{n=1}^{\infty} \alpha_{-n}^\mu \cdot \alpha_{\mu n} . \tag{2.7.27}$$

Commutators as usual

$$[\alpha_m^\mu, \alpha_n^\nu] = m\delta_{m,-n}\eta^{\mu\nu}\,, \qquad [x^\mu, p^\nu] = i\eta^{\mu\nu}\,. \tag{2.7.28}$$

For bc theory, the boundary conditions related to strings are

$$c(z) = \tilde{c}(\bar{z})\,, \quad b(z) = \tilde{b}(\bar{z})\,, \quad \mathrm{Im}\, z = 0\,, \tag{2.7.29}$$

which is the z coordinate form where the boundary is on the real axis. Then we can use the doubling trick to write the holomorphic and anti-holomorphic fields in the upper half plane in the form of holomorphic fields in the entire plane

$$c(z) \equiv \tilde{c}(\bar{z}')\,, \quad b(z) \equiv \tilde{b}(\bar{z}')\,, \quad \mathrm{Im}(z) \leq 0\,,\ z' = \bar{z}\,. \tag{2.7.30}$$

So for open strings, each b and c has a set of Laurent expansions.

2.8 Vertex operators

In quantum field theory, on one hand, there is the state space of the theory; on the other hand, there is the set of local operators. In conformal field theory, there is a simple and useful isomorphism between them when quantizing the CFT on a circle. Examine the semi-infinite cylinder in w-coordinates,

$$0 \leq \mathrm{Re}\, w \leq 2\pi\,, \quad w \sim w + 2\pi\,, \quad \mathrm{Im}\, w \leq 0\,, \tag{2.8.1}$$

under $z = \exp(-iw)$, this semi-infinite cylinder is mapped to a unit disk. As shown in Figure 2.7.

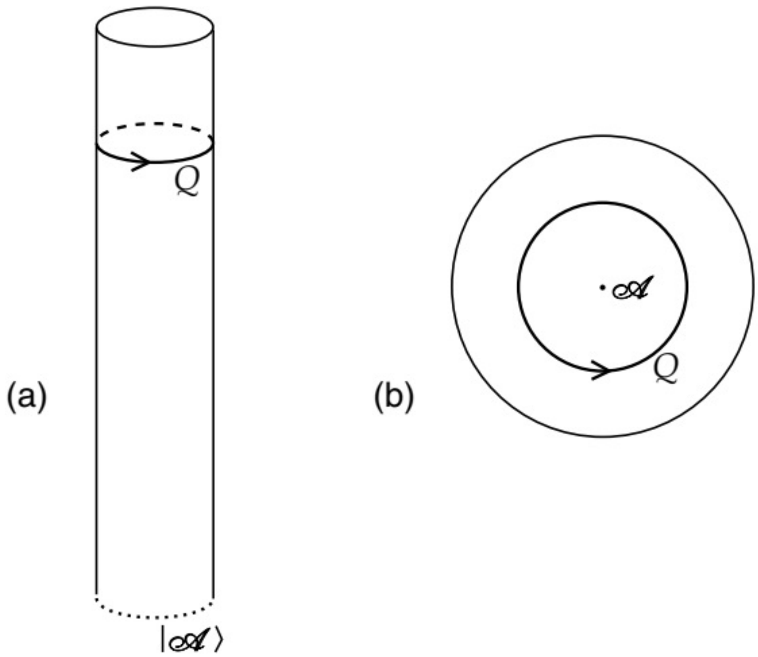


Figure 2.7: (a) Semi-infinite cylinder in w-coordinates, with initial state $|\mathscr{A}\rangle$ and charge Q. (b) Conformal equivalent unit disk, with local operator $\mathscr{A}$ and contour integral of Q.

For free field theory, it is easy to get the detailed form of this isomorphism. Suppose there is a conserved charge Q acting on state $|\mathscr{A}\rangle$ as in Figure 2.7(a). By using the OPE to calculate the contour integral in Figure 2.7(b), the corresponding local operator can be found. Corresponding to the identity operator, we equate the state $|1\rangle$. When there is no operator at the origin, ∂X^μ and $\bar{\partial} X^\mu$ are holomorphic and anti-holomorphic inside the contour of Q in Figure 2.7(b). Define the contour integral (2.7.2) of α^μ_m and $\tilde{\alpha}^\mu_m$ for $m \geq 0$ to have no pole, thus yielding zero. $|1\rangle$ is annihilated by these modes, indicating it is the ground state,

$$|1\rangle = |0;0\rangle \,, \tag{2.8.2}$$

now examine, for example, the state $\alpha^\mu_{-m}|1\rangle$ for positive m. Transitioning to Figure 2.7(b), $Q = \alpha^\mu_{-m}$, the field is holomorphic within the contour, thus for $m \geq 1$ we get

$$\alpha^\mu_{-m} = \left(\frac{2}{\alpha'}\right)^{1/2} \oint \frac{\mathrm{d}z}{2\pi} z^{-m} \partial X^\mu(z) \to \left(\frac{2}{\alpha'}\right)^{1/2} \frac{i}{(m-1)!} \partial^m X^\mu(0) \tag{2.8.3}$$

so

$$\alpha^\mu_{-m}|1\rangle \cong \left(\frac{2}{\alpha'}\right)^{1/2} \frac{i}{(m-1)!} \partial^m X^\mu(0) \,, \quad m \geq 1 \,. \tag{2.8.4}$$

Similarly

$$\tilde{\alpha}^\mu_{-m}|1\rangle \cong \left(\frac{2}{\alpha'}\right)^{1/2} \frac{i}{(m-1)!} \bar{\partial}^m X^\mu(0) \,, \quad m \geq 1 \,. \tag{2.8.5}$$

When α^μ_{-m} or $\tilde{\alpha}^\mu_{-m}$ act on general state $|\mathscr{A}\rangle$, this correspondence still holds. If $:\mathscr{A}(0,0):$ is any normal-ordered operator, there may be singularities in the operator product of $\partial X^\mu(z)$ and $:\mathscr{A}(0,0):$, but it is not hard to verify for $m > 0$

$$\alpha^\mu_{-m} :\mathscr{A}(0,0):=:\alpha^\mu_{-m}\mathscr{A}(0,0): \;. \tag{2.8.6}$$

This is because the contour integral for this contraction will not have a simple singularity, then the same operation (2.8.3) can be performed within the normal ordering, thus states with several α oscillator excitations appear as normal products $: \,:$ of corresponding operators

$$\alpha_{-m} \to i\left(\frac{2}{\alpha'}\right)^{1/2} \frac{1}{(m-1)!} \partial^m X^\mu(0) \,, \quad m \geq 1 \,, \tag{2.8.7a}$$

$$\tilde{\alpha}^\mu_{-m} \to i\left(\frac{2}{\alpha'}\right)^{1/2} \frac{1}{(m-1)!} \bar{\partial}^m X^\mu(0) \,, \quad m \geq 1 \,, \tag{2.8.7b}$$

similarly

$$x^\mu_0 \to X^\mu(0,0) \,. \tag{2.8.8}$$

Any operator can be obtained by acting with operators on the left of (2.8.7) and (2.8.8) on $|1\rangle$. The operator corresponding to that state is given by the normal product $: \,:$ of the corresponding local operator on the right. For example,

$$|0;k\rangle \cong :\mathrm{e}^{ik\cdot X(0,0)}: \;. \tag{2.8.9}$$

This is easy to understand: under the translation $X^\mu \to X^\mu + a^\mu$, the state and the corresponding operator have the same translation, both being multiplied by $\mathrm{e}^{ik\cdot a}$.

One can show that

$$\alpha_m\,|0;k\rangle = 0 \qquad \forall m > 0.$$

and

$$\hat{P}\,|0;k\rangle = k\,|0;k\rangle$$

Proof. By direct calculation:

$$\begin{aligned}
\alpha_m |0;k\rangle &= i\sqrt{\frac{2}{\alpha'}}\oint \frac{\mathrm{d}z}{2\pi i} z^m \partial X(z) : \sum_{n=0}^{\infty} \frac{(ik)^n}{n!} X^n(0) : \\
&= i\sqrt{\frac{2}{\alpha'}}\oint \frac{\mathrm{d}z}{2\pi i} z^m \left(-\frac{ik\alpha'}{2} \frac{:\mathrm{e}^{ikX}:}{z} + \dots \right) && \text{(using (2.3))} \\
&= \sqrt{\frac{\alpha'}{2}} k \,\mathrm{Res}\left[z^m \frac{:\mathrm{e}^{ikX}:}{z} \right] \\
&= 0 \qquad \forall m > 0.
\end{aligned}$$

To show latter we utilize $[\hat{P}, e^{ik\hat{X}}] = ke^{ik\hat{X}}$:

$$\begin{aligned}
\hat{P}|0;k\rangle = \hat{P}e^{ik\hat{X}}|0;0\rangle &= e^{ik\hat{X}}\hat{P}|0;0\rangle + ke^{ik\hat{X}}|0;0\rangle \\
&= k|0;k\rangle
\end{aligned}$$

The same method applies to bc theory. For clarity, we specify the case $\lambda = 2$. This is of interest to bosonic string theory. From Laurent expansion and contour discussion gives

$$b_m|1\rangle = 0, \quad m \geq -1, \qquad c_m|1\rangle = 0, \quad m \geq 2\,. \tag{2.8.10}$$

Note the shifts in the Laurent expansion exponents, which come from the conformal weights of b and c. The identity operator is no longer mapped onto the ground state. Instead, relationship (2.8.11) is determined

$$|1\rangle = b_{-1}|{\downarrow}\rangle\,. \tag{2.8.11}$$

The translation of raising operators is straightforward,

$$b_{-m} \to \frac{1}{(m-2)!}\partial^{m-2}b(0)\,, \quad m \geq 2\,, \tag{2.8.12a}$$

$$c_{-m} \to \frac{1}{(m+1)!}\partial^{m+1}c(0)\,, \quad m \geq -1\,. \tag{2.8.12b}$$

Note that the state $b_{-1}|{\downarrow}\rangle$ with ghost number $-\frac{3}{2}$ is mapped onto the identity operator with ghost number 0, while the state $|{\downarrow}\rangle$ with ghost number $-\frac{1}{2}$ is mapped onto the operator c with ghost number 1. This difference comes from the non-tensor property (2.5.17) of the ghost number current,

$$\partial_z w) j_w(w) = j_z(z) + q_0 \frac{\partial_z^2 w}{\partial_z w} = j_z(z) - \frac{q_0}{z}\,, \tag{2.8.13}$$

where $q_0 = \lambda - \frac{1}{2} = \frac{3}{2}$. The expression for the cylindrical reference frame (2.7.22) typically defines the ghost number of states, and the contour discussion in Figure 2.7 relates the ghost number of vertex operators with the polar coordinate system charge

$$Q^{\mathrm{g}} \equiv \frac{1}{2\pi i}\oint \mathrm{d}z\, j_z = N^{\mathrm{g}} + q_0\,. \tag{2.8.14}$$

This also applies to other charges: for tensor currents, the charge of the state and the corresponding operator are equal.

Most of the above discussion can be extended to $\beta\gamma$ theory, but there are some difficulties under superstring.

All concepts in this section can be extended to open strings. Half-infinite band

$$0 \leq \mathrm{Re}\, w \leq \pi, \quad \mathrm{Im}\, w \leq 0\,, \tag{2.8.15}$$

is mapped to a semi-disk under $z = -\exp(-iw)$, i.e., the overlap region of the upper half plane and the unit circle. Initial state is again mapped to the origin, which is on the boundary, so there is an isomorphism

$$\text{local operator on the boundary } \leftrightarrow \text{state on the boundary} . \tag{2.8.16}$$

Details same as above. Doubling trick is useful in contour discussion.

Path Integral Derivation

Examine a disk in the z-plane, with local operator $\mathscr{A}$ at the origin, the field to be path integrated is ϕ, this field takes fixed boundary value ϕ_b on the unit circle. Path integrating over the internal field ϕ_i while keeping ϕ_b constant gives the functional $\Psi_\mathscr{A}[\phi_\mathrm{b}]$,

$$\Psi_\mathscr{A}[\phi_\mathrm{b}] = \int [\mathrm{d}\phi_\mathrm{i}]_{\phi_\mathrm{b}} \exp(-S[\phi_\mathrm{i}])\mathscr{A}(0) . \tag{2.8.17}$$

The functional of fields is the Schrödinger representation of states, so this is the mapping from operators to states. The Schrödinger representation assigns a complex amplitude to each field configuration, with many uses.

Alternatively, starting from some state $\Psi[\phi_\mathrm{b}]$. Examine path integration in the annular region $1 \geq |z| \geq r$, where the field ϕ_b on the outer circle is kept fixed, and path integrated over ϕ'_b along the inner circle as follows:

$$\int [d\phi_\mathrm{i}]_{\phi_\mathrm{b},\phi'_\mathrm{b}} [\mathrm{d}\phi'_\mathrm{b}] \ \exp(-S[\phi_\mathrm{i}])\, r^{-L_0-\tilde{L}_0} \Psi[\phi'_\mathrm{b}] \tag{2.8.18}$$

that is, this integration is performed under weight $r^{-L_0-\tilde{L}_0}\Psi[\phi'_\mathrm{b}]$. Now, path integration over the disk exactly corresponds to the propagation from $|z| = r$ to $|z| = 1$, which is equivalent to acting with the operator $r^{+L_0+\tilde{L}_0}$. This cancels out the operator acting on Φ, so (2.8.18) is once again $\Psi[\phi_\mathrm{b}]$. Now take the limit $r \to 0$. The annulus becomes a disk, and the limit of path integration over the inner circle can be considered as defining some local operator at the origin. According to this construction, path integration over the disk and this operator reproduce $\Psi[\phi_\mathrm{b}]$ on the boundary.

We use a free scalar field X as an example. Expanded on the unit circle

$$X_\mathrm{b}(\theta) = \sum_{n=-\infty}^{\infty} X_n \mathrm{e}^{in\theta} , \quad X_n^* = X_{-n} . \tag{2.8.19}$$

The boundary state $\Psi[X_b]$ can be thought of as a function of all X_n. First identify the state corresponding to the identity operator, which is given by path integration without inserting any operators

$$\Psi_1[X_\mathrm{b}] = \int [\mathrm{d}X_\mathrm{i}]_{X_\mathrm{b}} \exp\left(-\frac{1}{2\pi\alpha'}\int \mathrm{d}^2z\, \partial X \bar{\partial} X\right) . \tag{2.8.20}$$

Calculate it using the usual Gaussian method. Decompose X_i into classical part and fluctuation

$$X_\mathrm{i} = X_\mathrm{cl} + X'_\mathrm{i} , \tag{2.8.21a}$$

$$X_\mathrm{cl}(z,\bar{z}) = X_0 + \sum_{n=1}^{\infty} (z^n X_n + \bar{z}^n X_{-n}) . \tag{2.8.21b}$$

Under this definition, X_cl satisfies the equations of motion, X'_i is zero on the boundary. Then the path integral becomes

$$\Psi_1[X_\mathrm{b}] = \exp(-S_\mathrm{cl}) \int [\mathrm{d}X'_\mathrm{i}]_{X_\mathrm{b}=0} \exp\left(-\frac{1}{2\pi\alpha'}\int \mathrm{d}^2z\, \partial X' \bar{\partial} X'\right) , \tag{2.8.22}$$

where

$$S_{\mathrm{cl}} = \frac{1}{2\pi\alpha'}\sum_{m,n=1}^{\infty} mnX_mX_{-n}\int_{|z|<1} \mathrm{d}^2z\, z^{m-1}\bar{z}^{n-1}$$
$$= \frac{1}{\alpha'}\sum_{m=1}^{\infty} mX_mX_{-m}\,. \tag{2.8.23}$$

The integral over X_i' is a constant independent of boundary conditions, so

$$\Psi_1[X_{\mathrm{b}}] \propto \exp\left(-\frac{1}{\alpha'}\sum_{m=1}^{\infty} mX_mX_{-m}\right)\,. \tag{2.8.24}$$

This is Gaussian type, actually the ground state. To see this, write raising and lowering operators in Schrödinger basis,

$$\alpha_n = -\frac{in}{(2\alpha')^{1/2}}X_{-n} - i\left(\frac{\alpha'}{2}\right)^{1/2}\frac{\partial}{\partial X_n}\,, \tag{2.8.25a}$$
$$\tilde{\alpha}_n = -\frac{in}{(2\alpha')^{1/2}}X_{n} - i\left(\frac{\alpha'}{2}\right)^{1/2}\frac{\partial}{\partial X_{-n}}\,, \tag{2.8.25b}$$

which comes from the Laurent expansion at $|z| = 1$ and the mode algebra. Applying them on (2.8.24), we find

$$\alpha_n\Psi_1[X_{\mathrm{b}}] = \tilde{\alpha}_n\Psi_1[X_{\mathrm{b}}] = 0\,, \quad n \geq 0\,, \tag{2.8.26}$$

so this is indeed the ground state. Therefore

$$|1\rangle \propto |0;0\rangle\,. \tag{2.8.27}$$

We do not track the total normalization factor of the path integral, but define $|1\rangle = |0;0\rangle$.

Another simple calculation is the state corresponding to $\partial^k X$; this is just adding a factor $\partial^k X_{\mathrm{cl}}(0) = k!X_k$ to existing result, so

$$|\partial^k X\rangle = k!X_k\Psi_1 = -i\left(\frac{\alpha'}{2}\right)^{1/2}(k-1)!\alpha_{-k}|0;0\rangle\,, \tag{2.8.28}$$

similar for $\bar{\partial}^k X$. This can be generalized to all products of derivatives and exponentials of X. Conformal normal ordering only cancels out effect of X_{i}' path integral.

2.9 More on states and operators

OPE

In this section we make some other applications of the state-operator correspondence. The first is a general extension of the OPE, as shown in Figure 2.8. Examine product

$$\mathscr{A}_i(z,\bar{z})\mathscr{A}_j(0,0)\,, \tag{2.9.1}$$

where $|z| < 1$. We can split the path integral into integration over field ϕ_{i} inside unit circle, field ϕ_{b} on unit circle, field ϕ_{e} outside unit circle. As discussed at the end of last section, integration over ϕ_{i} yields some functional of ϕ_{b}, denoted as $\Psi_{ij,z,\bar{z}}[\phi_{\mathrm{b}}]$. Via state–operator correspondence, this is equivalent to gluing disk together with appropriate operator $\mathscr{A}$. To turn it into standard form expanded with complete set of $L_0, \tilde{L}_0$ eigenstates

$$\mathscr{A}_{ij,z,\bar{z}} = \sum_k z^{h_k-h_i-h_j}\bar{z}^{\tilde{h}_k-\tilde{h}_i-\tilde{h}_j}c^k{}_{ij}\mathscr{A}_k\,, \tag{2.9.2}$$

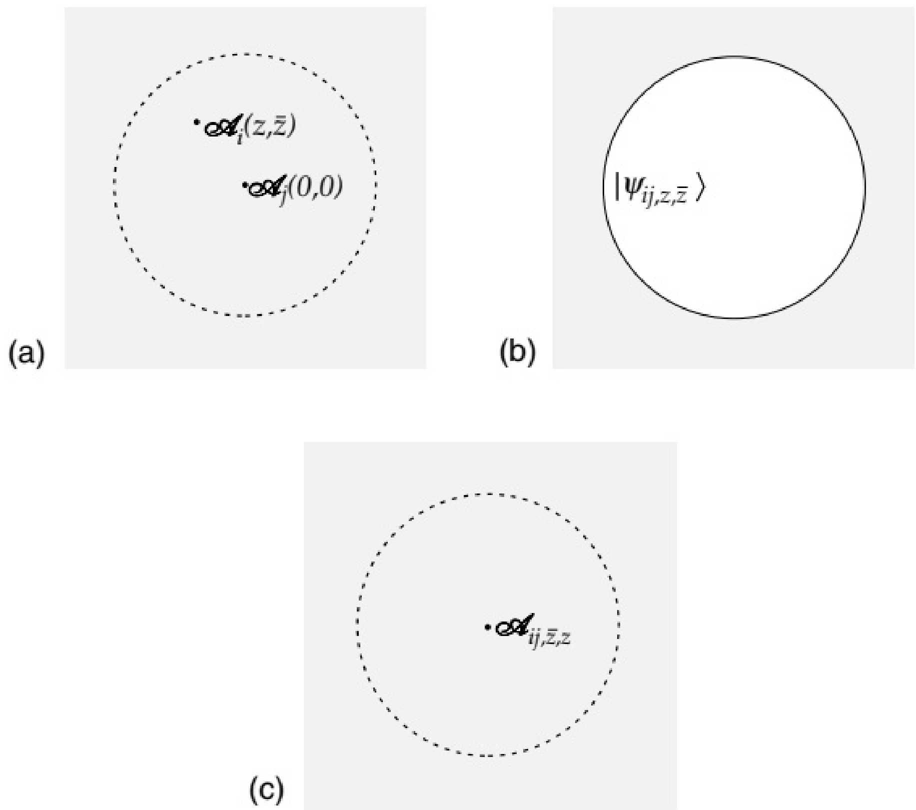


Figure 2.8: (a) Worldsheet with two local operators. (b) Integrating over fields inside the disk gives boundary state $|\psi_{ij,z,\bar{z}}\rangle$. (c) Cutout inside a disk with corresponding local operators. Expanding to operators with definite weights gives OPE.

where z dependence and $\bar{z}$ dependence are determined by conformal weights like (2.4.21). Convergence of OPE is exactly convergence of a complete set in quantum mechanics. As long as there is no other operator in $|z'| \leq |z|$, this construction is possible. Allowing us to cut out circle of radius $|z| + \epsilon$. For three operators

$$\mathscr{A}_i(0,0)\mathscr{A}_j(1,1)\mathscr{A}_k(z,\bar{z}) \,, \tag{2.9.3}$$

convergence region $|z| < 1$ of $z \to 0$ OPE overlaps with convergence region $|1-z| < 1$ of $z \to 1$ OPE. Then coefficient of $\mathscr{A}_l$ in triple product can be written as summation including $c^l{}_{ik}c^m{}_{lj}$ or summation including $c^l{}_{jk}c^m{}_{li}$. Associativity requires these sums to be equal. Graphic representation see Figure 2.9.

Virasoro Algebra and Highest Weight States

Now apply discussion in Figure 2.7 to Virasoro generators:

$$\begin{aligned} L_m|\mathscr{A}\rangle &\cong \oint \frac{\mathrm{d}z}{2\pi i}\, z^{m+1}T(z)\mathscr{A}(0,0) \\ &\cong L_m \cdot \mathscr{A}(0,0) \,. \end{aligned} \tag{2.9.4}$$

Here we entered new notation and concept. Given state-operator isomorphism, each operator acting on Hilbert space has an image acting on local operator space. In other words, formed corresponding contour around operator and calculated corresponding local operator using OPE. Corresponding to L_m mapping states to states is $L_m\cdot$ mapping local operators to local operators.

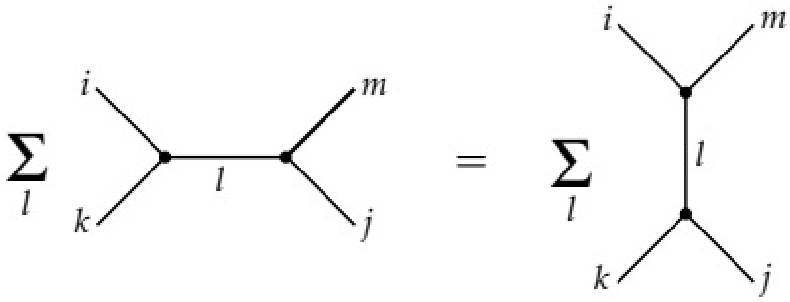


Figure 2.9: Graphical representation of OPE associativity

Generally there will be local operators at different positions z_i, and each position will have a different version of Virasoro algebra. These different versions are obtained by Laurent expansion centered at operator position. A standard basis for generators will also exist in some geometries. For example for cylinder then expansion centered at $z = 0$. "$\cdot$" acts as a reminder: what we discuss is Laurent expansion coefficient around specific operator.

In this notation, $T\mathscr{A}$ OPE is

$$T(z)\mathscr{A}(0,0) = \sum_{m=-\infty}^{\infty} z^{-m-2} L_m \cdot \mathscr{A}(0,0)\,. \tag{2.9.5}$$

To relate to earlier notation (2.4.11), for $n \geq 0$, $\mathscr{A}^{(n)} = L_{n-1} \cdot \mathscr{A}$, conformal transformation of $\mathscr{A}$ is

$$\delta\mathscr{A}(z,\bar{z}) = -\epsilon \sum_{n=0}^{\infty} \frac{1}{n!} \left[\partial^n v(z) L_{n-1} + (\partial^n v(z))^* \tilde{L}_{n-1}\right] \cdot \mathscr{A}(z,\bar{z})\,. \tag{2.9.6}$$

Using general form of $T\mathscr{A}$ OPE (2.4.14), we have very useful results

$$L_{-1} \cdot \mathscr{A} = \partial\mathscr{A}, \quad \tilde{L}_{-1} \cdot \mathscr{A} = \bar{\partial}\mathscr{A}\,, \tag{2.9.7a}$$

$$L_0 \cdot \mathscr{A} = h\mathscr{A}, \quad \tilde{L}_0 \cdot \mathscr{A} = \tilde{h}\mathscr{A}\,. \tag{2.9.7b}$$

OPE (2.9.5) implies state corresponding to primary field $\mathcal{O}$ with weight $(h, \tilde{h})$ satisfies

$$L_0|\mathcal{O}\rangle = h|\mathcal{O}\rangle\,, \quad \tilde{L}_0|\mathcal{O}\rangle = \tilde{h}|\mathcal{O}\rangle\,, \tag{2.9.8a}$$

$$L_m|\mathcal{O}\rangle = \tilde{L}_m|\mathcal{O}\rangle = 0\,, \quad m > 0\,. \tag{2.9.8b}$$

Such state is called highest weight state: by repeatedly acting with lowering operators on arbitrary state, we eventually obtain state annihilated by further lowering operators, called highest weight state.

An interesting special case is identity operator, operator product $T1$ is non-singular, so in arbitrary CFT get

$$L_m|1\rangle = \tilde{L}_m|1\rangle = 0, \quad m \geq -1\,. \tag{2.9.9}$$

As marked in 2.4, operators L_0 and $L_{\pm 1}$ form closed algebra, so do $\tilde{L}_0$ and $\tilde{L}_{\pm 1}$. Entire algebra is

$$SL(2,\mathbb{R}) \times SL(2,\mathbb{R}) = SL(2,\mathbb{C})\,. \tag{2.9.10}$$

Therefore $|1\rangle$ is also called $SL(2,\mathbb{C})$-invariant state. It is the only such state, because (2.9.7) implies any operator $\mathscr{A}$ corresponding to $SL(2,\mathbb{C})$-invariant state is independent of position, so must be c-number, as explained after (2.4.25).

Unitary CFT

Highest weight states play important role in string theory and representation theory of Virasoro algebra. Now we derive several important results generally valid in unitary CFT. Unitary CFT has positive definite inner product, making

$$L_m^\dagger = L_{-m}\,, \quad \tilde{L}_m^\dagger = \tilde{L}_{-m}\,. \tag{2.9.11}$$

Recall: inner product via $\langle\alpha|A\beta\rangle = \langle A^\dagger\alpha|\beta\rangle$ defined adjoint. For example X^μ CFT for *spacelike* μ is unitary. If we take inner product of ground states as

$$\langle 0;k|0;k'\rangle = 2\pi\delta(k-k') \tag{2.9.12}$$

and define

$$\alpha_m^\dagger = \alpha_{-m}\,, \quad \tilde{\alpha}_m^\dagger = \tilde{\alpha}_{-m}\,. \tag{2.9.13}$$

This implicitly defines inner product of all higher states. This CFT for $X^\mu = 0$ is not unitary, because there is negative sign in commutation relation there.

First constraint in unitary CFT is any state must have $h, \tilde{h} \geq 0$. If this holds for highest weight states, then it holds for all states. For highest weight states, Virasoro algebra gives

$$2h_\mathcal{O}\langle\mathcal{O}|\mathcal{O}\rangle = 2\langle\mathcal{O}|L_0|\mathcal{O}\rangle = \langle\mathcal{O}|[L_1, L_{-1}]|\mathcal{O}\rangle = \|L_{-1}|\mathcal{O}\rangle\|^2 \geq 0\,, \tag{2.9.14}$$

so $h_\mathcal{O} \geq 0$. If $h_\mathcal{O} = \tilde{h}_\mathcal{O} = 0$, then

$$L_{-1}\cdot\mathcal{O} = \tilde{L}_{-1}\cdot\mathcal{O} = 0\,, \tag{2.9.15}$$

thus $\mathcal{O}$ is independent of position. As previously marked, $\mathcal{O}$ must be c-number, i.e., in unitary CFT identity operator is unique $(0,0)$ operator. Surprisingly, X^μ is unique exception: corresponding state $x^\mu|0;0\rangle$ is non-normalizable due to infinite range of X^μ, and equation (2.9.14) no longer holds. For CFT corresponding to compact dimension, general theorem is most useful, such exceptions cannot occur.

In same way, in unitary CFT, if and only if $\tilde{h} = 0$, operator is holomorphic, and if and only if $h = 0$, operator is anti-holomorphic; this is an important result, repeated as follows:

$$\partial\mathscr{A} = 0 \Leftrightarrow h = 0\,, \quad \bar{\partial}\mathscr{A} = 0 \Leftrightarrow \tilde{h} = 0\,. \tag{2.9.16}$$

Finally, using previous discussion and commutator $[L_n, L_{-n}]$ it can be proven that in unitary CFT $c, \tilde{c} \geq 0$. Actually, unique unitary CFT with $c = 0$ is trivial: $L_n = 0$.

Zero-point energy

State-operator mapping gives another simple derivation of different normal-ordering constants. In any CFT, we know $L_0|1\rangle = 0$, and this determines another normalization of L_0. In X CFT, $|1\rangle = 0$ is ground state, so $a^X = 0$. In bc theory, $|1\rangle = 0$ is excited state $b_{-1}|{\downarrow}\rangle$, so to be consistent with earlier $\lambda = 2$ result (2.7.1), weight of $|{\downarrow}\rangle$ is -1.

This also provides physical interpretation of central charge. In unitary CFT, ground state is $|1\rangle$ with $L_0 = \tilde{L}_0 = 0$. Conformal transformation (2.6.10) between radial generator and time translation generator implies for ground state

$$E = -\frac{c+\tilde{c}}{24}\,. \tag{2.9.17}$$

This is Casimir energy, arising from finite scale of system, and dependent on central charge. Based on dimensional analysis, this energy is inversely proportional to size ℓ of system, here 2π, so general result is

$$E = -\frac{\pi(c+\tilde{c})}{12\ell}\,. \tag{2.9.18}$$

We now have three ways to calculate normal-ordering constants:

1. Like in (2.7.8), starting from Virasoro algebra;
2. Like in (2.7.11), relating two forms of normal ordering;
3. Starting from previous state-operator mapping.

However method of adding up zero-point energy is intuitive, so we give another description:

1. For each boson mode, zero-point energy is $\frac{1}{2}\omega$; for each fermion mode, zero-point energy is $-\frac{1}{2}\omega$. Add up energy.
2. Encountering divergent sum of form $\sum_{n=1}^{\infty}(n-\theta)$, where θ comes from non-trivial periodicity conditions, define it as
$$\sum_{n=1}^{\infty}(n-\theta) = \frac{1}{24} - \frac{1}{8}(2\theta-1)^2 \,. \tag{2.9.19}$$
This value can be obtained, as in (1.3.32), by regularizing and discarding the quadratically divergent part.
3. The normal ordering constant for the generator T_0 in the w frame was provided earlier, where T_0 is defined by (2.6.8). For L_0, we must add the non-tensor correction $\frac{1}{24}c$.

For the free bosonic field, the modes are integer-valued; thus, after step 2, for $\theta = 0$, we obtain half of the sum, which is $-\frac{1}{24}$. This is exactly canceled by the correction in step 3, giving $a^X = 0$. For the ghost fields, we similarly obtain $\frac{2}{24} - \frac{26}{24} = -1$.

Chapter 3

The Polyakov path integral

We now undertake a systematic study of string theory using the Polyakov path integral and conformal field theory. After introducing the path integral picture, we discuss gauge fixing, the Weyl anomaly, and vertex operators. Subsequently, we will generalize string theory to curved spacetime.

3.1 Sums over world-sheets

The Feynman path integral is one way to represent quantum theory, and it is a very natural method for describing string interactions. In path integral quantization, amplitudes are obtained by summing over all possible histories between initial and final states. Each history is assigned a weight

$$\exp(iS_{\rm cl}/\hbar)\,, \tag{3.1.1}$$

where $S_{\rm cl}$ is the classical action for the given history. Thus, summing over all worldsheets connecting given initial and final curves defines the amplitude in string theory. Figure 3.1a refers to open strings, and Figure 3.1b to closed strings. A similar sum for a relativistic point particle yields the free propagator.

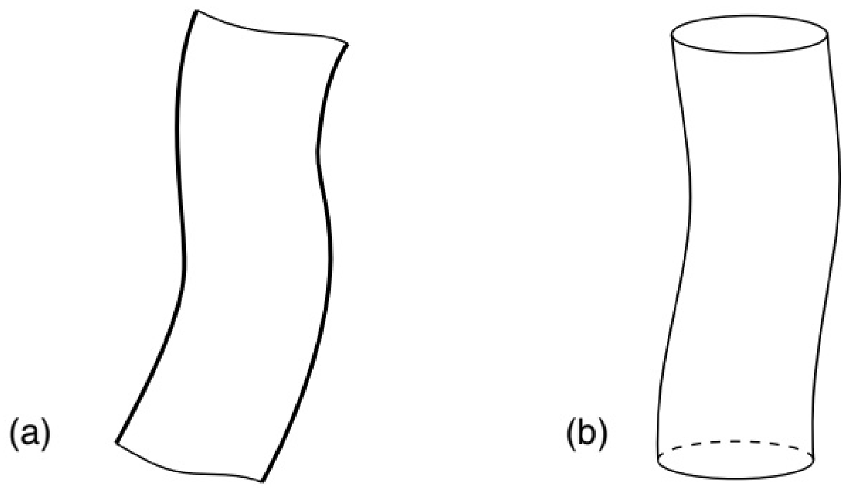


Figure 3.1: (a) Open string worldsheet with the topology of a long strip. The bold curves are the worldlines of the string endpoints. (b) Closed string worldsheet with the topology of a cylinder.

One can imagine several ways for strings to interact. One is a contact interaction, the energy when two strings intersect; another is through long-range forces via some quantum field.

However, interactions added manually to string theory cannot be consistent with symmetry; we will see the reasons for this later. Conversely, the only allowed interaction is already implicit in the sum over worldsheets. For example, consider the worldsheets shown in Figure 3.2. Figure 3.2a looks like a quantum correction to the open string amplitude in Figure 3.1a. This correction includes an intermediate state with two open strings. Figure 3.2b has three external closed strings, representing a string decaying into two. From this, we see that this is a correct way to introduce interactions in strings. We will see that these interactions produce consistent theories containing gravity that are finite and unitary.

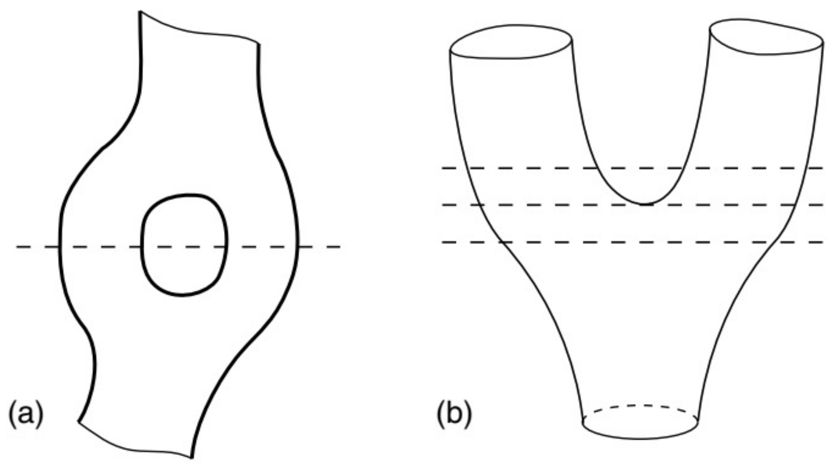


Figure 3.2: (a) Quantum correction to the open string propagator. (b) A closed string decaying into two. Dashed lines represent time slices.

It is interesting to examine the process of Figure 3.2b at several consecutive moments, as shown in Figure 3.3. In closed string theory, all particles are obtained as different vibrational states of the string. All interactions (gauge, gravitational, Yukawa) originate from the single process in Figure 3.3. For theories with open strings, there are several additional possible processes, as shown in Figure 3.4. At a given moment, the processes in Figures 3.3 and 3.4 occur at definite spacetime points, but this is an illusion. A different slicing places the apparent interaction at different points; there are no distinguishable points in spacetime. Interactions originate only from the global topology of the worldsheet, while the local properties of the worldsheet are the same as in the free case. As discussed in the introduction, this "smeared out" interaction cuts off short-distance divergences in gravity.

In fact, there are several different string theories, depending on the topologies introduced in the sum over worldsheets. Note that the worldsheets we have drawn have two types of boundaries: source boundaries corresponding to the initial and final string configurations, and endpoint boundaries corresponding to the worldlines of the open string ends. For now, ignore the source boundaries; we will discuss source details later in this chapter. Therefore, in subsequent discussions, "boundary" refers to the endpoint boundary. Thus, there are 4 ways to define the sum over worldsheets, corresponding to the 4 free string theories listed at the end of Section 1.4:

1. Closed oriented: all oriented worldsheets without boundaries;
2. Closed unoriented: all worldsheets without boundaries;
3. Closed plus open oriented: all oriented worldsheets with any number of boundaries;
4. Closed plus open unoriented: all worldsheets with any number of boundaries;

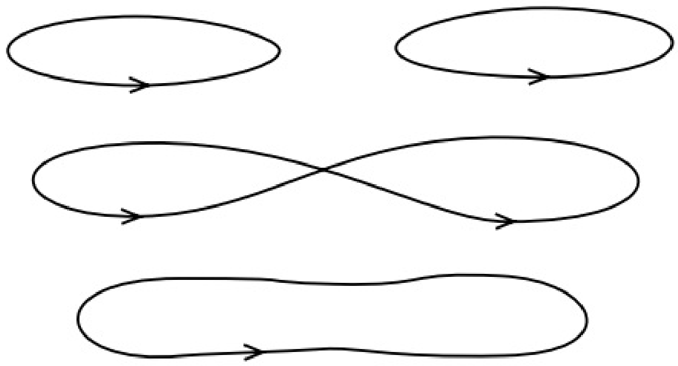


Figure 3.3: Several consecutive slices of Figure 3.2b. Arrows indicate the orientation of the string.

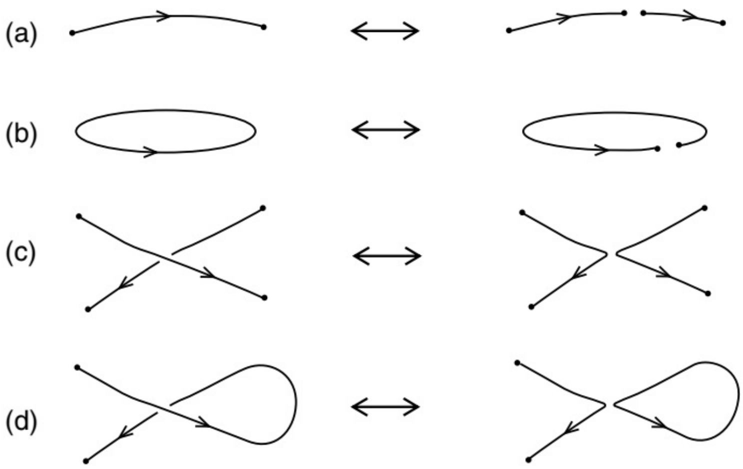


Figure 3.4: Processes involving open strings: (a) open string $\leftrightarrow$ open string + open string; (b) closed string $\leftrightarrow$ open string; (c) open string + open string $\leftrightarrow$ open string + open string; (d) open string $\leftrightarrow$ open string + closed string.

There are two things to elaborate on here. One is the connection between the inclusion of unoriented worldsheets and the $\Omega = +1$ projection described in Section 1.4. Figure 3.5 shows two worldsheets counted in the open unoriented theory. The worldsheet in Figure 3.5b is equivalent to cutting Figure 3.5a along the dashed line. We parameterize it with $0 \leq \sigma \leq \pi$ and require the fields at the cut to satisfy

$$X^{\mu}_{\text{upper}}(\sigma) = X^{\mu}_{\text{lower}}(\pi - \sigma) . \tag{3.1.2}$$

This in turn is equivalent to acting with the operator Ω on the open string states. Summing these two surfaces with appropriate weights is equivalent to inserting the operator $\frac{1}{2}(1+\Omega)$, which projects onto $\Omega = +1$ states. In the sum over all oriented and unoriented surfaces, inserting projection operators on all intermediate states reduces the spectrum to the $\Omega = +1$ part.

Second, the list above does not include a theory with only open strings. Let's explain in detail why this is the case. Consider a worldsheet with the topology of an annulus, as shown in Figure 3.6a. For convenience, we have only drawn the vacuum amplitude, but the same discussion

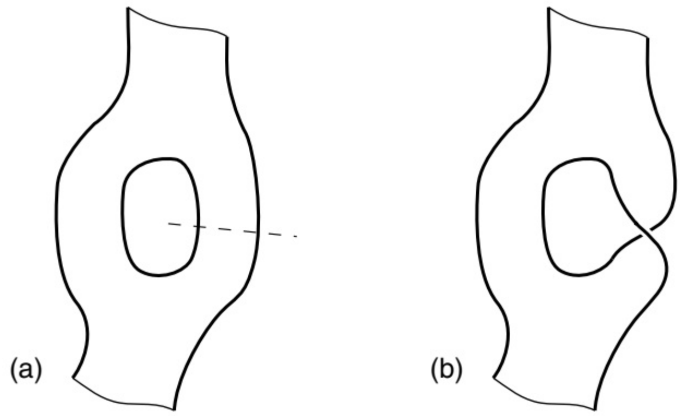


Figure 3.5: Oriented (a) and unoriented (b) contributions to the 2-open-string amplitude.

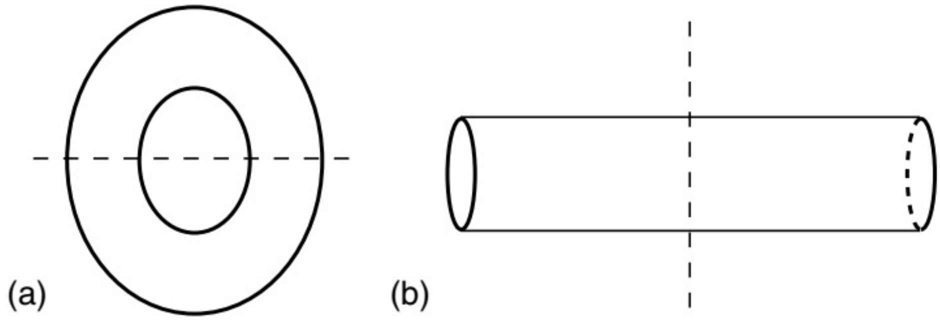


Figure 3.6: (a) Annulus: the intermediate state consists of two open strings. (b) Cylinder with the same topology: the intermediate state is a single closed string.

applies to scattering amplitudes by connecting appropriate sources. As can be seen from the dashed lines, this is a process with two open strings in the intermediate state. Thus, worldsheets of this topology must appear in any theory with open strings. In Figure 3.6b, we draw the same topology as a cylinder and cut it at an intermediate state of a single closed string. Therefore, even if we start only with open strings, the sum over worldsheets will inevitably introduce such processes—the scattering of open strings produces closed strings. As we study the divergence of amplitudes, we will see in detail how this works.

The necessity of closed strings in open string theory can be seen from another perspective. Consider the interactions in Figure 3.4a and b, but time-reversed: two open strings merge into one, and an open string becomes closed. Joining them at the interaction point is the same; merging two ends allows the first interaction but not the second. This would require attaching some non-local constraints to the dynamics of the string, which would prove to be inconsistent. The same discussion applies to Figures 3.4c and d, where the interaction is the local reconnection of a pair of strings; thus, if any open string interaction is allowed, the same process producing a closed string from open strings is also allowed.

3.2 The Polyakov path integral

We now begin to study the sum over worldsheets, expressed as a functional integral in the Polyakov formalism. There is a difference here from Chapter 1: the Minkowski metric γ_{ab} is replaced by the Euclidean metric $g_{ab}(\sigma^1, \sigma^2)$ with signature $(+,+)$. The integral is taken over all Euclidean metrics and the embedding $X^\mu(\sigma^1, \sigma^2)$ of the worldsheet in Minkowski spacetime:

$$\int [\mathrm{d}X \mathrm{d}g] \exp(-S) \,. \tag{3.2.1}$$

The Euclidean action is

$$S = S_X + \lambda\chi \,, \tag{3.2.2}$$

where

$$S_X = \frac{1}{4\pi\alpha'} \int_M \mathrm{d}^2\sigma\, g^{1/2} g^{ab} \partial_a X^\mu \partial_b X_\mu \,, \tag{3.2.3a}$$

$$\chi = \frac{1}{4\pi} \int_M \mathrm{d}^2\sigma\, g^{1/2} R + \frac{1}{2\pi} \int_{\partial M} \mathrm{d}s\, k \,. \tag{3.2.3b}$$

Here $\mathrm{d}s$ is the proper time along the boundary, and k is the geodesic curvature of the boundary $k = \pm t^a n_b \nabla_a t^b$.

Remark. *t^a is a unit vector tangent to the boundary, and n^a is a unit vector pointing outward and perpendicular to t^a; the $+$ sign corresponds to a timelike boundary, and $-$ to a spacelike boundary. Here the boundary is always spacelike.*

Proof. We wish to show that the Euler term

$$\chi = \frac{1}{4\pi} \int_M d^2\sigma \sqrt{g} R + \frac{1}{2\pi} \int_{\partial M} ds\, k \tag{3.2.4}$$

with k the geodesic curvature given by

$$k = \pm t^a n_b \nabla_a t^b \tag{3.2.5}$$

is invariant under a Weyl rescaling. From (1.2.32) we know that under a Weyl rescaling $g'_{ab} = e^{2\omega(\sigma)} g_{ab}$ the Riemann curvature transforms, in Euclidean spacetime, as

$$\sqrt{g'} R' = \sqrt{g}(R - 2\nabla^2\omega) = \sqrt{g}(R - 2\nabla^a \partial_a \omega) \tag{3.2.6}$$

We have used the fact that ω is a scalar so the $\nabla_a \omega = \partial_a \omega$. We also know how a connection transforms under a Weyl rescaling

$$\begin{aligned} \Gamma'^{\,a}_{\ bc} &= \Gamma^a_{bc} + g^{ad}(g_{cd}\partial_b\omega + g_{bd}\partial_c\omega - g_{bc}\partial_d\omega) \\ &= \Gamma^a_{bc} + \delta^a_c \partial_b\omega + \delta^a_b \partial_c\omega - \gamma^{ad}\gamma_{bc}\partial_d\omega \end{aligned} \tag{3.2.7}$$

The tangent and normal vectors t^a and n^a are normalised, $g_{ab} t^a t^b = \mp 1$ (with the minus sign for timelike boundaries and the plus sign for spacelike boundaries). Therefore

$$\mp 1 = g'_{ab} t'^a t'^b = e^{2\omega} t'^a t'^b \tag{3.2.8}$$

and similarly for n^a. This can be satisfied if

$$t'^a = e^{-\omega} t^a \quad \text{and} \quad n'^a = e^{-\omega} n^a \tag{3.2.9}$$

From this we find that

$$n'^a = g'_{ab} n^b = e^{2\omega} \gamma_{ab} e^{-\omega} n^b = e^{\omega} n_a \tag{3.2.10}$$

The term with the geodesic curvature then becomes

$$\begin{aligned} k' &= \pm t'^a n_b \nabla_a t'^b = \pm t'^a n_b (\partial_a t'^b + \Gamma^b_{ac} t'^c) \\ &= \pm e^{-\omega} t^a e^{\omega} n_b \left[\partial_a (e^{-\omega} t^b) + (\Gamma^b_{ac} + \delta^b_c \partial_a \omega + \delta^b_a \partial_c \omega - \gamma^{bd} \gamma_{ac} \partial_d \omega) e^{-\omega} t^c \right] \\ &= \pm t^n b e^{-\omega} \left(-\partial_a \omega t^b + \partial_a t^b + \Gamma^b_{ac} t^c + \partial_a \omega t^b + \delta^b_a \partial_c \omega t^c - \gamma^{bd} \gamma_{ac} \partial_d \omega t^c \right) \\ &= e^{-\omega} (k \mp t^n b \gamma^{bd} \gamma_{ac} \partial_d \omega t^c) = e^{-\omega} (k \mp t^n a n_b \partial^b \omega) \end{aligned} \tag{3.2.11}$$

where we have used that the tangent and normal vectors are orthonormal $t^a n_a = 0$. Now, if $t^a t_a = -1$ then we are to chose the upper sign, which is minus and so the second term in the brackets gets a plus sign. If $t^a t_a = +1$ we need to take the lower sign and we once again find a plus sign for the second term. Thus

$$k' = e^{-\omega}(k + n^a \partial_a \omega) \tag{3.2.12}$$

It remains to work out the transformation of ds. We have

$$ds'^2 = g'_{ab} dx^a dx^b = e^{2\omega} dx^a dx^b = e^{2\omega} ds^2 \tag{3.2.13}$$

and thus

$$ds' = e^{\omega} ds \tag{3.2.14}$$

Bringing everything together we have

$$\begin{aligned} \chi' &= \frac{1}{4\pi} \int_M d^2\sigma \sqrt{g'} R' + \frac{1}{2\pi} \int_{\partial M} ds' k' \\ &= \frac{1}{4\pi} \int_M d^2\sigma \sqrt{g} (R - 2\nabla^2 \omega) + \frac{1}{2\pi} \int_{\partial M} e^{\omega} ds e^{-\omega} (k + n^a \partial_a \omega) \\ &= \frac{1}{4\pi} \int_M d^2\sigma \sqrt{g} R + \frac{1}{2\pi} \int_{\partial M} ds k + \frac{1}{2\pi} \left[- \int_M d^2\sigma \sqrt{g} \nabla^a \partial_a \omega + \int_{\partial M} ds n^a \partial_a \omega \right] \\ &= \chi \end{aligned} \tag{3.2.15}$$

where the term between brackets vanishes due to Stokes' theorem

$$\int_M d^2\sigma \sqrt{g} \nabla^a v_a = \int_{\partial M} ds n^a v_a \tag{3.2.16}$$

□

The advantage of the Euclidean path integral is that the integral over the metric is better defined. The topologically non-trivial worldsheets we describe can have non-singular Euclidean metrics, whereas Minkowski metrics cannot. We will take the Euclidean theory (3.2.1) as the starting point. We will show how to give it a precise definition and that it defines a consistent spacetime theory. However, we will give a brief formal discussion that it is equivalent to our starting point Minkowski theory.

We start with the point particle example, where the path integral

$$\int [\mathrm{d}X] \exp\left[iS_{\mathrm{NG}}\right] = \int [\mathrm{d}\eta\, \mathrm{d}X] \exp\left[\frac{i}{2} \int \mathrm{d}\tau \left(\eta^{-1} \dot{X}^\mu \dot{X}_\mu - \eta m^2 \right)\right] \tag{3.2.17}$$

is oscillatory.

Both Nambu-Goto and Polyakov action have gauge redundancy, therefore to show that the partition function are equivalent we define the Euclidean action for a free relativistic particle of mass m propagating in a D-dimensional flat spacetime using the auxiliary einbein formalism:

$$S[X,e] = \int_0^1 d\tau \left(\frac{\dot{X}^2}{2e(\tau)} + \frac{1}{2}e(\tau)m^2 \right) \tag{3.2.18}$$

where:

- $\tau \in [0,1]$ is the arbitrary parameter along the worldline.
- $X^\mu(\tau)$ are the target-space coordinates.
- $\dot{X}^2 = \delta_{\mu\nu}\frac{dX^\mu}{d\tau}\frac{dX^\nu}{d\tau}$ is the squared worldline velocity.
- $e(\tau)$ is the einbein (the 1D independent metric component, intrinsically local).

The path integral requires integrating over all field configurations.

$$Z = \int \mathcal{D}x\,\mathcal{D}p\,\mathcal{D}e\ \exp\left(-\int d\tau \left(p_\mu \dot{x}^\mu - \frac{e}{2}p^2 - \frac{e}{2}m^2\right)\right).$$

Often the partition function is given as integral over x and e, therefore the integral over p sets the integration measure for e in the following manner.

$$\begin{aligned}
\mathcal{D}e\mathcal{D}p\ \exp\left\{\int d\tau \left[-p_\mu \dot{X}^\mu + \frac{e}{2}(p^2+m^2)\right]\right\} &= \mathcal{D}e\prod_{i=1}^{N}\int dp_i \exp\left(-p_{\mu\, i}\dot{X}_i^\mu + \frac{e_i}{2}(p_i^2+m^2)\right) \\
&= \mathcal{D}e\prod_{i=1}^{N}\left(\frac{2\pi}{e_i}\right)^{1/2}\exp\left\{\frac{\dot{X}_i^2}{2e_i} + \frac{e_i}{2}m^2\right\} \\
&= \mathcal{N}\prod_{i=1}^{N}\exp\left\{\frac{\dot{X}_i^2}{2e_i} + \frac{e_i}{2}m^2\right\}e_i^{-1/2}de_i
\end{aligned}$$

We will keep the factor of $\mathcal{N}$ implicit and keep absorbing all the constants in the normalization until at the end. The factor of $e_i^{-1/2}$ coming from path integral over momenta is absorbed into $\mathcal{D}e$ as integration measure. The total partition function evaluating the auxiliary field is:

$$Z_e[X] = \int \mathcal{D}e \exp\left(-S[X,e]\right) \tag{3.2.19}$$

To perform the exact functional integration, we must regularize the continuous integral by discretizing the worldline τ into N intervals, each of uniform width $\epsilon = d\tau = 1/N$. The fields are evaluated at discrete slices i: $\int \mathrm{d}\tau \to \sum_{i=1}^N \epsilon$, $e(\tau) \to e_i$ and $\dot{X}(\tau) \to \dot{X}_i$. The continuum action S and measure $\mathcal{D}e$ become discrete limits:

$$S_N = \sum_{i=1}^{N}\epsilon\left(\frac{\dot{X}_i^2}{2e_i} + \frac{1}{2}e_i m^2\right) \tag{3.2.20}$$

$$\mathcal{D}e_N = \prod_{i=1}^{N} e_i^{-1/2}de_i \tag{3.2.21}$$

Upon Substituting the discrete forms back into the partition function, the exponential factorizes into a product of independent, 1D integrals evaluated at each slice:

$$Z_{e,N}[X] = \prod_{i=1}^{N}\left[\int_0^\infty \frac{de_i}{\sqrt{e_i}}\exp\left(-\frac{\epsilon\dot{X}_i^2}{2e_i} - \frac{\epsilon m^2 e_i}{2}\right)\right] \tag{3.2.22}$$

We evaluate the integral enclosed in the square brackets using the exact mathematical identity for the modified Bessel function of the second kind:

$$\int_0^\infty x^{-1/2} \exp\left(-\frac{A}{x} - Bx\right) dx = \sqrt{\frac{\pi}{B}} \exp\left(-2\sqrt{AB}\right) \tag{3.2.23}$$

For a single local slice i, we define the constants $A = \frac{\epsilon \dot{X}_i^2}{2}$ and $B = \frac{\epsilon m^2}{2}$.
We evaluate the exponent for the slice:

$$-2\sqrt{AB} = -2\sqrt{\left(\frac{\epsilon \dot{X}_i^2}{2}\right)\left(\frac{\epsilon m^2}{2}\right)} = -2\sqrt{\frac{\epsilon^2 m^2 \dot{X}_i^2}{4}} = -m\epsilon\sqrt{\dot{X}_i^2} \tag{3.2.24}$$

We evaluate the prefactor for the slice:

$$\sqrt{\frac{\pi}{B}} = \sqrt{\frac{2\pi}{\epsilon m^2}} \tag{3.2.25}$$

Substitute the evaluated local terms back into the global discrete product:

$$Z_{e,N}[X] = \prod_{i=1}^N \left[\sqrt{\frac{2\pi}{\epsilon m^2}} \exp\left(-m\epsilon\sqrt{\dot{X}_i^2}\right)\right] \tag{3.2.26}$$

Separate the prefactors from the exponential sum:

$$Z_{e,N}[X] = \left(\prod_{i=1}^N \sqrt{\frac{2\pi}{\epsilon m^2}}\right) \exp\left(-m\sum_{i=1}^N \epsilon\sqrt{\dot{X}_i^2}\right) \tag{3.2.27}$$

The infinite product of prefactors evaluates to an overall, field-independent normalization constant $\mathcal{N}_N$. Taking the continuum limit as $N \to \infty$ and $\epsilon \to d\tau$, the discrete sum over slices perfectly restores the continuous Riemann integral over the worldline:

$$\lim_{N\to\infty} Z_{e,N}[X] = \mathcal{N} \exp\left(-m\int_0^1 d\tau\sqrt{\dot{X}^2}\right) \tag{3.2.28}$$

By strictly preserving the locality of the functional measure and performing the integration slice-by-slice, the auxiliary path integral analytically collapses into the square-root geometric action $S_{\text{geo}} = m\int d\tau\sqrt{\dot{X}^2}$.

The path integral over η, X^μ is a product of ordinary integrals, so we can deform the contour just like an ordinary integral. If we take

$$\eta(\tau) \to e^{-i\theta}\eta(\tau)\,, \quad X^0(\tau) \to e^{-i\theta}X^0(\tau)\,, \tag{3.2.29}$$

for an infinitesimal θ, all terms in the exponent are required to have a negative real part for convergence. Now we can rotate the contour in the field space $\eta = -ie, X^0 = -iX^D$, and the integral becomes

$$\int [\mathrm{d}e\,\mathrm{d}X] \exp\left[-\frac{1}{2}\int \mathrm{d}\tau \left(e^{-1}\sum_{\mu=1}^D \dot{X}^\mu \dot{X}_\mu + em^2\right)\right]. \tag{3.2.30}$$

This is the Euclidean point particle analog of the path integral (3.2.1). We have merely performed a contour rotation, so the Euclidean path integral gives the same amplitude as the Minkowski one.

If we write the metric in the tetrad form $\gamma_{ab} = -e_a^0 e_b^0 + e_a^1 e_b^1$ and perform the same rotation on e_a^0, the same treatment applies to the Polyakov action. This provides a formal proof of the equivalence of the Minkowski and Euclidean path integrals. Explicit calculations show they yield the same amplitude in the light-cone and conformal gauges. In (3.2.3a), we left the rotation of X^0 to emphasize the Minkowski character of spacetime. Written in terms of X^0, the equations of motion and OPE are covariant with metric $\eta^{\mu\nu}$.

We note that in two dimensions, χ is locally a total derivative, thus it depends on the topology of the worldsheet—it is the Euler characteristic of the worldsheet. Hence, the $\mathrm{e}^{-\lambda\chi}$ factor in the path integral only affects the relative weights of different topologies in the sum over worldsheets. If we add an extra strip to the worldsheet, as we did from Figure 3.1a to Figure 3.2a, the Euler characteristic decreases by 1, and the path integral is assigned an extra weight e^{λ}. Since adding a strip corresponds to emitting and reabsorbing a virtual open string, the amplitude for emitting an open string from any process is proportional to $\mathrm{e}^{\lambda/2}$. Adding a handle to any worldsheet reduces the Euler characteristic by 2, thus increasing the factor by $\mathrm{e}^{2\lambda}$. Since this corresponds to emitting and reabsorbing a closed string, the amplitude for emitting a closed string is proportional to e^{λ}. The Euler term in the action thus controls the coupling constants in string theory, where

$$g_{\mathrm{o}}^2 \sim g_{\mathrm{c}} \sim \mathrm{e}^{\lambda} \,. \tag{3.2.31}$$

Incidentally, λ can be viewed as a free parameter in the theory, which contradicts the statement in the introduction that string theory has no such parameters. This is a key point, and we will address it in Section 3.7.

On the other hand, the counting of couplings extends in a simple way to worldsheets with string sources. Sources for closed strings are closed loops, and sources for open strings are line segments with boundaries. For convenience, we will require the open string source boundaries to be at right angles with the endpoint boundaries in the worldsheet metric. We must be careful because the boundary curvature diverges at corners: the Euler characteristic is

$$\chi = \tilde{\chi} + \frac{1}{4} n_{\mathrm{c}} \,, \tag{3.2.32}$$

where $\tilde{\chi}$ only includes integrals over smooth boundary parts. n_{c} is the number of corners. The correct weight for the path integral is

$$\exp(-\lambda\tilde{\chi}) = \exp(-\lambda\chi + \lambda n_{\mathrm{c}}/4) \,. \tag{3.2.33}$$

This stems from unitarity: if we cut the worldsheet, the λ dependence must be invariant. This is indeed the case for weight (3.2.33), because the area integral in $\tilde{\chi}$ cancels at the edges. Adding a closed string source is equivalent to cutting a hole in the worldsheet and reducing χ by 1. Adding an open string source leaves χ unchanged but increases n_c by 2. Thus, we obtain the same amplitudes for emitting closed or open strings as in (3.2.31).

3.3 Gauge fixing

The path integral (3.2.1) is not entirely correct, as it contains a massive overcounting. This is because configurations (X, g) and (X', g') related by diff $\times$ Weyl symmetry represent the same physical configuration. In fact, we need to divide by the volume of the local symmetry group,

$$\int \frac{[\mathrm{d}X\,\mathrm{d}g]}{V_{\text{diff} \times \text{Weyl}}} \exp(-S) \equiv Z \,. \tag{3.3.1}$$

We will achieve this through gauge fixing, integrating over a slice that passes through each gauge equivalence class once, and obtaining the correct measure on the slice using the Faddeev-Popov method.

In Chapter 1, we applied the light-cone gauge. While useful for some purposes, this hides certain symmetries of the theory, so we now make a different choice. Note that the metric is symmetric with three independent components, and there are three gauge functions: two coordinates and the local scaling of the metric. Thus, there are exactly enough gauge degrees of freedom to eliminate the integral over the metric, fixing it to a specific functional form, which we call the fiducial metric

$$g_{ab}(\sigma) \to \hat{g}_{ab}(\sigma)\,. \tag{3.3.2}$$

A simple choice is the flat metric or *unit gauge* metric

$$\hat{g}_{ab}(\sigma) = \delta_{ab}\,. \tag{3.3.3}$$

Sometimes it is desirable to consider the effects of the diff group separately. We can turn any metric into a Weyl transformation of the unit form. This is the *conformal gauge*

$$\hat{g}_{ab}(\sigma) = \exp[2\omega(\sigma)]\delta_{ab}\,. \tag{3.3.4}$$

Thus, any metric can be turned into a flat form, at least locally. First, perform a Weyl transformation to make the Ricci scalar zero. The Weyl transformation of the Ricci scalar is

$$g'^{1/2}R' = g^{1/2}\left(R - 2\nabla^2\omega\right)\,. \tag{3.3.5}$$

Setting $R' = 0$ requires solving $2\nabla^2\omega = R$. This is always possible, similar to discussions in electrostatics. In two dimensions, a zero Ricci scalar implies a zero Riemann tensor, because the symmetries of the Riemann tensor imply

$$R_{abcd} = \frac{1}{2}(g_{ac}g_{bd} - g_{ad}g_{bc})R\,. \tag{3.3.6}$$

So the metric is flat and the coordinates are equivalent to the unit metric (3.2.3). Introducing complex coordinates $z = \sigma^1 + i\sigma^2$ here, as in Chapter 2, the flat metric is $\mathrm{d}s^2 = \mathrm{d}z\mathrm{d}\bar{z}$. Consider a coordinate transformation such that z' is a holomorphic function of z

$$z' \equiv \sigma'^1 + i\sigma'^2 = f(z)\,, \tag{3.3.7}$$

combined with a Weyl transformation. The new metric is

$$\mathrm{d}s'^2 = \exp(2\omega)|\partial_z f|^{-2}\mathrm{d}z'\mathrm{d}\bar{z}'\,. \tag{3.3.8}$$

Then for

$$\omega = \ln|\partial_z f|\,, \tag{3.3.9}$$

this metric is invariant.

Most of this additional freedom will be eliminated when we examine the entire worldsheet. There are two issues here. First, in the discussion of Noether's theorem and Ward identities, we emphasized under (2.3.10) that these results depend on symmetries locally defined on the worldsheet. Transformations (3.3.7) and (3.3.9) will thus give conserved currents and Ward identities. This is the conformal invariance studied in Chapter 2. We see that this originates from the subgroup of diff $\times$ Weyl transformations that preserve the unit metric.

The second issue is what happens to our counting of metrics and gauge degrees of freedom when we do examine the entire worldsheet. we will deal with these in Chapter 5.

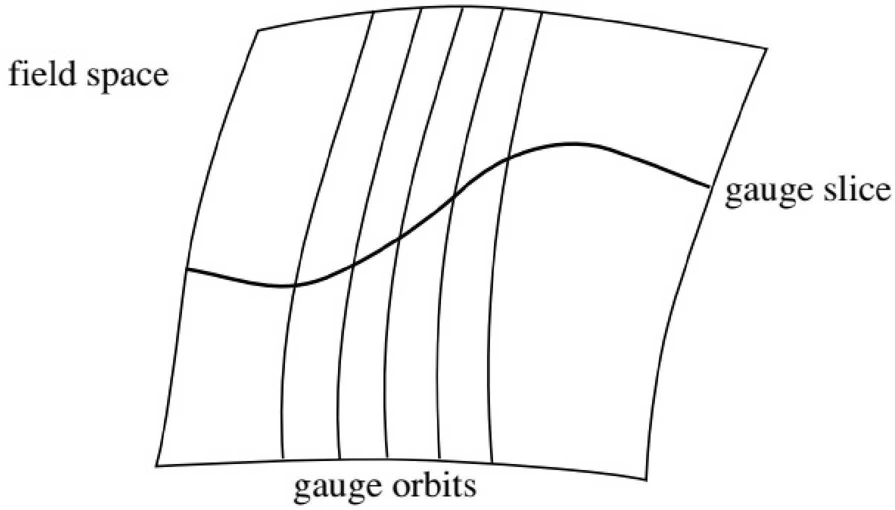


Figure 3.7: Graphical representation of the Faddeev–Popov procedure. A gauge orbit is a family of gauge-equivalent configurations. Integrating over the entire field space and dividing by the orbit volume is equivalent to integrating on a slice that intersects all orbits exactly once.

Faddeev-Popov determinant

After fixing the metric, the functional integral is taken over the slice parameterized by X^μ. To obtain the correct measure, we use the Faddeev-Popov procedure. This method is the same as that used to obtain the gauge-fixing measure in Yang-Mills theory. Figure 3.7 illustrates this method.

Let ζ be the combination of coordinate transformations and Weyl transformations

$$\zeta : g \to g^\zeta , \quad g^\zeta_{ab}(\sigma') = \exp[2\omega(\sigma)] \frac{\partial\sigma^c}{\partial\sigma'^a}\frac{\partial\sigma^d}{\partial\sigma'^b} g_{cd}(\sigma) \,. \tag{3.3.10}$$

According to standard procedure, define the Faddeev-Popov measure Δ_{FP}

$$1 = \Delta_{\text{FP}}(g) \int [\mathrm{d}\zeta]\, \delta(g - \hat{g}^\zeta) \,, \tag{3.3.11}$$

where $\hat{g}_{ab}$ is the fiducial metric. In (3.3.11), $[\mathrm{d}\zeta]$ is the gauge-invariant measure on the diff $\times$ Weyl group. The δ-function is actually a δ-functional, requiring $g_{ab}(\sigma) = \hat{g}^\zeta_{ab}(\sigma)$ at every point. Substituting (3.3.11) into the functional integral (3.3.1)

$$Z[\hat{g}] = \int \frac{[\mathrm{d}\zeta\, \mathrm{d}X\, \mathrm{d}g]}{V_{\text{diff} \times \text{Weyl}}} \Delta_{\text{FP}}(g)\delta(g - \hat{g}^\zeta) \exp(-S[X, g]) \,. \tag{3.3.12}$$

Integrating over g_{ab} and renaming the integration variable $X \to X^\zeta$ gives

$$Z[\hat{g}] = \int \frac{[\mathrm{d}\zeta\, \mathrm{d}X^\zeta]}{V_{\text{diff} \times \text{Weyl}}} \Delta_{\text{FP}}(\hat{g}^\zeta) \exp\Big(-S[X^\zeta, \hat{g}^\zeta]\Big) \,. \tag{3.3.13}$$

Now, utilizing the gauge invariance of Δ_{FP}[1], the invariance of the functional integral measure

[1] Gauge invariance of Δ_{FP}:

$$\Delta_{\text{FP}}(g^\zeta)^{-1} = \int [\mathrm{d}\zeta']\, \delta(g^\zeta - \hat{g}^{\zeta'}) = \int [\mathrm{d}\zeta']\, \delta(g - \hat{g}^{\zeta^{-1}\cdot\zeta'})$$
$$= \int [\mathrm{d}\zeta'']\, \delta(g - \hat{g}^{\zeta''}) = \Delta_{\text{FP}}(g)^{-1} \,,$$

where $\zeta'' = \zeta^{-1} \cdot \zeta'$.

[dX], and the gauge invariance of the action, we get:

$$Z[\hat{g}] = \int \frac{[\mathrm{d}\zeta\, \mathrm{d}X]}{V_{\text{diff}\times\text{Weyl}}} \Delta_{\text{FP}}(\hat{g}) \exp(-S[X,\hat{g}])\,. \tag{3.3.14}$$

Finally, since nothing in the integrand depends on ζ, the integral over ζ merely yields the volume of the gauge group and cancels with the denominator, leaving

$$Z[\hat{g}] = \int [\mathrm{d}X]\, \Delta_{\text{FP}}(\hat{g}) \exp(-S[X,\hat{g}])\,. \tag{3.3.15}$$

Thus, $\Delta_{\text{FP}}(\hat{g})$ is the correct measure on the slice.

To calculate the expression for the Faddeev-Popov measure (3.3.11), we assume that the choice of gauge perfectly fixes the gauge symmetry, so that for exactly one value of ζ, $\delta(g - \hat{g}^\zeta)$ is non-zero; for $\delta(g - \hat{g}^\zeta)$, this is an equality. As mentioned earlier, there is a small global mismatch, which we will address later when time is involved—it does not affect the local properties of the worldsheet. For ζ near 1, we can expand

$$\begin{aligned}\delta g_{ab} &= 2\delta\omega\, g_{ab} - \nabla_a \delta\sigma_b - \nabla_b \delta\sigma_a \\ &= (2\delta\omega - \nabla_c \delta\sigma^c) g_{ab} - 2(P_1 \delta\sigma)_{ab}\,.\end{aligned} \tag{3.3.16}$$

where P_1 is the differential operator that takes a vector into a traceless symmetric 2-tensor

$$(P_1\delta\sigma)_{ab} = \frac{1}{2}(\nabla_a\delta\sigma_b + \nabla_b\delta\sigma_a - g_{ab}\nabla_c\delta\sigma^c)\,. \tag{3.3.17}$$

Proof. We need to work out

$$\delta g_{ab}(\sigma) = g'_{ab}(\sigma) - g_{ab}(\sigma)$$

for an infinitesimal version of (3.3.10), i.e. with $e^{2\omega} = 1 + 2\omega$ and $\sigma'_a = \sigma_a + \delta\sigma_a$ First, we see that

$$\frac{\partial \sigma^a}{\partial \sigma'^b} = \delta^a_b - \partial_b \delta^a$$

Thus, to first order,

$$\begin{aligned}\delta g_{ab}(\sigma) &= g'_{ab}(\sigma) - g_{ab}(\sigma) = g'_{ab}(\sigma' - \delta\sigma) - g_{ab}(\sigma) \\ &= g'_{ab}(\sigma') - \delta\sigma^c \partial_c g'_{ab}(\sigma') - g_{ab}(\sigma) \\ &= (1+2\omega)(\delta^c_a - \partial_a\delta^c)(\delta^d_b - \partial_b\delta^d) g_{cd}(\sigma) - \delta^c\partial_c g_{ab}(\sigma') - g_{ab}(\sigma) \\ &= \delta^c_a\delta^d_b g_{cd} - \delta^d_b \partial_a\delta^c g_{cd} - \delta^c_a\partial_b\delta^d g_{cd} + 2\omega\delta^c_a\delta^d_b g_{cd} - \delta^c\partial_c g_{ab} - g_{ab} \\ &= 2\omega g_{ab} - \partial_a\delta^c g_{bc} - \partial_b\delta^d g_{ad} - \delta^c\partial_c g_{ab}\end{aligned}$$

We now use $\delta\sigma^a = g^{ab}\delta\sigma_b$:

$$\begin{aligned}\delta g_{ab}(\sigma) &= 2\omega g_{ab} - \partial_a(g^{cd}\delta\sigma_d) g_{bc} - \partial_b(g^{cd}\delta\sigma_c) g_{ad} - \delta^c\partial_c g_{ab} \\ &= 2\omega g_{ab} - \partial_a g^{cd}\delta\sigma_d g_{bc} - \partial_a\delta\sigma_d g^{cd} g_{bc} - \partial_b g^{cd}\delta\sigma_c g_{ad} - g^{cd}\partial_b\delta\sigma_c g_{ad} - \delta^c\partial_c g_{ab} \\ &= 2\omega g_{ab} + \delta\sigma_d g^{cd}\partial_a g_{bc} - \partial_a\delta\sigma_d g^d_b + \delta\sigma_c g^{cd}\partial_b g_{ad} - \partial_b\delta\sigma_c\delta^c_a - \delta^c\partial_c g_{ab} \\ &= 2\omega g_{ab} - \partial_a\delta\sigma_b - \partial_b\delta\sigma_a + \delta\sigma_d g^{cd}\partial_a g_{bc} + \delta\sigma_c g^{cd}\partial_b g_{ad} - g^{cd}\delta\sigma_d\partial_c g_{ab} \\ &= 2\omega g_{ab} - \partial_a\delta\sigma_b - \partial_b\delta\sigma_a + \delta\sigma_d g^{cd}(\partial_a g_{bc} + \partial_b g_{ac} - \partial_c g_{ab})\end{aligned}$$

Let us now work out $\nabla_a\delta\sigma_b + \nabla_b\delta\sigma_a$ keeping in mind that because the indices are downstairs, the connection gets a minus sign

$$
\begin{aligned}
\nabla_a \delta\sigma_b + \nabla_b \delta\sigma_a &= \partial_a \delta\sigma_b - \Gamma^c_{ab}\delta\sigma_c + \partial_b \delta\sigma_a - \Gamma^c_{ba}\delta\sigma_c \\
&= \partial_a \delta\sigma_b + \partial_b \delta\sigma_a - (\Gamma^c_{ab} + \Gamma^c_{ba})\delta\sigma_c \\
&= \partial_a \delta\sigma_b + \partial_b \delta\sigma_a - 2\Gamma^d_{ab}\delta\sigma_d \\
&= \partial_a \delta\sigma_b + \partial_b \delta\sigma_a - \delta\sigma_d g^{cd}(\partial_a g_{bc} + \partial_b g_{ac} - \partial_c g_{ab})
\end{aligned}
$$

and we see that indeed

$$\delta g_{ab}(\sigma) = 2\omega g_{ab} - \nabla_a \delta\sigma_b - \nabla_b \delta\sigma_a$$

If we now just fill in the definition of P_1 in (3.3.16) we get

$$
\begin{aligned}
\delta g_{ab} &= 2\delta\omega g_{ab} - \nabla_c \delta\sigma^c g_{ab} - (\nabla_a \delta\sigma_b + \nabla_b \delta\sigma_a - g_{ab}\nabla_c \delta\sigma^c) \\
&= 2\delta\omega g_{ab} - \nabla_a \delta\sigma_b - \nabla_b \delta\sigma_a
\end{aligned}
$$

Which shows that the first and second line of (3.3.16) are equal. Next, we show that P_1 takes vectors into traceless symmetric tensors. First, it is obvious from the definition (3.3.17) that

$$(P_1 \delta\sigma)_{ab} = (P_1 \delta\sigma)_{ba}$$

Next, also the tracelessness is obvious

$$
\begin{aligned}
g^{ab}(P_1\delta\sigma)_{ab} &= g^{ab}\frac{1}{2}(\nabla_a \delta\sigma_b + \nabla_b \delta\sigma_a - g_{ab}\nabla_c \delta\sigma^c) \\
&= \frac{1}{2}(2\nabla_a \delta\sigma^a - \delta^a_a \nabla_c \delta\sigma^c) = 0
\end{aligned}
$$

But notice that the tracelessness is only valid in two dimensions. □

Near 1, the inverse determinant (3.3.11) becomes

$$
\begin{aligned}
\Delta_{\mathrm{FP}}(\hat{g})^{-1} &= \int [\mathrm{d}\delta\omega\, \mathrm{d}\delta\sigma]\, \delta\Big[-(2\delta\omega - \hat{\nabla}\cdot\delta\sigma)\hat{g} + 2\hat{P}_1\delta\sigma\Big] \\
&= \int [\mathrm{d}\delta\omega\, \mathrm{d}\beta\, \mathrm{d}\delta\sigma]\, \exp\Big\{2\pi i \int \mathrm{d}^2\sigma\, \hat{g}^{1/2}\beta^{ab}\Big[-(2\delta\omega - \hat{\nabla}\cdot\delta\sigma)\hat{g} + 2\hat{P}_1\delta\sigma\Big]_{ab}\Big\} \\
&= \int [\mathrm{d}\beta'\, \mathrm{d}\delta\sigma]\, \exp\Big\{4\pi i \int \mathrm{d}^2\sigma\, \hat{g}^{1/2}\beta'^{ab}(\hat{P}_1\delta\sigma)_{ab}\Big\}\,. \qquad (3.3.18)
\end{aligned}
$$

The hat on the differential operator indicates that it contains the fiducial metric $\hat{g}$. In the second equality, we used the integral representation of the functional δ-function, introducing the symmetric tensor field β^{ab}. In the third equality, we integrated out $\delta\omega$, which produced a δ-functional forcing β^{ab} to be traceless; then the functional integral $[\mathrm{d}\beta']$ is over *traceless* symmetric tensors. We now have an expression for $\Delta_{\mathrm{FP}}(\hat{g})^{-1}$ as a functional integral over the vector field $\delta\sigma^a$ and the traceless symmetric tensor β'^{ab}.

By replacing the bosonic fields with corresponding Grassmann *ghost* fields

$$\delta\sigma^a \to c^a\,, \qquad (3.3.19a)$$

$$\beta'_{ab} \to b_{ab}\,, \qquad (3.3.19b)$$

we can transform the path integral, where b^{ab} is traceless like β'^{ab}. Hence,

$$\Delta_{\mathrm{FP}}(\hat{g}) = \int [\mathrm{d}b\, \mathrm{d}c]\, \exp(-S_{\mathrm{g}})\,, \qquad (3.3.20)$$

where, under a convenient normalization of the fields, the ghost action is

$$S_{\mathrm{g}} = \frac{1}{2\pi}\int \mathrm{d}^2\sigma\, \hat{g}^{1/2} b_{ab}\hat{\nabla}^a c^b = \frac{1}{2\pi}\int \mathrm{d}^2\sigma \hat{g}^{1/2} b_{ab}(\hat{P}_1, c)^{ab}\,. \qquad (3.3.21)$$

Locally on the worldsheet, the Polyakov path integral is

$$Z[\hat{g}] = \int [\mathrm{d}X\,\mathrm{d}b\,\mathrm{d}c]\ \exp(-S_X - S_{\mathrm{g}})\,. \tag{3.3.22}$$

Since this action is quadratic in the fields, we can express the path integral in the form of determinants. This calculation will be described in more detail in Chapter 5; the result is

$$Z[\hat{g}] = (\det \hat{\nabla}^2)^{-D/2} \det \hat{P}_1\,, \tag{3.3.23}$$

where the first determinant comes from the X integration and the second from the ghost integration.

In the conformal gauge, $\hat{g}_{ab}(\sigma) = \exp[2\omega(\sigma)]\delta_{ab}$, the ghost action (3.3.21) is

$$\begin{aligned} S_{\mathrm{g}} &= \frac{1}{2\pi}\int \mathrm{d}^2z\,\left(b_{zz}\nabla_{\bar{z}}c^z + b_{\bar{z}\bar{z}}\nabla_z c^{\bar{z}}\right) \\ &= \frac{1}{2\pi}\int \mathrm{d}^2z\,\left(b_{zz}\partial_{\bar{z}}c^z + b_{\bar{z}\bar{z}}\partial_z c^{\bar{z}}\right)\,. \end{aligned} \tag{3.3.24}$$

Note that $\omega(\sigma)$ does not appear in the final form. This is because if a tensor has only z indices, its covariant $\bar{z}$ derivative simplifies to a simple derivative, and vice-versa, which the reader can verify by solving for the connection tensor. Since the action (3.3.24) is Weyl invariant, b_{ab} and c^a are neutral under local Weyl transformations (while b^{ab} and c_a are not; they contain extra metric factors). Because a conformal transformation is a combination of a coordinate transformation (3.3.7) and a Weyl transformation (3.3.9), we can immediately read off the conformal transformations of b_{ab} and c^a from the tensor indices: these are the bc CFT with $(h_b, h_c) = (2, -1)$ and the $\tilde{b}\tilde{c}$ CFT with $(\tilde{h}_b, \tilde{h}_c) = (2, -1)$.

The method we used to derive the gauge-fixed path integral is sometimes called a "heuristic" derivation. The Polyakov functional integral (3.3.1) is not reasonably defined due to massive gauge overcounting. On the other hand, the gauge-fixed functional (3.3.22) is quite well defined and can be calculated clearly. For practical reasons, we should perhaps regard the gauge-fixed form as the starting point, while the gauge-invariant one gives a direct interpretation. In the gauge-fixed form, the substance of gauge symmetry is expressed as BRST invariance, which we will address in the next chapter.

For open strings, we must also consider parts containing boundary segments. It is convenient to view the functional integral as an integral of fields over a fixed region, so diff invariance is restricted to coordinate transformations that map boundaries to themselves. Thus, the variation $\delta\sigma^a$ has a zero normal component

$$n_a\delta\sigma^a = 0\,. \tag{3.3.25}$$

This boundary condition is inherited by the corresponding ghost field

$$n_a c^a = 0\,, \tag{3.3.26}$$

so c^a is proportional to the tangent vector t^a. Then the equations of motion provide a boundary condition for b_{ab}. They have surface terms

$$\int_{\partial M} \mathrm{d}s\, n^a b_{ab}\delta c^b = 0\,. \tag{3.3.27}$$

Combining with the boundary condition on c^a, this implies

$$n_a t_b b^{ab} = 0\,. \tag{3.3.28}$$

These are the boundary conditions used in Section 2.7.

3.4 The Weyl anomaly

A key feature of string theory is that it is self-consistent only in specific spacetime backgrounds. For bosonic string theory in flat spacetime, the condition is $D = 26$. It was first discovered from a pathology in scattering amplitudes. In light-cone analysis, it originated from the loss of Lorentz symmetry. The underlying cause of this constraint is an *anomaly* in local worldsheet symmetry. In our current formalism, the anomaly is in the Weyl symmetry: $T^a{}_a$ is zero in the classical sense, but not in the quantum theory.

A global Weyl transformation, $\omega(\sigma) =$ constant, is equivalent to a general scaling of length. Several field theories in four dimensions are invariant under such a scaling. These include the ϕ^4 massless scalar field theory and non-Abelian gauge field theories. Scale invariance is apparent because the Lagrangian contains only dimensionless parameters—the coupling constant in each case is dimensionless. However, due to divergences in quantum theory, there is a non-zero renormalization group β function in each theory, implying that the effective coupling constant is actually a function of the length scale. Correspondingly, the trace of the energy-momentum tensor is zero in the classical theory, but after quantum effects are considered, the trace is non-zero and proportional to the β function.

In the previous section, we ignored the possibility of anomalies. We need to verify whether the gauge-fixed path integral $Z[g]$ is truly independent of the choice of fiducial metric,

$$Z[g^\zeta] = Z[g]\ ? \tag{3.4.1}$$

In Section 4.2, we will generalize this to more general gauge transformations. For convenience, hereafter g without a hat denotes the fiducial metric. We are also interested in path integrals with extra insertions

$$\langle \cdots \rangle_g \equiv \int [\mathrm{d}X\, \mathrm{d}b\, \mathrm{d}c]\ \exp(-S[X, b, c, g]) \cdots\ . \tag{3.4.2}$$

Then, we require

$$\langle \cdots \rangle_{g^\zeta} = \langle \cdots \rangle_g\ . \tag{3.4.3}$$

Currently, we are not interested in the details of the insertions—that will be the subject of the next two sections—we restrict our attention to gauge transformations where ζ is zero near the field positions in "$\ldots$". That is, we want to know if Weyl invariance holds as an operator equation.

It is easy to guarantee diff invariance and Poincaré invariance in quantum theory. For example, the gauge-fixed path integral can be defined using Pauli-Villars regularization, dividing the path integral by a massive regulator field. A massive regulator field can be coupled to the metric in a diff-invariant and Poincaré-invariant way. However, for a regulator field Y^μ, the diff-invariant and Poincaré-invariant term $\mu^2 \int \mathrm{d}^2\sigma\ g^{1/2} Y^\mu Y_\mu$ is not Weyl-invariant. Any Weyl transformation can be obtained by repeated infinitesimal transformations, so studying infinitesimal transformations is sufficient. The energy-momentum tensor $T^{\mu\nu}$ is the infinitesimal variation of the path integral with respect to the metric,

$$\delta \langle \cdots \rangle_g = -\frac{1}{4\pi} \int \mathrm{d}^2\sigma\ g(\sigma)^{1/2} \delta g_{ab}(\sigma) \Big\langle T^{ab}(\sigma) \cdots \Big\rangle_g \tag{3.4.4}$$

Classically, $T^{\mu\nu}$ originates entirely from the variation of the action, and this is consistent with the Noether definition

$$T^{ab}(\sigma) \overset{\text{classical}}{=} \frac{4\pi}{g(\sigma)^{1/2}} \frac{\delta}{\delta g_{ab}(\sigma)} S\ . \tag{3.4.5}$$

In the flat worldsheet limit, it also reduces to the earlier definition. If we take δg_{ab} to have the form of a coordinate transformation, the diff invariance of the path integral implies that T^{ab} is conserved.

In the following analysis, we will first ignore possible boundary terms. For a Weyl transformation, the definition of T^{ab} in (3.4.4) implies

$$\delta_{\mathrm{W}}\langle\cdots\rangle_g = -\frac{1}{2\pi}\int \mathrm{d}^2\sigma\, g(\sigma)^{1/2}\delta\omega(\sigma)\langle T^a{}_a(\sigma)\cdots\rangle_g\,, \tag{3.4.6}$$

so for general insertions "...", Weyl invariance can be stated as an operator statement: the energy-momentum tensor is traceless

$$T^a{}_a \stackrel{?}{=} 0\,. \tag{3.4.7}$$

The classical action is Weyl invariant, but in the quantum theory, since we haven't found a completely gauge-invariant regulator, this trace might be non-zero. This trace must be diff invariant and Weyl invariant because we protected these symmetries; it is zero on a flat worldsheet because we know from the previous chapter that the theory is conformally flat. This leaves only one possibility:

$$T^a{}_a = a_1 R\,, \tag{3.4.8}$$

where a_1 is some constant and R is the worldsheet Ricci scalar. Terms with two or more derivatives are forbidden for dimensional reasons. When using worldsheet length as the unit, g_{ab} and X^μ are dimensionless, so the constant a_1 in (3.4.8) is also dimensionless. Terms with multiple derivatives would have a coefficient with a positive power of the cutoff length scale used to define the path integral, and thus vanish in the limit as the cutoff is taken to zero.

Calculation of the Weyl anomaly

The possible obstacle to constructing a gauge-invariant theory has reduced to a single constant a_1. In fact, this number is proportional to a quantity we have calculated—the *central charge* c of the CFT on a flat worldsheet. Both a_1 and c are related to the two-point function of the energy-momentum tensor, though through different components. To get the precise relationship between them, we will need to use diff invariance.

In complex coordinates,

$$T_{z\bar z} = \frac{a_1}{2} g_{z\bar z} R\,. \tag{3.4.9}$$

Taking the covariant derivative,

$$\nabla^{\bar z} T_{\bar z z} = \frac{a_1}{2}\nabla^{\bar z}(g_{z\bar z}R) = \frac{a_1}{2}\nabla_z R = \frac{a_1}{2}\partial_z R\,, \tag{3.4.10}$$

where we used the property that the metric is covariant constant. Then, through the conservation of T_{ab},

$$\nabla^z T_{zz} = -\nabla^{\bar z} T_{\bar z z} = -\frac{a_1}{2}\partial_z R\,. \tag{3.4.11}$$

To fix a_1, we compare the Weyl transformations on both sides. The Weyl transformation of the right side, (3.3.5), is

$$a_1 \partial_z \nabla^2 \delta\omega \approx 4 a_1 \partial_z^2 \partial_{\bar z}\delta\omega\,, \tag{3.4.12}$$

where we expand near a flat worldsheet. To get the Weyl transformation on T_{zz}, we first use the conformal transformation (2.4.25),

$$\epsilon^{-1}\delta T_{zz}(z) = -\frac{c}{12}\partial_z^3 v^z(z) - 2\partial_z v^z(z) T_{zz}(z) - v^z(z)\partial_z T_{zz}(z)\,. \tag{3.4.13}$$

From the discussion in Section 3.3, this conformal transformation consists of a coordinate transformation $\delta z = \epsilon v$ and a Weyl transformation $2\delta\omega = \epsilon\partial v + \epsilon(\partial v)^*$. The last two terms in the variation are coordinate transformations of tensors, so to leading order in flat space, the Weyl transformation is

$$\delta_{\mathrm{W}} T_{zz} = -\frac{c}{6}\partial_z^2\delta\omega\,. \tag{3.4.14}$$

Acting with $\partial^z = 2\partial_{\bar{z}}$ and comparing with the transformation (3.4.12) gives:

$$c = -12a_1, \qquad T^a{}_a = -\frac{c}{12}R\,. \tag{3.4.15}$$

Proof. Starting from

$$\sqrt{g'}R' = \sqrt{g}(R - 2\nabla^2\omega) \tag{3.4.16}$$

For an infinitesimal Weyl transformation $g'_{ab}(\sigma') = e^{2\delta\omega}g_{ab}(\sigma) = (1+2\delta\omega)g_{ab}(\sigma)$ we have for the RHS of (3.4.11)

$$\delta_W \nabla^{\bar{z}} T_{\bar{z}z} = -\frac{a_1}{2}\partial_z \delta_W R \tag{3.4.17}$$

now

$$\begin{aligned}\delta_W R = R'(\sigma) - R(\sigma) &= \sqrt{\frac{g}{g'}}(R - 2\nabla^2\omega) - R \\ &= e^{-2\omega}(R - 2\nabla^2\omega) - R = (1-2\delta\omega)(R - 2\nabla^2\delta\omega) - R \\ &= -2\delta\omega R - 2\nabla^2\delta\omega\end{aligned} \tag{3.4.18}$$

Therefore

$$\delta_W \nabla^{\bar{z}} T_{\bar{z}z} = -\frac{a_1}{2}\partial_z(-2\delta\omega R - 2\nabla^2\delta\omega) \tag{3.4.19}$$

We now expand this near a flat worldsheet, where we have $R = 0$ and $\nabla^2 = 2g^{z\bar{z}}\nabla_z\nabla_{\bar{z}} = 4\partial_z\partial_{\bar{z}}$. This gives

$$\delta_W \nabla^{\bar{z}} T_{\bar{z}z} = -\frac{a_1}{2}\partial_z(-2\times 4\partial_z\partial_{\bar{z}}\delta\omega) = 4a_1\partial_z^2\partial_{\bar{z}}\delta\omega \tag{3.4.20}$$

Combining (3.4.12) and (3.4.14) we find

$$4a_1\partial_z^2\partial_{\bar{z}}\delta\omega = \nabla^z\left(-\frac{c}{6}\partial_z^2\delta\omega\right) = -\frac{c}{6}g^{z\bar{z}}\partial_z\partial_{\bar{z}}\delta\omega = -\frac{c}{3}\partial_{\bar{z}}\partial_z^2\delta\omega \tag{3.4.21}$$

and so

$$a_1 = -\frac{c}{12} \tag{3.4.22}$$

□

We re-derive this once more with a slightly longer procedure, in which we will obtain a useful intermediate result. In the conformal gauge,

$$R = -2\exp(-2\omega)\partial_a\partial_a\omega\,, \tag{3.4.23a}$$

$$\nabla^2 = \exp(-2\omega)\partial_a\partial_a\,. \tag{3.4.23b}$$

By contracting two lower indices, i.e., $\delta^{ab}\partial_a\partial_b$. The Weyl variation of $Z[g]$, (3.4.6), becomes

$$\delta_{\rm W} Z[\exp(2\omega)\delta] = \frac{a_1}{\pi}Z[\exp(2\omega)\delta]\int \mathrm{d}^2\sigma\;\delta\omega\partial_a\partial_a\omega\,, \tag{3.4.24}$$

where $\exp(2\omega)\delta$ represents the conformal gauge. Integration immediately yields

$$Z[\exp(2\omega)\delta] = Z[\delta]\exp\left(-\frac{a_1}{2\pi}\int \mathrm{d}^2\sigma\;\partial_a\omega\partial_a\omega\right)\,. \tag{3.4.25}$$

Since every metric is diffeomorphic to a conformal metric, (3.4.25) actually determines the entire dependence of $Z[g]$ on the metric. We need to find a diff-invariant expression that can reduce to (3.4.25) in the conformal gauge. Utilizing (3.4.23), we can prove the expected expression:

$$Z[g] = Z[\delta] \exp\left[\frac{a_1}{8\pi}\int \mathrm{d}^2\sigma \int \mathrm{d}^2\sigma'\, g^{1/2} R(\sigma) G(\sigma,\sigma') g^{1/2} R(\sigma')\right], \tag{3.4.26}$$

where the scalar Green's function is defined by

$$g(\sigma)^{1/2}\nabla^2 G(\sigma,\sigma') = \delta^2(\sigma-\sigma') \tag{3.4.27}$$

This is an interesting result: the path integral is completely determined by the anomaly equation.

Proof. Let us check that (3.4.26) reduces to (3.4.25) in the conformal gauge

$$Z[g]\Big|_{g=e^{2\omega}\delta_+} = Z[\delta_+] \exp\left\{\frac{a_1}{8\pi}\int d^2\sigma \int d^2\sigma' e^{2\omega(\sigma)}\left[-2e^{-2\omega(\sigma)}\partial_a\partial_a\omega(\sigma)\right]\right\} \tag{3.4.28}$$

$$\times G(\sigma,\sigma') e^{2\omega(\sigma')}\left[-2e^{-2\omega(\sigma')}\partial'_a\partial'_a\omega(\sigma')\right] \tag{3.4.29}$$

$$= Z[\delta_+] \exp\left[\frac{a_1}{2\pi}\int d^2\sigma \int d^2\sigma' \partial_a\partial_a\omega(\sigma) G(\sigma,\sigma')\partial'_a\partial'_a\omega(\sigma')\right] \tag{3.4.30}$$

$$= Z[\delta_+] \exp\left[\frac{a_1}{2\pi}\int d^2\sigma \int d^2\sigma' \omega(\sigma)\partial_a\partial_a G(\sigma,\sigma')\partial'_a\partial'_a\omega(\sigma')\right] \tag{3.4.31}$$

where we have used partial integration in the last line. We now use (3.4.23b) and (3.4.27) to rewrite

$$\partial_a\partial_a G(\sigma,\sigma') = e^{2\omega}\nabla^2 G(\sigma,\sigma') = e^{2\omega} g^{-1/2}\delta^2(\sigma-\sigma') \tag{3.4.32}$$

$$= e^{2\omega}e^{-2\omega}\delta^2(\sigma-\sigma') = \delta^2(\sigma-\sigma') \tag{3.4.33}$$

Thus

$$Z[g]\Big|_{g=e^{2\omega}\delta_+} = Z[\delta_+] \exp\left[\frac{a_1}{2\pi}\int d^2\sigma \int d^2\sigma'\, \omega(\sigma)\delta^2(\sigma-\sigma')\partial_a\partial'_a\omega(\sigma')\right] \tag{3.4.34}$$

$$= Z[\delta_+] \exp\left[\frac{a_1}{2\pi}\int d^2\sigma\, \omega(\sigma)\partial_a\partial_a\omega(\sigma)\right] \tag{3.4.35}$$

$$= Z[\delta_+] \exp\left[-\frac{a_1}{2\pi}\int d^2\sigma\, \partial_a\omega(\sigma)\partial_a\omega(\sigma)\right] \tag{3.4.36}$$

where we have, once more, used partial integration in the last line. Note also that (3.4.26) is manifestly diffeomorphism invariant. Indeed R and G are scalar functions and the measure in the integrals is the diffeomorphism invariant measure $d^2\sigma\sqrt{g}$. □

Now expand $Z[g]$ around a flat background, $g_{ab} = \delta_{ab} + h_{ab}$, and retain terms up to second order in $h_{\bar{z}\bar{z}}$. To first order in $h_{\bar{z}\bar{z}}$, the Ricci scalar is $4\partial_z^2 h_{\bar{z}\bar{z}}$, so

$$\begin{aligned}\ln\frac{Z[\delta+h]}{Z[\delta]} &\approx \frac{a_1}{8\pi^2}\int \mathrm{d}^2z \int \mathrm{d}^2z'\, (\partial_z^2 \ln|z-z'|^2) h_{\bar{z}\bar{z}}(z,\bar{z})\partial_{z'}^2 h_{\bar{z}\bar{z}}(z',\bar{z}') \\ &= -\frac{3a_1}{4\pi^2}\int \mathrm{d}^2z \int \mathrm{d}^2z'\, \frac{h_{\bar{z}\bar{z}}(z,\bar{z})h_{\bar{z}\bar{z}}(z',\bar{z}')}{(z-z')^4}\,.\end{aligned} \tag{3.4.37}$$

We can calculate it using second-order perturbation theory in the metric, which is given by (3.4.4):

$$\ln \frac{Z[\delta+h]}{Z[\delta]} \approx \frac{1}{8\pi^2} \int \mathrm{d}^2 z \int \mathrm{d}^2 z' \, h_{\bar{z}\bar{z}}(z,\bar{z}) h_{\bar{z}\bar{z}}(z',\bar{z}') \big\langle T_{zz}(z) T_{zz}(z') \big\rangle_\delta \, . \tag{3.4.38}$$

Now use the standard TT OPE (2.4.26). Except for the first term containing a non-zero spin operator, all terms are zero due to rotational invariance, leaving $\big\langle T_{zz}(z) T_{zz}(z') \big\rangle_1 = \frac{1}{2} c (z-z')^{-4}$. Comparing with the result (3.4.37) from the Weyl anomaly gives (3.4.15) once again.

Proof. We first compute the Ricci scalar in the linear limit. If $g_{ab} = \delta_{ab} + h_{ab}$ then in complex coordinates we need a linear deformation from $g_{zz} = g_{\bar{z}\bar{z}} = 0$ and $g_{z\bar{z}} = 1/2$. Thus

$$g_{-} = \begin{pmatrix} h_{zz} & \frac{1}{2} + h_{z\bar{z}} \\ \frac{1}{2} + h_{z\bar{z}} & h_{\bar{z}\bar{z}} \end{pmatrix} \tag{3.4.39}$$

The inverse metric is

$$g^{-} = \begin{pmatrix} -4h_{z\bar{z}} & 2(1-2h_{z\bar{z}}) \\ 2(1-2h_{z\bar{z}}) & -4h_{zz} \end{pmatrix} \tag{3.4.40}$$

This is easily checked by multiplying the two matrices with one another and showing that they are equal to the identity matrix plus terms of second order in h. To calculate the determinant $\sqrt{g}$ we first revert to the ordinary worldsheet coordinates. We have

$$\begin{aligned} ds^2 &= g_{zz} dz\, dz + g_{\bar{z}\bar{z}} d\bar{z}\, d\bar{z} + 2 g_{z\bar{z}} dz\, d\bar{z} \\ &= h_{zz}(d\sigma^1 + i d\sigma^2)^2 + h_{\bar{z}\bar{z}}(d\sigma^1 - i d\sigma^2)^2 + (1 + 2h_{z\bar{z}}(d\sigma^1 + i d\sigma^2))(d\sigma^1 - i d\sigma^2) \\ &= (1 + h_{zz} + h_{\bar{z}\bar{z}} + 2h_{z\bar{z}}) d\sigma^1 d\sigma^1 + (1 - h_{zz} - h_{\bar{z}\bar{z}} + 2h_{z\bar{z}}) d\sigma^2 d\sigma^2 \\ &\quad + 2i(h_{zz} - h_{\bar{z}\bar{z}}) d\sigma^1 d\sigma^2 \end{aligned} \tag{3.4.41}$$

and therefore

$$g_{11} = 1 + h_{zz} + h_{\bar{z}\bar{z}} + 2h_{z\bar{z}} \tag{3.4.42}$$

$$g_{22} = 1 - h_{zz} - h_{\bar{z}\bar{z}} + 2h_{z\bar{z}} \tag{3.4.43}$$

$$g_{12} = i(h_{zz} - h_{\bar{z}\bar{z}}) \tag{3.4.44}$$

The determinant is therefore

$$\begin{aligned} g &= (1 + h_{zz} + h_{\bar{z}\bar{z}} + 2h_{z\bar{z}})(1 - h_{zz} - h_{\bar{z}\bar{z}} + 2h_{z\bar{z}}) + (h_{zz} - h_{\bar{z}\bar{z}})^2 \\ &= 1 + 4h_{zz} + o(h^2) \end{aligned} \tag{3.4.45}$$

and

$$\sqrt{g} = 1 + 2h_{zz} + o(h^2) \tag{3.4.46}$$

The calculation of R is rather messy, so it was evaluated using mathematica, focussing on the linear terms, the result is quite simple

$$R = 4\partial_{\bar{z}}^2 h_{zz} + 4\partial_z^2 h_{\bar{z}\bar{z}} - 8\partial_z \partial_{\bar{z}} h_{z\bar{z}} + o(h^2) \tag{3.4.47}$$

We now focus on the terms with $h_{\bar{z}\bar{z}}$ only. This means that to first order we can take $\sqrt{g} = 1$ and $R = 4\partial_z^2 h_{\bar{z}\bar{z}}$. Moreover the solution of (3.4.27) is something we already know; it is given by (2.1.27), i.e. $\partial\bar{\partial} \ln |z|^2 = 2\pi \delta^2(z,\bar{z})$. Using $\nabla^2 = 2\partial\bar{\partial} + o(h)$ we get

$$G(\sigma,\sigma') = \frac{1}{4\pi}\ln|z-z'|^2 \tag{3.4.48}$$

We can now use all this in (3.4.26):

$$\begin{aligned}
Z[g] &= Z[\delta]\exp\frac{a_1}{8\pi}\int\frac{1}{2}d^2z\int\frac{1}{2}d^2z' \times 1\times 4\partial_z^2 h_{\bar{z}\bar{z}}(z,\bar{z})\times\frac{1}{4\pi}\ln|z-z'|^2\times 1\times 4\partial_z^2 h_{\bar{z}\bar{z}}(z',\bar{z}') \\
&= Z[\delta]\exp\frac{a_1}{8\pi^2}\int d^2z\int d^2z'\partial_z^2 h_{\bar{z}\bar{z}}(z,\bar{z})\ln|z-z'|^2\partial_z^2 h_{\bar{z}\bar{z}}(z',\bar{z}')
\end{aligned} \tag{3.4.49}$$

Using partial integration and $g=\delta+h$ this gives

$$\ln\frac{Z[\delta+h]}{Z[\delta]} = \frac{a_1}{8\pi^2}\int d^2z\int d^2z' h_{\bar{z}\bar{z}}(z,\bar{z})\partial_z^2\partial_z^2(\ln|z-z'|^2)h_{\bar{z}\bar{z}}(z',\bar{z}') \tag{3.4.50}$$

Now $\partial_z^2\partial_z^2\ln|z-z'|^2 = -6/(z-z')^4$ so that we find (3.4.37).

$$\ln\frac{Z[\delta+h]}{Z[\delta]} = -\frac{3a_1}{4\pi^2}\int d^2z\int d^2z'\frac{h_{\bar{z}\bar{z}}(z,\bar{z})h_{\bar{z}\bar{z}}(z',\bar{z}')}{(z-z')^4} \tag{3.4.51}$$

□

Discussion

For a worldsheet theory consisting of X^μ with central charge D and ghosts with central charge -26, the total central charge is

$$c = c^X + c^g = D - 26\,, \tag{3.4.52}$$

where the ghost central charge is given by (2.5.12) with $\lambda = 2$. This theory is Weyl invariant only for $D = 26$, the same condition for Lorentz invariance found in Chapter 1. When the Weyl anomaly is non-zero, different gauge choices are not equivalent, and just as in gauge theories with anomalies, any one choice leads to an inconsistency, such as loss of covariance or unitarity. For example, in the light-cone gauge, choosing a different spacetime background would require a specific gauge range, thus transforming the potential Weyl anomaly into a Lorentz anomaly.

There is another thing to try: ignore the Weyl anomaly and only treat diff invariance as gauge symmetry. Then the metric will have a real degree of freedom $\omega(\sigma)$ that must be integrated out instead of being gauge-fixed. Since the quantum effects in (3.4.25) act on ω like an action for an X^μ; when $D < 26$, this sign is spacelike. However, the physics is a bit strange, because the diff transformations of ω and the subsequent energy-momentum tensor are different from those of X^μ. In fact, this theory is a linear dilaton CFT. Thus, there is no $D+1$ dimensional Lorentz invariance, and this theory is useless for our current purposes, namely understanding physics as part of our universe. It is an interesting model, and we will return to it briefly in Section 9.9.

Above, we found that a flat worldsheet CFT can couple to a curved metric in a Weyl-invariant way if and only if the total central charge c is zero. However, there is a contradiction, since the same discussion applies to the anti-holomorphic central charge $\tilde{c}$, and we have seen from the example of bc CFT that c does not need to equal $\tilde{c}$. The key point is that in (3.4.11) and (3.4.26), we implicitly assumed worldsheet diff invariance. If $c \neq \tilde{c}$, then no diff-invariant expression exists that is consistent with both $O(h_{zz}^2)$ and $O(h_{\bar{z}\bar{z}}^2)$ calculations, so the CFT *cannot* couple to a curved metric in a diff-invariant way; there is a *diff* anomaly or *gravitational* anomaly. The contradiction indicates that $c = \tilde{c}$ is necessary for a CFT to be free of gravitational anomalies, and it can be proven to be sufficient.

Introducing boundaries

First, return to equation (3.4.8) and consider a more general possibility

$$T^a{}_a = a_1 R + a_2\,. \tag{3.4.53}$$

This is allowed by diff invariance but is non-zero in the flat limit. Earlier we set $a_2 = 0$ through the choice of T_{ab} in (2.3.15b). Equivalently, having more than one conserved T_{ab} on a flat worldsheet means there is freedom in how to couple to a curved metric. Specifically, we can add the action

$$S_{\text{ct}} = b \int \mathrm{d}^2\sigma\, g^{1/2}\,, \tag{3.4.54}$$

which is a two-dimensional cosmological constant and is not Weyl invariant. It adds $2\pi b g_{ab}$ to T_{ab}, so the trace (3.4.53) becomes

$$T^a{}_a = a_1 R + (a_2 + 4\pi b) \tag{3.4.55}$$

By setting $b = -a_2/4\pi$, we can make the second term zero; this is exactly what we did in (2.3.15b). The values of a_2 and b depend on definitions. No local counterterm can eliminate a_1. For example, $\int \mathrm{d}^2\sigma\, g^{1/2} R$ is Weyl invariant. So if a_1 is non-zero, there truly exists an anomaly in Weyl symmetry. We now examine the effects of boundaries. Including boundary terms, the most general possible variation is

$$\begin{aligned}\delta_{\mathrm{W}} \ln Z[g] = &-\frac{1}{2\pi}\int_M \mathrm{d}^2\sigma\, g^{1/2}(a_1 R + a_2)\delta\omega \\ &\qquad -\frac{1}{2\pi}\int_{\partial M} \mathrm{d}s\,(a_3 + a_4 k + a_5 n^a\partial_a)\delta\omega\,,\end{aligned} \tag{3.4.56}$$

where $k = -t^a n_b \nabla_a t^b$ is the geodesic curvature of the boundary. Possible counterterms are

$$S_{\text{ct}} = \int_M \mathrm{d}^2\sigma\, g^{1/2} b_1 + \int_{\partial M} \mathrm{d}s\,(b_2 + b_3 k)\,. \tag{3.4.57}$$

The b_3 term is added to the geodesic curvature term in the Euler characteristic. The Weyl variation of S_{ct}, including the $\mathrm{d}s$ dependence and the dependence of the unit vectors t and n on the metric, is

$$\delta_{\mathrm{W}} S_{\text{ct}} = 2\int_M \mathrm{d}^2\sigma\, g^{1/2} b_1 \delta\omega + \int_{\partial M} \mathrm{d}s\,(b_2 + b_3 n^a\partial_a)\delta\omega\,. \tag{3.4.58}$$

We can choose b_1, b_2 and b_3 such that a_2, a_3 and a_5 are zero, leaving a_1 and a_4 as potential anomalies. In our treatment, there is another tool, the Wess–Zumino consistency condition. Taking the second-order functional variation

$$\begin{aligned}&\delta_{\mathrm{W}_1}(\delta_{\mathrm{W}_2} \ln\langle\cdots\rangle_g) \\ &\quad = \frac{a_1}{\pi}\int_M \mathrm{d}^2\sigma\, g^{1/2}\,\delta\omega_2 \nabla^2 \delta\omega_1 - \frac{a_4}{2\pi}\int_{\partial M} \mathrm{d}s\,\delta\omega_2\, n^a\partial_a\delta\omega_1 \\ &\quad = -\frac{a_1}{\pi}\int_M \mathrm{d}^2\sigma\, g^{1/2}\,\partial_a\delta\omega_2 \partial^a\delta\omega_1 + \frac{2a_1 - a_4}{2\pi}\int_{\partial M} \mathrm{d}s\, n^a\delta\omega_2\,\partial_a\delta\omega_1\,.\end{aligned} \tag{3.4.59}$$

This is a second-order functional derivative, so by definition, it is symmetric with respect to $\delta\omega_1$ and $\delta\omega_2$. The second term is not, so $a_4 = 2a_1$. From this, it follows that in open string boundary conditions, no new Weyl anomaly is generated. In quantum theory, Weyl invariance holds if and only if $a_1 = -c/12$ is zero. This is determined by the physics *internal* to the worldsheet.

The Wess–Zumino consistency condition is a powerful constraint on the form of anomalies. A second application allows us to prove that the central charge is a constant. This follows from Lorentz invariance and dimensional analysis, but in a more general CFT, we could have

$$T^a{}_a(\sigma) = -\frac{\mathscr{C}(\sigma)}{12} R(\sigma)\,, \tag{3.4.60}$$

where $\mathscr{C}(\sigma)$ is some local operator. Specifically, this trace would still be zero on a flat worldsheet. Repeating the calculation (3.4.59), replacing a_1 with $-\mathscr{C}(\sigma)/12$, we can see through integration by parts that there is an extra term proportional to $\partial_a \mathscr{C}(\sigma)$. The consistency condition thus implies that the expectation value containing $\mathscr{C}(\sigma)$ is independent of position, indicating that $\mathscr{C}(\sigma)$ itself is just a numerical constant c.

3.5 Scattering amplitudes

The idea of summing over all worldsheets bounded by initial and final curves seems very natural. However, defining it in a way compatible with local worldsheet symmetry is difficult, and the corresponding amplitudes are quite complex. There is a special case where the amplitudes simplify—namely taking the limit as the string sources tend to infinity. This corresponds to the scattering amplitude with specified incoming and outgoing string states, the S-matrix elements. For much of this book, we will restrict ourselves to S-matrix elements and similar "on-shell" problems. This is not entirely satisfactory, because we want to discuss systems in finite time, not just asymptotic processes. For now, since the S-matrix is easy to understand, we focus on it. Although the correct way to think about off-shell string theory is unclear, we will briefly address this question in the next section and in Chapter 9.

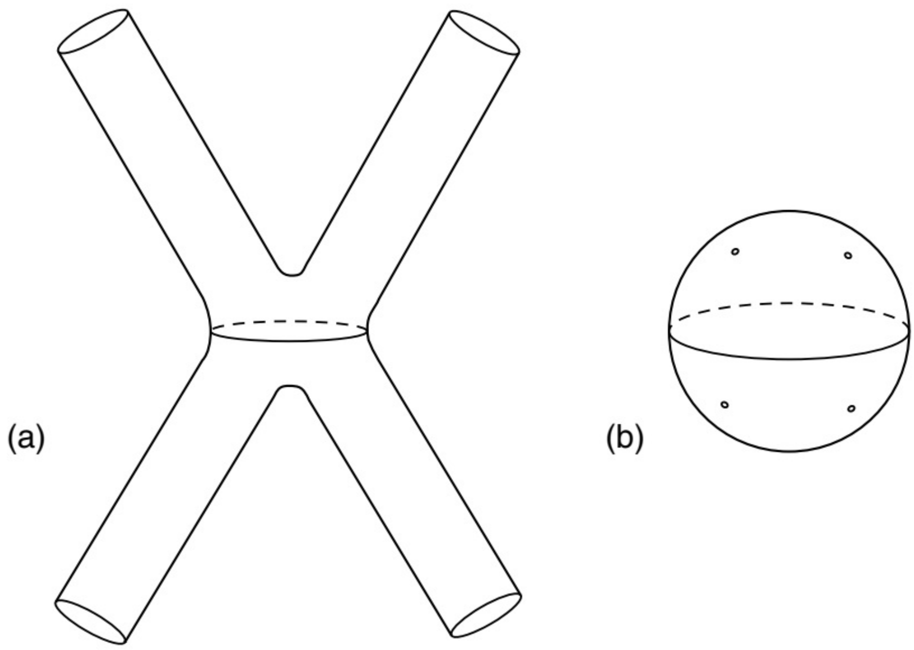


Figure 3.8: (a) Scattering of closed strings, where the string sources tend toward $X^0 = \pm\infty$. (b) Conformally equivalent image of 4-closed-string scattering: a sphere with very small holes.

So, we now examine the process shown in Figure 3.8, where the sources are pulled to infinity. We will discuss how these sources should be represented in the path integral. Away from the scattering process, the strings propagate freely. Each outgoing and incoming string is a long cylinder, which can be described by a complex coordinate w,

$$-2\pi t \leq \operatorname{Im} w \leq 0, \qquad w \cong w + 2\pi\,. \tag{3.5.1}$$

The lower end of the cylinder, $\operatorname{Im} w = -2\pi t$, is the end of the source, the upper end is inserted into the rest of the worldsheet, and the circumference is the periodic $\operatorname{Re} w$ direction. The limit corresponding to the scattering process is $t \to \infty$. It might seem that we are confusing long distances in spacetime with long cylinders in worldsheet coordinates, but this is quite correct: as we will see later, propagation over long spacetime distances originates precisely from such worldsheets where the cylinders become long in the above sense. We know from Chapter 2 that the cylinder has an equivalent description:

$$z = \exp(-iw), \quad \exp(-2\pi t) \leq |z| \leq 1\,. \tag{3.5.2}$$

In this picture, the long cylinder is mapped to a unit disc, where the external string state is the circle at the center. In the limit as $t \to \infty$, the small circle contracts to a point, and the worldsheet reduces to a sphere with a point inserted for each external state.

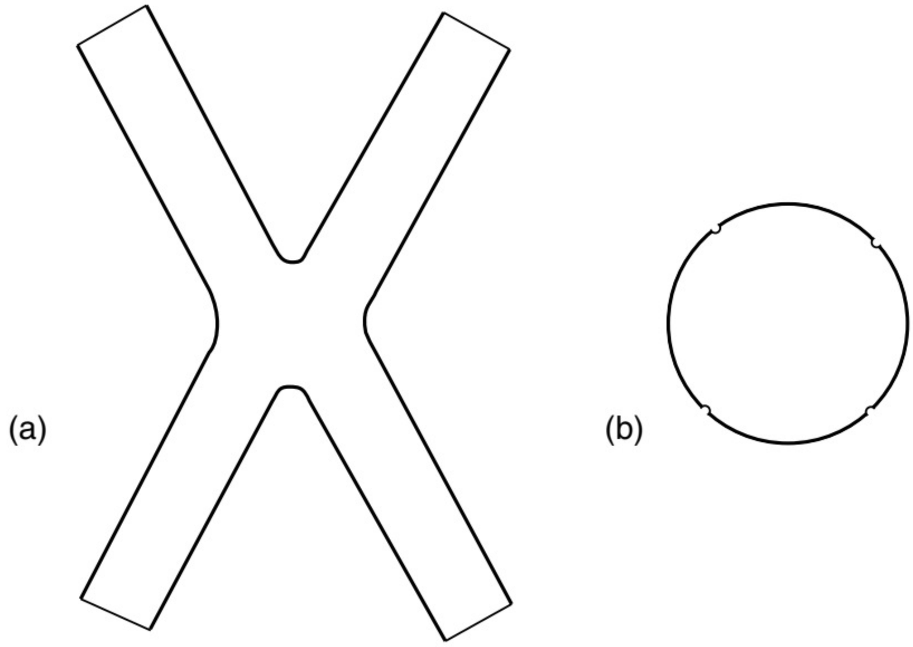


Figure 3.9: (a) Scattering of four open strings, where the string sources tend toward $X^0 = \pm\infty$. (b) Conformally equivalent image: a disc with indentations.

The same concept holds for external open string states. The long strip in Figure 3.9a can be described as

$$-2\pi t \leq \operatorname{Im} w \leq 0, \quad 0 \leq \operatorname{Re} w \leq \pi\,. \tag{3.5.3}$$

where $\operatorname{Im} w = -2\pi t$ is the source, and $\operatorname{Re} w = 0, \pi$ are the endpoint boundaries. Under $z = -\exp(-iw)$, this maps to the intersection of the unit disc and the upper half-plane. A small semicircle is cut at the origin

$$\exp(-2\pi t) \leq |z| \leq 1, \quad \operatorname{Im} z \geq 0\,. \tag{3.5.4}$$

The scattering process now looks like Figure 3.9b. In the limit $t \to \infty$, the source contracts to a point on the (endpoint) boundary.

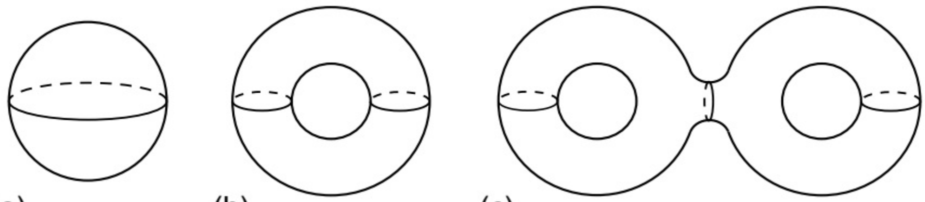
This is the state-operator correspondence we have already seen. Each source becomes a local perturbation on the worldsheet. For a given incoming or outgoing string, which has D-momentum k^μ and internal state j, there exists a corresponding local *vertex operator* $\mathscr{V}_j(k)$ determined by the limiting process. Incoming and outgoing states are distinguished by the sign of k^0; for incoming states $k^\mu = (E, \mathbf{k})$ and outgoing states $k^\mu = -(E, \mathbf{k})$, Figures 3.8 and 3.9 only describe the lowest order amplitudes, but this construction is quite general: we can restrict our attention to compact worldsheets with no tubes tending to infinity, but with a point-like insertion in the interior for each external closed string. For each external open string, there is a point-like insertion on the boundary. Thus, a connected n-point S-matrix element is

$$\begin{aligned} &S_{j_1 \ldots j_n}(k_1, \ldots, k_n) \\ &\quad = \sum_{\substack{\text{compact} \\ \text{topologies}}} \frac{[\mathrm{d}X\, \mathrm{d}g]}{V_{\text{diff} \times \text{Weyl}}} \exp(-S_X - \lambda\chi) \prod_{i=1}^{n} \int \mathrm{d}^2\sigma_i\, g(\sigma_i)^{1/2} \mathscr{V}_{j_i}(k_i, \sigma_i)\,. \end{aligned} \tag{3.5.5}$$

To make the vertex operator insertion diff-invariant, we integrate them over the worldsheet. In the rest of this volume, we will see that our guess is correct and defines a reasonable string S-matrix.

Depending on which of the 4 theories we are examining, the sum over topologies may include unoriented worldsheets and/or worldsheets with boundaries. In general, the sum over topologies is not restricted to connected worldsheets; the overall process may include two or more independent particle scatterings. It is convenient to restrict the sum to connected worldsheets, focusing on the connected S-matrix.

To obtain the connected S-matrix, we must sum over all compact, connected topologies of the worldsheet. In two dimensions, the classification of topologies is well-known. Any compact, connected, oriented surface without boundary is topologically equivalent to a sphere with g handles, where g is called the genus of the surface. In Figure 3.10, we show the $g = 0, 1, 2$ cases.



(a) (b) (c)

Figure 3.10: Compact connected oriented closed surfaces with genus (a) 0, (b) 1, (c) 2 respectively.

Boundaries can be added by cutting holes in the closed surface. Any compact, connected, oriented two-dimensional surface is topologically equivalent to a sphere with g handles and b holes, for example $(g, b) = (0, 1)$ is a disc, $(0, 2)$ is an annulus, and $(0, 3)$ is a pair of pants.

To describe unoriented surfaces, it is useful to introduce cross-caps: cutting a hole in the surface and gluing opposite points together. In detail, taking complex coordinates, cut out a disc slightly smaller than unit radius, and glue together opposite points defined by $z' = -1/\bar{z}$. Since the gluing is anti-holomorphic, the resulting surface is unoriented. In addition, unlike the case of a boundary, a cross-cap does not introduce a boundary or other local features. Since the boundary produced by cutting is only the boundary of the coordinate patch, any compact connected closed surface, whether oriented or unoriented, can always be obtained by adding g handles and c cross-caps to a sphere; any compact connected surface can be obtained by adding g handles, b boundaries, and c cross-caps. In fact, these descriptions are somewhat redundant: for both cases with and without boundaries, they can be obtained precisely by restricting the sum to only one of g or c being non-zero. For example, $(g, b, c) = (0, 0, 1)$ is a projective plane, $(g, b, c) = (0, 1, 1)$ is a Mobius strip, $(g, b, c) = (0, 0, 2)$ is a Klein bottle, and a torus with one cross-cap can be obtained via $(g, b, c) = (0, 0, 3)$ or $(g, b, c) = (1, 0, 1)$, with every two cross-caps being exchangeable for one handle. The Euler characteristic is

$$\chi = 2 - 2g - b - c\,. \tag{3.5.6}$$

The number of different topologies is much smaller than the number of different Feynman diagrams in a given order of field theory. For example, in closed oriented theory, in each order of perturbation theory, there exists exactly one topology. In field theory, the number of diagrams grows geometrically. A single string diagram includes all field theory diagrams. In different limits, handles are much longer than their circumference, we can approximate handles as lines, and string diagrams approximate corresponding Feynman diagrams in this limit.

3.6 Vertex operators

Using the state-operator correspondence, the vertex operator for the closed string tachyon is

$$\begin{aligned} V_0 &= 2g_\mathrm{c} \int \mathrm{d}^2\sigma \, g^{1/2} \, \mathrm{e}^{ik\cdot X} \\ &\to g_\mathrm{c} \int \mathrm{d}^2 z \ :\!\mathrm{e}^{ik\cdot X}\!: \ . \end{aligned} \tag{3.6.1}$$

We introduced a factor g_c in the mapping. This is the closed string coupling constant, originating from adding an extra string to the process: we use the normalization of the vertex operator as the definition of the coupling. In the second line, we transition to a flat worldsheet. The vertex operator must be diff and Weyl invariant, so, in particular, it must be conformally invariant on a flat worldsheet. To cancel the transformation of $\mathrm{d}^2 z$, the operator must be a tensor of weight $(1, 1)$. Through a direct OPE calculation, $\mathrm{e}^{ik\cdot X}$ is a tensor of weight $h = \tilde{h} = \alpha' k^2/4$, so the condition is

$$m^2 = -k^2 = -\frac{4}{\alpha'} \ , \tag{3.6.2}$$

which is exactly the mass discovered in light-cone quantization. Similarly, closed string tensor states at the first excited level have the flat-worldsheet vertex operator

$$\frac{2g_\mathrm{c}}{\alpha'} \int \mathrm{d}^2 z \ :\!\partial X^\mu \bar{\partial} X^\nu \mathrm{e}^{ik\cdot X}\!: \ . \tag{3.6.3}$$

The normalization relative to the tachyon vertex operator originates from the state-operator correspondence, and as we will see in Chapter 6, the same relative normalization is obtained from the normalization of the S-matrix. The weights are

$$h = \tilde{h} = 1 + \frac{\alpha' k^2}{4} \ , \tag{3.6.4}$$

so they are massless, again consistent with light-cone quantization. For it to be a tensor, there are further conditions, which we will handle below.

Vertex operators in the Polyakov formalism

Within the state-operator mapping, we have a systematic method to write down flat-worldsheet vertex operators for any state. Since the worldsheet can always be made flat, in principle, this is all we need. However, it is now useful to study vertex operators on curved worldsheets in the Polyakov framework. The method we use will be somewhat cumbersome and not as systematic as the state-operator mapping, but some results are useful.

Any operator included in the Polyakov path integral (3.5.5) must reflect the local diff $\times$ Weyl symmetry of the theory. We have not yet specified how the operator in the first line of (3.6.1) is defined on the worldsheet. As in our previous discussion of the Weyl anomaly, it is convenient to use a method that automatically preserves diff invariance and then manually check the Weyl transformation. For most purposes, dimensional regularization is used in the literature. However, we do not have a great need for it here, so we will not introduce it; instead, we will generalize the normal ordering introduced earlier.

Define a renormalized operator:

$$[\mathscr{F}]_{\mathrm{r}}=\exp\left(\frac{1}{2}\int d^2\sigma d^2\sigma'\Delta\left(\sigma,\sigma'\right)\frac{\delta}{\delta X^{\mu}(\sigma)}\frac{\delta}{\delta X_{\mu}\left(\sigma'\right)}\right)\mathscr{F} \tag{3.6.5}$$

where

$$\Delta\left(\sigma,\sigma'\right)=\frac{\alpha'}{2}\ln d^2\left(\sigma,\sigma'\right) \tag{3.6.6}$$

and $d\left(\sigma,\sigma'\right)$ is the geodesic distance between σ and σ'.

Similar to normal ordering, (3.6.5) tells us to sum over all ways of contracting pairs in $\mathscr{F}$ using $\Delta\left(\sigma,\sigma'\right)$. On a flat worldsheet, $d^2\left(\sigma,\sigma'\right)=\left|z-z'\right|^2$ and this reduces to the conformal normal ordering we have studied carefully. On a curved worldsheet, it cancels the singular integrals from self-contractions of the fields in $\mathscr{F}$. Diff invariance is manifest, but the contractions depend on the metric. Introduce the Weyl variation:

$$\delta_{\mathrm{W}}[\mathscr{F}]_{\mathrm{r}}=\left[\delta_{\mathrm{W}}\mathscr{F}\right]_{\mathrm{r}}+\frac{1}{2}\int d^2\sigma d^2\sigma'\delta_{\mathrm{W}}\Delta\left(\sigma,\sigma'\right)\frac{\delta}{\delta X^{\mu}(\sigma)}\frac{\delta}{\delta X_{\mu}\left(\sigma'\right)}[\mathscr{F}]_{\mathrm{r}} \tag{3.6.7}$$

The first term represents the explicit Weyl variation in the operator. For the vertex operator of a state with momentum k_μ, under translation, it must transform in the same way as the state, and thus takes the form of $e^{ik\cdot X}$ multiplied by derivatives of X^μ. Operators with different numbers of derivatives do not mix under Weyl transformations.

Proof.

$$\begin{aligned}
\delta_{\mathrm{W}}[\mathcal{F}]_{\mathrm{r}}&=\delta_{\mathrm{W}}e^{\frac{1}{2}\int d^2\sigma d^2\sigma'\Delta(\sigma,\sigma')\frac{\delta}{\delta X^{\mu}(\sigma)}\frac{\delta}{\delta X_{\mu}(\sigma')}}\mathcal{F}\\
&=\delta_{\mathrm{W}}\left[\frac{1}{2}\int d^2\sigma d^2\sigma'\Delta\left(\sigma,\sigma'\right)\frac{\delta}{\delta X^{\mu}(\sigma)}\frac{\delta}{\delta X_{\mu}\left(\sigma'\right)}\right]e^{\frac{1}{2}\int d^2\sigma d^2\sigma'\Delta(\sigma,\sigma')\frac{\delta}{\delta X^{\mu}(\sigma)}\frac{\delta}{\delta X_{\mu}(\sigma')}}\mathcal{F}\\
&\quad+e^{\frac{1}{2}\int d^2\sigma d^2\sigma'\Delta(\sigma,\sigma')\frac{\delta}{\delta X^{\mu}(\sigma)}\frac{\delta}{\delta X_{\mu}(\sigma')}}\delta_{\mathrm{W}}\mathcal{F}\\
&=\frac{1}{2}\int d^2\sigma d^2\sigma'\left(\delta_{\mathrm{W}}\Delta\left(\sigma,\sigma'\right)\right)\frac{\delta}{\delta X^{\mu}(\sigma)}\frac{\delta}{\delta X_{\mu}\left(\sigma'\right)}[\mathcal{F}]_{\mathrm{r}}+\left[\delta_{\mathrm{W}}\mathcal{F}\right]_{\mathrm{r}}
\end{aligned} \tag{3.6.8}$$

□

We begin by examining the zero-derivative case, operator (3.6.1). The Weyl variation originates from the explicit $g^{1/2}$ factor and the renormalization given in (3.6.7):

$$\delta_{\mathrm{W}}V_0=2g_{\mathrm{c}}\int d^2\sigma g^{1/2}\left(2\delta\omega(\sigma)-\frac{k^2}{2}\delta_{\mathrm{W}}\Delta(\sigma,\sigma)\right)\left[e^{ik\cdot X(\sigma)}\right]_{\mathrm{r}} \tag{3.6.9}$$

Proof. with $\delta_W g_{ab}=2\delta\omega g_{ab}$ we find, using (3.6.7)

$$\begin{aligned}
\delta_W V_0&=2g_c\int d^2\sigma\delta_W\sqrt{g}\left[e^{ik\cdot X(\sigma)}\right]_r\\
&=2g_c\int d^2\sigma\left\{\frac{1}{2}g^{-1/2}g^{ab}2\delta\omega g_{ab}\left[e^{ik\cdot X(\sigma)}\right]_r+\sqrt{g}\left[\delta_W e^{ik\cdot X(\sigma)}\right]_r\right.\\
&\quad\left.+\frac{1}{2}\sqrt{g}\int d^2\sigma' d^2\sigma''\delta_W\Delta(\sigma',\sigma'')\frac{\delta}{\delta X^{\mu}(\sigma')}\frac{\delta}{\delta X_{\mu}(\sigma'')}\left[e^{ik\cdot X(\sigma)}\right]_r\right\}
\end{aligned}$$

$$
\begin{aligned}
&= 2g_c \int d^2\sigma \left\{ \sqrt{g} 2\delta\omega \left[e^{ik\cdot X(\sigma)} \right]_r \right. \\
&\qquad \left. + \frac{1}{2}\sqrt{g} \int d^2\sigma' d^2\sigma'' \delta_W \Delta(\sigma', \sigma'')(-k^2)\delta^2(\sigma - \sigma')\delta^2(\sigma - \sigma'') \left[e^{ik\cdot X(\sigma)} \right]_r \right\} \\
&= 2g_c \int d^2\sigma \left\{ \sqrt{g} 2\delta\omega \left[e^{ik\cdot X(\sigma)} \right]_r - \frac{k^2}{2}\sqrt{g}\delta_W \Delta(\sigma, \sigma) \left[e^{ik\cdot X(\sigma)} \right]_r \right\} \\
&= 2g_c \int d^2\sigma \sqrt{g} \left(2\delta\omega - \frac{k^2}{2}\delta_W \Delta(\sigma, \sigma) \right) \left[e^{ik\cdot X(\sigma)} \right]_r
\end{aligned}
\tag{3.6.10}
$$

$\square$

At short distances,

$$d^2\left(\sigma, \sigma'\right) \approx \left(\sigma - \sigma'\right)^2 \exp(2\omega(\sigma)) \tag{3.6.11}$$

consequently,

$$\Delta\left(\sigma, \sigma'\right) \approx \alpha'\omega(\sigma) + \frac{\alpha'}{2}\ln\left(\sigma - \sigma'\right)^2 \tag{3.6.12}$$

In the limit $\sigma' \to \sigma$, the Weyl variation is non-singular:

$$\delta_{\mathrm{W}}\Delta(\sigma, \sigma) = \alpha'\delta\omega(\sigma) \tag{3.6.13}$$

The condition for the Weyl variation (3.6.9) to be zero is $k^2 = 4/\alpha'$, exactly the result derived previously.

Off-shell amplitudes

Consider a naive attempt to define off-shell amplitudes by taking k^μ not on the mass shell. This is inconsistent with local worldsheet symmetry. In position space, this corresponds to a local probe of the string worldsheet, given by an insertion:

$$\delta^D\left(X(\sigma) - x_0\right) = \int \frac{d^D k}{(2\pi)^D} \exp\left[ik\cdot\left(X(\sigma) - x_0\right)\right] \tag{3.6.14}$$

Since this contains all momenta, it is inconsistent. From several perspectives, this is not surprising. First, the previous discussion—that our point sources (vertex operators) for string states require a limiting process—remarkably restricts us to on-shell problems. Second, string theory contains gravity, and simple off-shell amplitudes do not exist in general relativity because we cannot specify the position of the probe without non-physical coordinates; coordinate-invariant off-shell quantities are much more complex. Third, in string theory, we have not introduced additional fields to measure local observables (similar to the electroweak processes used to probe strong interactions): we must use the string itself or other objects inherent in the theory (D-branes and solitons).

At least in perturbation theory, off-shell amplitudes can be defined. If we fix the gauge (for which light-cone gauge is simplest) and use string sources, then the S-matrix can be derived using a reduction formula (3.5.5) similar to quantum field theory. Finite-time transition amplitudes can be defined before taking the infinite-time limit. Incidentally, readers attempting to link the current discussion to field theory should note that the path integral expression (3.5.5) is not like a Green's function but rather has S-matrix element properties. The analogous object in field theory is a Green's function where external propagators are truncated and external momenta are restricted to the mass shell.

The discussion of local probes (3.6.14) also points to the problem of contact interactions between strings. The simplest form of such an interaction would be:

$$\int_M d^2\sigma g(\sigma)^{1/2} \int_M d^2\sigma' g\left(\sigma'\right)^{1/2} \delta^D\left(X(\sigma) - X\left(\sigma'\right)\right) \tag{3.6.15}$$

This is non-zero as long as the worldsheet intersects itself. However, the δ-function contains all momenta, as in (3.6.14). Thus, it is not Weyl invariant. The structure of string theory is quite fixed: strings can only couple to other strings in the manner described at the beginning of this chapter.

Massless closed string vertex operators

While somewhat messy, it is interesting to further pursue the Polyakov treatment of vertex operators. There is no way to construct a worldsheet scalar with exactly one derivative. Therefore, the next case is two derivatives; the possible diff-invariant cases are:

$$V_1 = \frac{g_c}{\alpha'} \int d^2\sigma g^{1/2} \Big\{ \left(g^{ab} s_{\mu\nu} + i\epsilon^{ab} a_{\mu\nu} \right) \left[\partial_a X^\mu \partial_b X^v e^{ik\cdot X} \right]_r + \alpha' \phi R \left[e^{ik\cdot X} \right]_r \Big\} \tag{3.6.16}$$

where $s_{\mu\nu}, a_{\mu\nu}$, and ϕ are a symmetric matrix, an antisymmetric matrix, and a constant, respectively.

The antisymmetric tensor ϵ^{ab} is normalized such that $g^{1/2}\epsilon^{12} = 1$, or $g^{1/2}\epsilon^{z\bar{z}} = -i$. The i accompanying the antisymmetric tensor in the vertex operator can be understood as arising from Euclidean continuation, because this term must have an odd number (one) of time derivatives. This result is widely applied in Euclidean actions and vertex operators. To extend the Weyl invariance analysis, we need to solve the higher-order geodesic distance Weyl correlation (3.6.13).

$$\partial_a \delta_W \Delta \left(\sigma, \sigma'\right)\big|_{\sigma'=\sigma} = \frac{1}{2}\alpha' \partial_a \delta\omega(\sigma) \tag{3.6.17a}$$

$$\partial_a \partial_b' \delta_W \Delta \left(\sigma, \sigma'\right)\big|_{\sigma'=\sigma} = \frac{1+\gamma}{2}\alpha' \nabla_a \partial_b \delta\omega(\sigma) \tag{3.6.17b}$$

$$\nabla_a \partial_b \delta_W \Delta \left(\sigma, \sigma'\right)\big|_{\sigma'=\sigma} = -\frac{\gamma}{2}\alpha' \nabla_a \partial_b \delta\omega(\sigma) \tag{3.6.17c}$$

where $\gamma = -2/3$. We leave it as a parameter for later reference; the third equation can be obtained from the gradient of the second or first. For the variation of curvature, we use (3.3.5). For renormalization variation, we use (3.6.7) and (3.6.17).

$$\delta_W V_1 = \tfrac{g_c}{2} \textstyle\int d^2\sigma g^{1/2} \delta\omega \Big\{ \left(g^{ab} S_{\mu v} + i\epsilon^{ab} A_{\mu v} \right) \left[\partial_a X^\mu \partial_b X^v e^{ik\cdot X} \right]_r + \alpha' F R \left[e^{ik\cdot X} \right]_r \Big\} \tag{3.6.18}$$

where

$$S_{\mu v} = -k^2 s_{\mu v} + k_v k^\omega S_{\mu\omega} + k_\mu k^\omega s_{v\omega} - (1+\gamma) k_\mu k_v s^\omega_\omega + 4 k_\mu k_v \phi \tag{3.6.19a}$$

$$A_{\mu v} = -k^2 a_{\mu\nu} + k_v k^\omega a_{\mu\omega} - k_\mu k^\omega a_{v\omega} \tag{3.6.19b}$$

$$F = (\gamma - 1) k^2 \phi + \frac{1}{2}\gamma k^\mu k^v s_{\mu v} - \frac{1}{4}\gamma(1+\gamma) k^2 s^v_v \tag{3.6.19c}$$

In deriving the Weyl variation (3.6.18), we performed integration by parts and used the relation:

$$\left[\nabla^2 X^\mu e^{ik\cdot X} \right]_r = i\frac{\alpha'\gamma}{4} k^\mu R \left[e^{ik\cdot X} \right]_r \tag{3.6.20}$$

The left side would naively be zero by the equations of motion, but here it is multiplied by another operator at the same point.

The general principle in this case is that the operator need not be zero, but is not independent—it can be expanded in terms of other local operators in the theory. Generally, the coefficients depend on the renormalization scheme; (3.6.20) can be obtained by taking the Weyl variation of both sides. Since $S_{\mu\nu}$, $A_{\mu\nu}$, and F multiply independent operators, the condition for Weyl invariance is $S_{\mu\nu} = A_{\mu\nu} = F = 0$. To get an accurate count of independent operators, we must note that not all V_1 of form (3.6.16) are independent.

Instead, under:

$$s_{\mu v} \to s_{\mu v} + \xi_\mu k_v + k_\mu \xi_v \tag{3.6.21a}$$

$$a_{\mu v} \to a_{\mu\nu} + \zeta_\mu k_v - k_\mu \zeta_v \tag{3.6.21b}$$

$$\phi \to \phi + \frac{\gamma}{2} k \cdot \xi \tag{3.6.21c}$$

the change in V_1 integrates to zero using (3.6.20).

Now choose n^μ such that $n \cdot k = 1$ and $n^2 = 0$. By the constraints:

$$n^\mu s_{\mu v} = n^\mu a_{\mu v} = 0 \tag{3.6.22}$$

we obtain a complete set of independent operators. This gives $2D$ equations for the $2D$ parameters ξ_μ, ζ_μ. It can be proven that one equation and one parameter are trivial, but the remaining system (3.6.22) is non-degenerate, defining independent vertex operators.

Using (3.6.22) and solving $S_{\mu\nu} n^\mu n^\nu = 0$, then $S_{\mu\nu} n^\mu = A_{\mu\nu} n^\mu = 0$, and finally $S_{\mu\nu} = A_{\mu\nu} = F = 0$, we find:

$$k^2 = 0 \tag{3.6.23a}$$

$$k^v s_{\mu v} = k^\mu a_{\mu v} = 0 \tag{3.6.23b}$$

$$\phi = \frac{1+\gamma}{4} s^\mu_\mu \tag{3.6.23c}$$

Once again, we find the mass-shell condition (3.6.23a), this time for the first excited state of the closed string. There is also the condition (3.6.23b) that polarization is transverse to momentum, required for massless tensor fields. In the frame:

$$k^\mu = (1, 1, 0, 0, \ldots, 0), \quad n^\mu = \frac{1}{2}(-1, 1, 0, 0, \ldots, 0) \tag{3.6.24}$$

(3.6.22) and (3.6.23b) imply that the 0 and 1 components are zero. (3.6.23c) fixes ϕ, leaving exactly $(D-2)^2$ operators, matching the light-cone result.

The condition $n^\mu s_{\mu\nu} = n^\mu a_{\mu\nu} = 0$ is needed to remove gauge-equivalent null polarizations. However, the introduction of n^μ explicitly breaks Lorentz invariance. The theory depends on the choice of n^μ only by a gauge transformation. For the theory to be Lorentz invariant and have a positive Hilbert space norm, (3.6.21) is critical. These relations characterize local spacetime symmetries. Specifically, ξ_μ represents infinitesimal spacetime coordinate transformations, and ζ_μ represents the gauge symmetry of the antisymmetric tensor field. We see that local spacetime symmetry is necessary for interacting theories. As discussed in Section 1.4, they must appear due to the presence of massless spin-1 and spin-2 fields.

We should mention a technical point. Different renormalization schemes give different names to operators. The most common scheme is dimensional regularization (DR); the relation between DR operators and those above is:

$$\left[e^{ik\cdot X}\right]_{\mathrm{DR}} = \left[e^{ik\cdot X}\right]_{\mathrm{r}} \tag{3.6.25a}$$

$$\left[\partial_a X^\mu e^{ik\cdot X}\right]_{\mathrm{DR}} = \left[\partial_a X^\mu e^{ik\cdot X}\right]_{\mathrm{r}} \tag{3.6.25b}$$

$$\left[\partial_a X^\mu \partial_b X^v e^{ik\cdot X}\right]_{\mathrm{DR}} = \left[\partial_a X^\mu \partial_b X^v e^{ik\cdot X}\right]_{\mathrm{r}} - \frac{\alpha'}{12} g_{ab} \eta^{\mu v} R \left[e^{ik\cdot X}\right]_{\mathrm{r}} \tag{3.6.25c}$$

$$\left[\nabla_a \partial_b X^\mu e^{ik\cdot X}\right]_{\mathrm{DR}} = \left[\nabla_a \partial_b X^\mu e^{ik\cdot X}\right]_{\mathrm{r}} + i\frac{\alpha'}{12} g_{ab} k^\mu R \left[e^{ik\cdot X}\right]_{\mathrm{r}} \tag{3.6.25d}$$

Comparing with (3.6.20), we see that in DR, $\gamma = 0$, so the equations of motion hold within the renormalized operator. This is a convenience; many equations simplify. Specifically, in the spacetime gauge transformation (3.6.21), ϕ is invariant.

Open string vertex operators

Extension to open strings is direct. The tachyon vertex operator is:

$$g_o \int_{\partial M} ds \left[e^{ik\cdot X}\right]_r \tag{3.6.26}$$

which is Weyl invariant for $k^2 = 1/\alpha'$. The photon vertex operator is:

$$-i\frac{g_o}{(2\alpha')^{1/2}} e_\mu \int_{\partial M} ds \left[\dot{X}^\mu e^{ik\cdot X}\right]_r \tag{3.6.27}$$

where the relative normalization is obtained via state-operator mapping or unitarity.

The appearance of i is the same as in the closed string case, and the sign is for convenience. If $k^2 = 0$ and $k\cdot e = 0$, this operator is Weyl invariant. There exists an equivalence $e_\mu \cong e_\mu + \lambda k_\mu$, which is the spacetime gauge transformation, leaving $D-2$ transverse polarizations.

3.7 Strings in curved spacetime

We now examine a major new topic: strings moving in curved spacetime. Recall the point particle case, which has a natural extension to curved spacetime by replacing the flat metric $\eta_{\mu\nu}$ with a general metric $G_{\mu\nu}$:

$$S_{\text{pp}} = \frac{1}{2}\int d\tau \left(\eta^{-1} G_{\mu\nu}(X)\dot{X}^\mu \dot{X}^v - \eta m^2\right) \tag{3.7.1}$$

After eliminating η, this becomes the invariant proper time along the worldline. Its variation gives the geodesic equation, representing particle motion in a gravitational field.

Making the same substitution in the Polyakov action gives:

$$S_\sigma = \frac{1}{4\pi\alpha'}\int_M d^2\sigma g^{1/2} g^{ab} G_{\mu\nu}(X)\partial_a X^\mu \partial_b X^v \tag{3.7.2}$$

This is a natural guess, but readers might object: we have learned that the graviton itself is a state of the string. Roughly speaking, a curved spacetime is a coherent background of gravitons. Thus, simply putting a curved metric in (3.7.2) seems insufficient.

To see why this is reasonable, consider a nearly flat spacetime $G_{\mu\nu}(X) = \eta_{\mu\nu} + \chi_{\mu\nu}(X)$. At first order in χ, the action corresponds to the insertion of the graviton vertex operator (3.6.16) with:

$$\chi_{\mu v}(X) = -4\pi g_c e^{ik\cdot X} S_{\mu v} \tag{3.7.3}$$

So indeed, inserting a curved spacetime matches what we know. (3.7.2) can be viewed as describing a coherent state of gravitons by exponentiating the vertex operator.

This provides a natural generalization: incorporate backgrounds for other massless string states. From the vertex operator form (3.6.16), we have:

$$S_\sigma = \frac{1}{4\pi\alpha'}\int_M d^2\sigma g^{1/2}\left[\left(g^{ab}G_{\mu\nu}(X) + i\epsilon^{ab}B_{\mu\nu}(X)\right)\partial_a X^\mu \partial_b X^v + \alpha' R\Phi(X)\right] \tag{3.7.4}$$

The field $B_{\mu\nu}(X)$ is the antisymmetric tensor, and the dilaton $\Phi(X)$ involves both Φ and the trace of $G_{\mu\nu}$, as implied by (3.6.23c). As a check, consider spacetime gauge invariance. Under a change of variables $X'^\mu(X)$ representing a field redefinition, (3.7.4) is invariant if G, B transform as tensors and Φ as a scalar. This is a coordinate transformation. The action is also invariant under:

$$\delta B_{\mu\nu}(X) = \partial_\mu \zeta_v(X) - \partial_v \zeta_\mu(X) \tag{3.7.5}$$

which adds a total derivative. This is the generalization of the electromagnetic gauge transformation to a two-form potential.

The three-index field strength:

$$H_{\omega\mu v} = \partial_\omega B_{\mu v} + \partial_\mu B_{v\omega} + \partial_v B_{\omega\mu} \tag{3.7.6}$$

is invariant. There is a similar generalization to n-th order antisymmetric tensor potentials, which play an important role in superstrings.

This theory thus depends only on gauge-invariant quantities constructed from the metric and other fields. We can consider X^μ as coordinates on a manifold called the target space; thus X^μ defines an embedding:

$$\text{worldsheet} \;\to\; \text{target} \tag{3.7.7}$$

In string theory, target space is spacetime. Field theories like (3.7.2), where the kinetic term is field-dependent, are called non-linear σ-models.

For example, pion fields can be approximated as coordinates on the group manifold $SU(2)$. The non-linear σ-model is no longer quadratic in X^μ, so the path integral is now an interacting 2D quantum field theory. Expanding near a classical solution $X^\mu = x_0^\mu + Y^\mu$:

$$\begin{aligned} G_{\mu\nu}(X)\partial_a X^\mu \partial_b X^v = {} & \left[G_{\mu\nu}(x_0) + G_{\mu v,\omega}(x_0)\, Y^\omega \right. \\ & \left. + \frac{1}{2} G_{\mu v,\omega\rho}(x_0)\, Y^\omega Y^\rho + \ldots\right] \partial_a Y^\mu \partial_b Y^v \end{aligned} \tag{3.7.8}$$

The first term is the kinetic term for Y^μ; the next is a cubic interaction, and so on.

The coupling constants $G_{\mu\nu,\omega}$ involve derivatives at x_0. In a target space with characteristic curvature radius R_c, derivatives are of order R_c^{-1}. Thus, the effective dimensionless coupling is $\alpha'^{1/2} R_c^{-1}$. If R_c is much larger than the string length, the coupling is weak, and 2D perturbation theory is useful. Since the wavelength is large, we can ignore the internal structure and use low-energy effective field theory. String theory enters when determining the effective action.

By restricting to massless backgrounds, we implicitly assume $\alpha'^{1/2} R_c^{-1} \ll 1$: when wavelengths are large, massive states are not created. The non-linear σ-model is renormalizable. However, it is more useful to define the metric function itself as the coupling, rather than its expansion coefficients.

Weyl invariance

We have emphasized that Weyl invariance is crucial for string consistency. Only if the 2D QFT is Weyl invariant will (3.7.4) define a consistent string theory. This action is the most general one classically invariant under rigid Weyl transformations where $\delta\omega$ is constant.

This is easy to see: under a rigid Weyl transformation, a contraction with n derivatives is proportional to $2-n$. Local Weyl transformations are also necessary. In that domain, G and B are invariant, but the Φ term is not and contains quantum contributions.

In the limit where B and Φ are small and G is near η, we can find the Weyl transformation by writing $S_\sigma = S_P - V_1 + \ldots$, where V_1 is the vertex operator (3.6.16).

To first order, the Weyl transformation is given in (3.6.18). For convenience, taking $\gamma = 0$, and relating the transformation to the trace of the energy-momentum tensor as in (3.4.6):

$$T^a{}_a = -\frac{1}{2\alpha'}\beta^G_{\mu v} g^{ab}\partial_a X^\mu \partial_b X^v - \frac{i}{2\alpha'}\beta^B_{\mu v}\epsilon^{ab}\partial_a X^\mu \partial_b X^v - \frac{1}{2}\beta^\Phi R \tag{3.7.9}$$

To linear order in χ, B, and Φ:

$$\beta^G_{\mu v} \approx -\frac{\alpha'}{2}\left(\partial^2 \chi_{\mu\nu} - \partial_v \partial^\omega \chi_{\mu\omega} - \partial_\mu \partial^\omega \chi_{\omega v} + \partial_\mu \partial_v \chi^\omega_\omega\right) + 2\alpha' \partial_\mu \partial_v \Phi \tag{3.7.10a}$$

$$\beta^B_{\mu\nu} \approx -\frac{\alpha'}{2}\partial^\omega H_{\omega\mu\nu} \tag{3.7.10b}$$

$$\beta^\Phi \approx \frac{D-26}{6} - \frac{\alpha'}{2}\partial^2\Phi \tag{3.7.10c}$$

In β^Φ, we include the flat spacetime anomaly, including the ghost contribution.

The symbol β is used because these are renormalization group beta functions. Weyl anomaly (3.7.10) has further contributions from higher orders. For example, expanding to second order in cubic terms produces divergences as interactions approach each other.

Their OPE contains singularities like $|z|^{-2}\partial X\bar{\partial}X$ times derivatives of G. Diff invariance requires the integral to be truncated at $|z|\exp(\omega) > a_0$. This introduces dependence on the metric scale. β^G thus gains contributions proportional to $O(G_{,\gamma})^2$, which combine with linear terms to form the spacetime Ricci tensor.

We quote the result keeping terms up to two derivatives:

$$\beta^G_{\mu\nu} = \alpha'\mathbf{R}_{\mu v} + 2\alpha'\nabla_\mu\nabla_v\Phi - \frac{\alpha'}{4}H_{\mu\lambda\omega}H_v{}^{\lambda\omega} + O\left(\alpha'^2\right) \tag{3.7.11a}$$

$$\beta^B_{\mu\nu} = -\frac{\alpha'}{2}\nabla^\omega H_{\omega\mu v} + \alpha'\nabla^\omega\Phi H_{\omega\mu\nu} + O\left(\alpha'^2\right) \tag{3.7.11b}$$

$$\beta^\Phi = \frac{D-26}{6} - \frac{\alpha'}{2}\nabla^2\Phi + \alpha'\nabla_\omega\Phi\nabla^\omega\Phi - \frac{\alpha'}{24}H_{\mu\nu\lambda}H^{\mu\nu\lambda} + O\left(\alpha'^2\right) \tag{3.7.11c}$$

Several terms in (3.7.11) are recognizable from the linear approximation, but are now covariant under spacetime coordinate changes. $\mathbf{R}_{\mu\nu}$ is the spacetime Ricci tensor. The condition for Weyl invariance is $\beta^G = \beta^B = \beta^\Phi = 0$.

These are reasonable equations of motion. $\beta^G = 0$ is like the Einstein equation with sources from the antisymmetric tensor and dilaton. $\beta^B = 0$ is the generalization of Maxwell's equations for the antisymmetric tensor.

Backgrounds

Another feature of the field equations: the dilaton Φ only appears with derivatives, so it is invariant under a constant shift. Such a shift in (3.7.4) changes the action by a term proportional to the Euler characteristic, not affecting local Weyl invariance. Specifically, the background $G = \eta, B = 0, \Phi = \Phi_0$ is Weyl invariant for any constant Φ_0. This matches the flat action (3.2.2) with $\lambda = \Phi_0$.

λ determines the coupling strength. We now see this parameter is not a constant of the theory, but corresponds to different backgrounds in a single theory. The Regge slope α' is also not a free parameter; it defines the unit of length.

From a string perspective, changing G, B, Φ is just observing the same theory in a different background (state). This is an attractive feature: unlike the Standard Model, string theory has no free parameters—couplings are determined by the state and dynamics. Of course, this moves the difficulty to understanding why certain backgrounds are chosen.

It is surprising that Einstein's equations emerge as the condition for 2D Weyl invariance. Strings can only propagate consistently in backgrounds satisfying the field equations. This parallels our discovery that only on-shell vertex operators are meaningful.

The condition $D = 26$ comes from the R term. In the σ-model, this generalizes to $\beta^\Phi = 0$. We see in (3.7.14c) that β^Φ is proportional to $D-26$ at leading order, but has gradient corrections. If the field is not constant, we can have other values for D.

Exact solutions for $D \neq 26$ are known. For example:

$$G_{\mu\nu}(X) = \eta_{\mu\nu}, \quad B_{\mu\nu}(X) = 0, \quad \Phi(X) = V_\mu X^\mu \tag{3.7.12}$$

where $\beta = 0$ if $V_\mu V^\mu = \frac{26-D}{6\alpha'}$. This is exact because the fields are linear, making the path integral Gaussian. Varying g_{ab} reveals this is exactly the dilaton CFT with $c = D + 6\alpha' V_\mu V^\mu = 26$. Since Φ requires a large gradient, it does not describe our flat universe but may have cosmological applications.

Spacetime action

The field equations (3.7.15) can be derived from the spacetime action:

$$\begin{aligned}\mathbf{S} = \frac{1}{2\kappa_0^2} \int d^D x (-G)^{1/2} e^{-2\Phi} [&\frac{2(D-26)}{3\alpha'} + \mathbf{R} - \frac{1}{12} H_{\mu\nu\lambda} H^{\mu\nu\lambda} \\ &+ 4\partial_\mu \Phi \partial^\mu \Phi + O\left(\alpha'\right)]\end{aligned} \tag{3.7.13}$$

The normalization κ_0 is not determined by the field equations. It can be shown that:

$$\begin{aligned}\delta \mathbf{S} = -\frac{1}{2\kappa_0^2 \alpha'} \int d^D x (-G)^{1/2} e^{-2\Phi} &\left[\delta G_{\mu v} \beta^{G\mu v} + \delta B_{\mu v} \beta^{B\mu v} \right. \\ &\left. + \left(2\delta\Phi - \frac{1}{2} G^{\mu v} \delta G_{\mu\nu}\right) \left(\beta_\omega^{G\omega} - 4\beta^\Phi\right)\right]\end{aligned} \tag{3.7.14}$$

This effective action controls low-energy spacetime fields.

Field redefinitions like $\tilde{G}_{\mu\nu}(x) = \exp(2\omega(x)) G_{\mu\nu}(x)$ are useful. This is a spacetime Weyl transformation. Let $\omega = 2\left(\Phi_0 - \Phi\right)/(D-2)$ and define $\tilde{\Phi} = \Phi - \Phi_0$. The action becomes:

$$\begin{aligned}\mathbf{S} = &\frac{1}{2\kappa^2} \int d^D X (-\tilde{G})^{1/2} \left[-\frac{2(D-26)}{3\alpha'} e^{4\tilde{\Phi}/(D-2)} + \tilde{\mathbf{R}} \right. \\ &\left. - \frac{1}{12} e^{-8\tilde{\Phi}/(D-2)} H_{\mu\nu\lambda} \tilde{H}^{\mu\nu\lambda} - \frac{4}{D-2} \partial_\mu \tilde{\Phi} \tilde{\partial}^\mu \tilde{\Phi} + O\left(\alpha'\right)\right]\end{aligned} \tag{3.7.15}$$

The metric $\tilde{G}_{\mu\nu}$ is the Einstein metric, and $G_{\mu\nu}$ is the string metric. $\kappa = \kappa_0 e^{\Phi_0}$ is the observed gravitational coupling.

Compactification and CFT

The four string theories studied so far (oriented/unoriented, with/without boundary) share features: automatic inclusion of general relativity, but also the need for $D = 26$, the existence of tachyons, and the absence of fermions. However, general relativity allows extra dimensions: some can be large and flat while others are small and curled.

The metric:

$$g_{MN} = \left[\begin{array}{cc} \eta_{\mu v} & 0 \\ 0 & g_{mn}\left(x^p\right) \end{array}\right] \tag{3.7.16}$$

where M, N are 26 coordinates, split into 4 spacetime coordinates and 22 internal coordinates. If the internal space is compact and Ricci flat ($\mathbf{R}_{mn} = 0$), this satisfies the field equations.

Physics on scales much larger than the internal space matches d-dimensional Minkowski space. A necessary condition for string consistency is 2D diff invariance. Weyl invariance is more technical; its loss produces extra degrees of freedom.

We make an additional technical assumption: X^0 appears only in the form:

$$-\frac{1}{4\pi\alpha'} \int_M d^2\sigma g^{1/2} g^{ab} \partial_a X^0 \partial_b X^0 \tag{3.7.17}$$

meaning the background is static. This assumption places the "wrong-sign" signature of X^0 in an explicit form.

Local invariance leads to a proposal: replace the 25 spatial fields X^μ with any unitary CFT with $c = \tilde{c} = 25$. This ensures the 2D theory, including X^0 and ghosts, couples to a curved metric consistently. Unitarity is needed to remove unphysical X^0 excitations.

For low-energy 4D physics, we keep $X^0 \ldots X^3$ and replace the rest with a $c = 22$ compact unitary CFT. Compactification implies a discrete spectrum. All these theories have tachyons (product of unitary $|1\rangle$ in internal CFT and 4D ground state).

How large is our generalization? We could introduce any fields with different quantum numbers. However, in 2D, many different theories are equivalent. Generally, all string theories based on different CFTs but same worldsheet symmetries and topologies are different ground states of a unified theory. Though we call them "different theories" due to different Lagrangians, they are just different vacua.

Chapter 4

The string spectrum

4.1 Old covariant quantization

In the conformal gauge, the worldsheet fields consist of X^μ and the Faddeev-Popov ghost fields b_{ab} and c^a. The Hilbert space is larger than the actual physical spectrum of the string: D sets of α^μ oscillators include unphysical oscillations of the coordinate system, and ghost oscillators are also present. As is general in covariant gauges, there are negative-norm states originating from timelike oscillators (due to the commutator being proportional to the spacetime metric $\eta_{\mu\nu}$) and from the ghost fields.

The actual physical space is smaller. How do we identify this smaller space? Consider the amplitude of an initial state $|i\rangle$ propagating on an infinite cylinder to a final state $|f\rangle$. Suppose we initially fix the metric to the form $g_{ab}(\sigma)$ using local symmetries. Now consider a different gauge where the metric is $g_{ab}(\sigma) + \delta g_{ab}(\sigma)$. Physical amplitudes should not depend on this choice. Of course, for a change in the metric, we know how the path integral changes. From the definition of T^{ab} in (3.4.4):

$$\delta\langle f|i\rangle = -\frac{1}{4\pi}\int \mathrm{d}^2\sigma\, g(\sigma)^{1/2}\delta g_{ab}(\sigma)\langle f|T^{ab}(\sigma)|i\rangle\,. \tag{4.1.1}$$

To make this zero for an arbitrary variation of the metric, we require for any physical states $|\psi\rangle$ and $|\psi'\rangle$:

$$\langle\psi|T^{ab}(\sigma)|\psi'\rangle = 0\,. \tag{4.1.2}$$

Consider another way to look at it. The original equations of motion obtained from the variation of g_{ab} are $T^{ab} = 0$. After fixing the gauge, this does not hold as an operator equation: since we do not vary g_{ab} in the gauge-fixed theory, an equation is missing. Condition (4.1.2) indicates that the missing equations of motion must hold for any matrix element between physical states. When we transform the gauge, the variation of the Faddeev–Popov determinant must be taken into account. Thus, the energy-momentum tensor in the matrix elements is the sum of the X^μ contribution and the ghost contribution:

$$T_{ab} = T^X_{ab} + T^{\mathrm{g}}_{ab}\,. \tag{4.1.3}$$

X^μ can be replaced by a more general CFT (which we refer to as the *matter* CFT). In this case:

$$T_{ab} = T^{\mathrm{m}}_{ab} + T^{\mathrm{g}}_{ab}\,. \tag{4.1.4}$$

In the remainder of this section, we will impose condition (4.1.2) in a simple but somewhat ad hoc manner, known as *old covariant quantization* (OCQ), which is sufficient for many purposes. In the next section, we will adopt a more systematic approach: BRST quantization. They are effectively equivalent, as will be proven in Section 4.4.

In this ad hoc method, we simply ignore the ghost fields and attempt to constrain the matter Hilbert space such that the missing equations of motion $T^{\mathrm{m}}_{ab} = 0$ hold for matrix elements. In terms of Laurent coefficients, this is $L^{\mathrm{m}}_n = 0$, and for closed strings, $\tilde{L}^{\mathrm{m}}_n = 0$. One might first attempt to require physical states to satisfy $L^{\mathrm{m}}_n|\psi\rangle = 0$ for all n, but this is too strong; acting with L^{m}_m on this equation and forming the commutator leads to inconsistencies due to the central charge in the Virasoro algebra. However, it is sufficient that the Virasoro lowering operators and the zero operator annihilate physical states:

$$(L^{\mathrm{m}}_n + A\delta_{n,0})|\psi\rangle = 0 \quad \text{for } n \geq 0 \,. \tag{4.1.5}$$

Then, for $n < 0$, we have:

$$\langle\psi|L^{\mathrm{m}}_n|\psi'\rangle = \langle L^{\mathrm{m}}_{-n}\psi|\psi'\rangle = 0 \,. \tag{4.1.6}$$

We used:

$$L^{\mathrm{m}\dagger}_n = L^{\mathrm{m}}_{-n} \tag{4.1.7}$$

which stems from the Hermiticity of the energy-momentum tensor. Equation (4.1.5) is consistent with the Virasoro algebra. When $n = 0$, we introduce a possible ordering constant as usual. States satisfying (4.1.5) are called physical. In the terminology of (2.9.8), physical states are highest weight states with weight $-A$. Equation (4.1.5) is analogous to the Gupta-Bleuler quantization in quantum electrodynamics.

Using the adjoint (4.1.7), one can see that states of the form:

$$|\chi\rangle = \sum_{n=1}^{\infty} L^{\mathrm{m}}_{-n}|\chi_n\rangle \tag{4.1.8}$$

are orthogonal to any $|\chi_n\rangle$. Such states are called spurious. A state that is both physical and spurious is called null. If $|\psi\rangle$ is physical and $|\chi\rangle$ is null, then $|\psi\rangle + |\chi\rangle$ is also physical, and the inner product of any physical state with it is the same as its inner product with $|\psi\rangle$. Thus, these two states are physically indistinguishable, and we have the equivalence relation:

$$|\psi\rangle \cong |\psi\rangle + |\chi\rangle \,. \tag{4.1.9}$$

The true physical states are then the set of equivalence classes:

$$\mathscr{H}_{\mathrm{OCQ}} = \frac{\mathscr{H}_{\mathrm{phys}}}{\mathscr{H}_{\mathrm{null}}} \,. \tag{4.1.10}$$

Let us see how this works for the first two levels of the open string in flat spacetime without assuming $D = 26$. The only relevant terms are:

$$L^{\mathrm{m}}_0 = \alpha' p^2 + \alpha_{-1}\cdot\alpha_1 + \cdots \,, \tag{4.1.11a}$$

$$L^{\mathrm{m}}_{\pm1} = (2\alpha')^{1/2} p\cdot\alpha_{\pm1} + \cdots \,. \tag{4.1.11b}$$

Proof. For the open string, according to (2.7.25):

$$\alpha^\mu_0 = (2\alpha')^{1/2}p^\mu \,, \quad \alpha' P^2 = \frac{\alpha'}{2\alpha'}\alpha_0\cdot\alpha_0 = \frac{1}{2}\alpha_0\cdot\alpha_0 \,, \quad (2\alpha')^{1/2}p\cdot\alpha_{\pm1} = \alpha_0\cdot\alpha_{\pm1}$$

□

At the lowest mass level, the only state is $|0;k\rangle$. At this level, the only non-trivial condition is $(L^{\mathrm{m}}_0 + A)|\psi\rangle = 0$, yielding $m^2 = A/\alpha'$. No null states exist at this level. Since the Virasoro generators in the spurious state (4.1.8) consist only of raising operators, there exists one equivalence class corresponding to a scalar particle.

Remark. *Proof for $m^2 = A/\alpha'$: Since $L_0^m|\psi\rangle = \alpha' p^2|\psi\rangle = -\alpha' m^2|\psi\rangle$, it follows that $\alpha' m^2 = A$.*

At the next level, there are D states:

$$|e;k\rangle = e\cdot\alpha_{-1}|0;k\rangle\ . \tag{4.1.12}$$

The norm is:

$$\begin{aligned}\langle e;k|e;k'\rangle &= \langle 0;k|e^*\cdot\alpha_1\, e\cdot\alpha_{-1}|0;k'\rangle\\ &= \langle 0;k|(e^*\cdot e + e^*\cdot\alpha_{-1}e\cdot\alpha_1)|0;k'\rangle\\ &= e^{\mu *}e_\mu(2\pi)^D\delta^D(k-k')\ .\end{aligned} \tag{4.1.13}$$

We used:

$$\alpha_n^{\mu\dagger} = \alpha_{-n}^\mu\ , \tag{4.1.14}$$

stemming from the Hermiticity of X^μ, and:

$$\langle 0;k|0;k'\rangle = (2\pi)^D\delta^D(k-k')\ , \tag{4.1.15}$$

which arises from momentum conservation. Timelike excitations result in one negative-norm state.

$$\alpha_{-1}^0\,|0;k\rangle = e_\mu\alpha_{-1}^\mu\,|0;k\rangle \qquad \text{(with } e_\mu = (1,0,0,\dots,0))$$

The $L_0^{\rm m}$ condition gives:

$$m^2 = \frac{1+A}{\alpha'}\ . \tag{4.1.16}$$

Proof.

$$\begin{aligned}L_0^m|e;k\rangle &= \alpha' p^2|e;k\rangle + \alpha_{-1}\cdot\alpha_1 e\cdot\alpha_{-1}|0;k\rangle\\ &= -\alpha' m^2|e;k\rangle + \alpha_{-1}^\mu e^\nu(\eta_{\mu\nu} + \alpha_{-1\nu}\alpha_{1,\mu})|0;k\rangle\\ &= -\alpha' m^2|e;k\rangle + e\cdot\alpha_{-1}|0;k\rangle\\ &= (1-\alpha' m^2)|e;k\rangle\end{aligned}$$

Therefore $A = \alpha' m^2 - 1$. □

The other non-trivial physical state condition is:

$$\begin{aligned}L_1^{\rm m}|e;k\rangle &\propto p\cdot\alpha_1\, e\cdot\alpha_{-1}|0;k\rangle = k_\mu\alpha_1^\mu e_\nu\alpha_{-1}^\nu\,|0;k\rangle\\ &= k_\mu e_\nu(\alpha_1^\mu\alpha_{-1}^\nu)\,|0;k\rangle = k_\mu e_\nu(\alpha_{-1}^\mu\alpha_1^\nu + \eta^{\mu\nu})\,|0;k\rangle\\ &= e\cdot k|0;k\rangle = 0\end{aligned} \tag{4.1.17}$$

thus we require $k\cdot e = 0$. For timelike polarization we have $k\cdot e = k^0$ which together with $k_\mu k^\mu = 0$ implies $k^0 \neq 0$ and therefore we have $L_1^{\rm m}|e;k\rangle \neq 0$. As such this state does not belong to the physical spectrum. At this level, there are spurious states that exist:

$$L_{-1}^{\rm m}|0;k\rangle = (2\alpha')^{1/2}k\cdot\alpha_{-1}|0;k\rangle\ . \tag{4.1.18}$$

That is, $e^\mu \propto k^\mu$ is spurious. One can classify this into three cases:

1. If $A > -1$, the mass squared is positive. Returning to the rest frame ($\mathbf{k} = 0, k^0 \neq 0$), the physical state condition ($e \cdot k = 0 \Rightarrow e^0 = 0$) removes the negative-norm timelike polarization. Spurious states are not physical, so there are no null states and the spectrum consists of $D - 1$ positive-norm states of a massive vector particle.

2. If $A = -1$, the mass squared is 0. $k \cdot k = 0$, so spurious states are physical and null. Thus:

$$k \cdot e = 0, \quad e_\mu \cong e_\mu + \gamma k_\mu \,. \tag{4.1.19}$$

   This describes $D - 2$ positive-norm states of a massless vector particle.

3. If $A < -1$, the mass squared is negative. The momentum is spacelike ($\mathbf{k} \neq 0, k^0 = 0$), so the physical state condition removes one positive-norm spacelike polarization. Spurious states are not physical, leaving a vector tachyon with $D - 2$ positive-norm states and one negative-norm state.

Case (3) is unacceptable. Case (2) is identical to light-cone quantization. Case (1) differs from the light-cone spectrum as it has a different mass and extra states at the first level, though without obvious inconsistencies.

The result for the next level is quite interesting: it depends on the constant A and the spacetime dimension D. Restricting $A = -1$ from the discovery at the first excited level, it coincides with the light-cone spectrum only if $D = 26$. If $D < 26$, the OCQ spectrum has positive-norm states but more than the light-cone quantization; for $D > 26$, there are negative-norm states. When $A = -1$ and $D = 26$, OCQ is identical to light-cone quantization at all levels:

$$\mathscr{H}_{\mathrm{OCQ}} = \mathscr{H}_{\text{light-cone}} \ , \tag{4.1.20}$$

which will be proven in Section 4.4. Only in this case are consistent interactions known.

Generalization to closed strings is straightforward. There are two sets of oscillators and two sets of Virasoro algebras, so at each level, the spectrum is the product of a pair of open string spectra. Thus, the first two levels are:

$$|0;k\rangle, \quad m^2 = -\frac{4}{\alpha'} \ ; \tag{4.1.21a}$$

$$e_{\mu\nu}\alpha^\mu_{-1}\tilde{\alpha}^\nu_{-1}|0;k\rangle, \quad m^2 = 0, k^\mu e_{\mu\nu} = k^v e_{\mu\nu} = 0 \,, \tag{4.1.21b}$$

$$e_{\mu\nu} \cong e_{\mu\nu} + a_\mu k_\nu + k_\mu b_\nu \ , \quad a \cdot k = b \cdot k = 0 \,. \tag{4.1.21c}$$

The relevant values are $A = -1$ and $D = 26$. As in light-cone quantization, there are $(D-2)^2$ massless states forming a traceless symmetric tensor, an antisymmetric tensor, and a scalar.

Mnemonic

For more general string theories, there is a mnemonic to obtain the zero-point constant. As derived in the next section, the L_0^{m} condition can be understood as:

$$(L_0^{\mathrm{m}} + L_0^{\mathrm{g}})|\psi, \downarrow\rangle = 0 \,. \tag{4.1.22}$$

That is, ghost contributions are introduced, where ghosts are in the ground state $|\downarrow\rangle$ which has $L_0^{\mathrm{g}} = -1$. The L_0 generator and the Hamiltonian differ by an offset (2.6.10) proportional to the central charge, but since the total central charge in string theory is 0, we can write this condition as:

$$(H^{\mathrm{m}} + H^{\mathrm{g}})|\psi, \downarrow\rangle = 0 \,. \tag{4.1.23}$$

Now, apply the mnemonic given at the end of Chapter 2 for zero-point energy. Specifically, ghosts always cancel $\mu = 0, 1$ oscillators because they have the same periodicity but opposite statistics. Thus, the rule is: A is given by the zero-point energy of the transverse oscillators. This is the same rule as in light-cone coordinates, which here gives $A = 24\left(-\frac{1}{24}\right) = -1$.

Incidentally, for counting physical states (not their precise form), one can always ignore the ghosts and $\mu = 0, 1$ oscillators and count transverse excitations as in the light-cone gauge.

Condition (4.1.22) requires the physical states to have weight 1. Since physical states require matter states to be highest weight states, the vertex operators must be weight 1 or $(1, 1)$ tensors. This matches the condition $A = -1$ from another perspective: that integrated vertex operators must be conformally invariant, as found in Section 3.6.

4.2 BRST quantization

We now turn to a more systematic study of the spectrum. Condition (4.1.2) is insufficient to guarantee gauge invariance. It implies invariance under an arbitrary fixed choice of g_{ab}, but this is not the most general gauge. In the light-cone gauge, we impose conditions on both X^μ and g_{ab}. To investigate the most general variations in gauge conditions, we must allow δg_{ab} to be an *operator*, that is, let it depend on the fields in the path integral.

To derive the complete invariance condition, it is useful to take a more general and abstract view. Consider a path integral with local symmetries, where the path integral fields are denoted ϕ_i; in our case, these are $X^\mu(\sigma)$ and $g_{ab}(\sigma)$. We use a compact notation where i also represents the coordinate σ. Gauge invariance is $\epsilon^\alpha \delta_\alpha$, where α also includes coordinates. Since we can always decompose a complex parameter into real and imaginary parts, we assume the gauge parameter ϵ^α is real. The gauge transformations satisfy the algebra:

$$[\delta_\alpha, \delta_\beta] = f^\gamma{}_{\alpha\beta}\delta_\gamma \,. \tag{4.2.1}$$

Now fix the gauge with the condition:

$$F^A(\phi) = 0 \tag{4.2.2}$$

where A again includes coordinates. Following the Faddeev-Popov procedure in Section 3.3, the path integral becomes:

$$\int \frac{[\mathrm{d}\phi_i]}{V_{\text{gauge}}} \exp(-S_1) \to \int [\mathrm{d}\phi_i \, \mathrm{d}B_A \, \mathrm{d}b_A \, \mathrm{d}c^\alpha] \, \exp(-S_1 - S_2 - S_3) \,, \tag{4.2.3}$$

where S_1 is the original gauge-invariant action, S_2 is the gauge-fixing action:

$$S_2 = -iB_A F^A(\phi) \,, \tag{4.2.4}$$

and S_3 is the Faddeev-Popov action:

$$S_3 = b_A c^\alpha \delta_\alpha F^A(\phi) \,. \tag{4.2.5}$$

We introduced field B_A to generate the integral representation of the gauge fixing $\delta(F^A)$.

Two things are noteworthy regarding this action. First, it is invariant under the Becchi–Rouet–Stora–Tyutin (BRST) transformation:

$$\delta_{\mathrm{B}}\phi_i = -i\epsilon c^\alpha \delta_\alpha \phi_i \,, \tag{4.2.6a}$$

$$\delta_{\mathrm{B}} B_A = 0 \,, \tag{4.2.6b}$$

$$\delta_{\mathrm{B}} b_A = \epsilon B_A \tag{4.2.6c}$$

$$\delta_{\mathrm{B}} c^\alpha = \frac{i}{2}\epsilon f^\alpha{}_{\beta\gamma} c^\beta c^\gamma \,. \tag{4.2.6d}$$

This transformation mixes commuting and anticommuting quantities, requiring ϵ to be anticommuting. Ghost number is conserved, with c^α having ghost number 1, b_A and ϵ having ghost number -1, and other fields having 0. Since the action of δ_{B} on ϕ_i is exactly a gauge transformation with parameter $i\epsilon c^\alpha$, the original action S_1 is itself invariant. The variation of S_2 cancels the variation of b_A in S_3, while the variations of $\delta_\alpha F^A$ and c^α in S_3 cancel. The second key property is:

$$\delta_{\mathrm{B}}(b_A F^A) = i\epsilon(S_2 + S_3) \,. \tag{4.2.7}$$

Now consider a small change in the gauge-fixing condition δF. The change in the gauge-fixing and ghost actions yields:

$$\begin{aligned}\epsilon\delta\langle f|i\rangle &= i\langle f|\delta_{\mathrm{B}}(b_A\delta F^A)|i\rangle \\ &= -\epsilon\langle f|\{Q_{\mathrm{B}}, b_A\delta F^A\}|i\rangle \,,\end{aligned} \tag{4.2.8}$$

where we expressed the BRST variation as the anticommutator with the corresponding conserved charge Q_B. For this to be zero for any δF, physical states must satisfy:

$$\langle\psi|\{Q_{\mathrm{B}}, b_A\delta F^A\}|\psi'\rangle = 0 \,. \tag{4.2.9}$$

To hold for arbitrary δF, we must have:

$$Q_{\mathrm{B}}|\psi\rangle = Q_{\mathrm{B}}|\psi'\rangle = 0 \,. \tag{4.2.10}$$

This is the important condition: physical states must be BRST invariant. We assumed $Q_{\mathrm{B}}^\dagger = Q_{\mathrm{B}}$. There are several ways to see this must be the case. One is that if $Q_{\mathrm{B}}^\dagger$ were different, there would have to be other symmetries, but no such candidates exist. A better argument is that c^α and b_A are like anticommuting versions of the gauge parameter ϵ^α and the Lagrange multiplier B_A, and thus inherit their real properties.

On the other hand, for any constant matrix M_{AB}, we can add a term to the action proportional to:

$$\epsilon^{-1}\delta_{\mathrm{B}}(b_A B_B M^{AB}) = -B_A B_B M^{AB} \tag{4.2.11}$$

By the above discussion, the amplitudes between physical states remain unaffected. Integrating over B_A now produces a Gaussian instead of a δ-function: these are Gaussian averaged gauges, which include the covariant α-gauges of gauge theory.

There is a more critical concept. For the BRST charge to remain conserved while moving through the space of gauge choices, it must commute with the change in the Hamiltonian:

$$\begin{aligned}0 &= [Q_{\mathrm{B}}, \{Q_{\mathrm{B}}, b_A\delta F^A\}] \\ &= Q_{\mathrm{B}}^2 b_A\delta F^A - Q_{\mathrm{B}} b_A\delta F^A Q_{\mathrm{B}} + Q_{\mathrm{B}} b_A\delta F^A Q_{\mathrm{B}} - b_A\delta F^A Q_{\mathrm{B}}^2 \\ &= [Q_{\mathrm{B}}^2, b_A\delta F^A] \,.\end{aligned} \tag{4.2.12}$$

For this to be zero for general gauge changes, we need:

$$Q_{\mathrm{B}}^2 = 0 \,. \tag{4.2.13}$$

That is, the BRST transformation is nilpotent. Since Q_B^2 has ghost number 2, the possibility of it being a constant is excluded. One can verify that all fields are invariant under two BRST transformations (4.2.6). Specifically:

$$\delta_{\mathrm{B}}(\delta'_{\mathrm{B}} c^\alpha) = -\frac{1}{2}\epsilon\epsilon' f^\alpha{}_{\beta\gamma} f^\gamma{}_{\delta\epsilon} c^\beta c^\delta c^\epsilon = 0 \,. \tag{4.2.14}$$

The product of ghosts is antisymmetric in β, δ, ϵ, and thus the product of structure constants is zero due to the Jacobi identity.

We should mention that we made two simplifying assumptions for the gauge algebra (4.2.1). First, the structure constants $f^\alpha{}_{\beta\gamma}$ are constants, independent of the fields; second, the algebra has no additional terms proportional to the equations of motion on the right side. More generally, both assumptions are violated. In these cases, the BRST system described does not yield nilpotent transformations, and we would need the Batalin–Vilkovisky (BV) system. The BV system has various applications in string theory, but we will not require it.

The nilpotency of Q_B has an important consequence. States of the form:

$$Q_{\mathrm{B}}|\chi\rangle\ , \tag{4.2.15}$$

where χ is arbitrary, will be annihilated by Q_{B} and are thus physical. However, such a state is orthogonal to all physical states, including itself: if $Q_{\mathrm{B}}|\psi\rangle = 0$, then:

$$\langle\psi|(Q_{\mathrm{B}}|\chi\rangle) = (\langle\psi|Q_{\mathrm{B}})|\chi\rangle = 0\ . \tag{4.2.16}$$

Thus, all physical amplitudes containing such null states are zero. Two physical states differing only by a null state:

$$|\psi'\rangle = |\psi\rangle + Q_{\mathrm{B}}|\chi\rangle \tag{4.2.17}$$

have identical inner products with all physical states. So, as in OCQ, we identify the true physical space with the set of equivalence classes, where states differing by a null state are equivalent. This is a natural construction for a nilpotent operator, called the *cohomology* of Q_{B}. Other examples of nilpotent operators include the exterior derivative in differential geometry and the boundary operator in topology. In cohomology, the term *closed* refers to states annihilated by Q_{B}, and *exact* refers to states of the form (4.2.15). Thus, our procedure is:

$$\mathscr{H}_{\mathrm{BRST}} = \frac{\mathscr{H}_{\mathrm{closed}}}{\mathscr{H}_{\mathrm{exact}}}\ . \tag{4.2.18}$$

We will clearly see in the remainder of this chapter that this space in string theory has the expected form. Clearly, the invariance condition removes one set of unphysical X^μ and ghost oscillators, while the equivalence relation removes another set.

Point particle example

Now let us examine the example of a point particle. Expanding the compact notation, the local symmetry is coordinate reparameterization $\delta\tau(\tau)$, so the index α becomes τ. A basis for infinitesimal transformations is $\delta_{\tau_1}\tau(\tau) = \delta(\tau - \tau_1)$. Their action on the fields is:

$$\delta_{\tau_1}X^\mu(\tau) = -\delta(\tau-\tau_1)\partial_\tau X^\mu(\tau), \qquad \delta_{\tau_1}e(\tau) = -\partial_\tau[\delta(\tau-\tau_1)e(\tau)]\ . \tag{4.2.19}$$

Applying a second transformation and forming the commutator, we have:

$$\begin{aligned}
&[\delta_{\tau_1},\delta_{\tau_2}]X^\mu(\tau) \\
&\quad= -\Big[\delta(\tau-\tau_1)\partial_\tau\delta(\tau-\tau_2) - \delta(\tau-\tau_2)\partial_\tau\delta(\tau-\tau_1)\Big]\partial_\tau X^\mu(\tau) \\
&\quad\equiv \int \mathrm{d}\tau_3\, f^{\tau_3}{}_{\tau_1\tau_2}\delta_{\tau_3}X^\mu(\tau)\ .
\end{aligned} \tag{4.2.20}$$

From the commutator, we have determined the structure constants:

$$f^{\tau_3}{}_{\tau_1\tau_2} = \delta(\tau_3-\tau_1)\partial_{\tau_3}\delta(\tau_3-\tau_2) - \delta(\tau_3-\tau_2)\partial_{\tau_3}\delta(\tau_3-\tau_1)\ . \tag{4.2.21}$$

The BRST transformation is then:

$$\delta_{\mathrm{B}}X^\mu = i\epsilon c\dot{X}^\mu\ , \tag{4.2.22a}$$

$$\delta_{\mathrm{B}}e = i\epsilon(\dot{ce})\ , \tag{4.2.22b}$$

$$\delta_{\mathrm{B}}B = 0\ , \tag{4.2.22c}$$

$$\delta_{\mathrm{B}} b = \epsilon B \,, \tag{4.2.22d}$$
$$\delta_{\mathrm{B}} c = i\epsilon c\dot{c} \,. \tag{4.2.22e}$$

Gauge $e(\tau) = 1$ is similar to the unit gauge of the string, using a single coordinate degree of freedom to fix each component of the tetrad, so $F(\tau) = 1 - e(\tau)$. The gauge-fixed action is:

$$S = \int \mathrm{d}\tau \left(\frac{1}{2} e^{-1} \dot{X}^\mu \dot{X}_\mu + \frac{1}{2} e m^2 + iB(e-1) - e\dot{b}c \right) \tag{4.2.23}$$

We find it convenient to integrate out B, thus fixing $e = 1$. This leaves the fields X^μ, b and c, with the action:

$$S = \int \mathrm{d}\tau \left(\frac{1}{2} \dot{X}^\mu \dot{X}_\mu + \frac{1}{2} m^2 - \dot{b}c \right) \tag{4.2.24}$$

and the BRST transformation:

$$\delta_{\mathrm{B}} X^\mu = i\epsilon c \dot{X}^\mu \,, \tag{4.2.25a}$$
$$\delta_{\mathrm{B}} b = i\epsilon \bigg(-\frac{1}{2} \dot{X}^\mu \dot{X}_\mu + \frac{1}{2} m^2 + b\dot{c} \bigg) \,, \tag{4.2.25b}$$
$$\delta_{\mathrm{B}} c = i\epsilon c\dot{c} \,. \tag{4.2.25c}$$

Since B no longer appears, we used the equation of motion from e to replace B in the transformation rule for b. One can verify that the new transformations (4.2.25) are a symmetry of the action and are nilpotent. This is applicable when field B^A is integrated out; although $\delta_{\mathrm{B}} \delta'_{\mathrm{B}} b$ is no longer identically zero, it is proportional to the equations of motion. This is satisfactory: $Q_{\mathrm{B}}^2 = 0$ holds as an operator equation, which is exactly what we need. The canonical commutators are:

$$[p^\mu, X^v] = -i\eta^{\mu v}, \qquad \{b, c\} = 1 \,, \tag{4.2.26}$$

where $p^\mu = i\dot{X}^\mu$, with i from the Euclidean spacetime signature. The Hamiltonian is $H = \frac{1}{2}\left(p^2 + m^2\right)$, and the Noether procedure yields the BRST operator:

$$Q_{\mathrm{B}} = cH \,. \tag{4.2.27}$$

The structure here is similar to what we seek for the string: the constraint (missing equation of motion) is $H = 0$, and the BRST operator is c times this operator.

The ghosts generate a system of two states, so a complete basis for the states is $|k, \downarrow\rangle, |k, \uparrow\rangle$, where:

$$p^\mu |k, \downarrow\rangle = k^\mu |k, \downarrow\rangle \,, \quad p^\mu |k, \uparrow\rangle = k^\mu |k, \uparrow\rangle \,, \tag{4.2.28a}$$
$$b|k, \downarrow\rangle = 0 \,, \quad b|k, \uparrow\rangle = |k, \downarrow\rangle \,, \tag{4.2.28b}$$
$$c|k, \downarrow\rangle = |k, \uparrow\rangle \,, \quad c|k, \uparrow\rangle = 0 \,. \tag{4.2.28c}$$

The action of the BRST operator on them is:

$$Q_{\mathrm{B}} |k, \downarrow\rangle = \frac{1}{2}(k^2 + m^2)|k, \uparrow\rangle, \quad Q_{\mathrm{B}} |k, \uparrow\rangle = 0 \,. \tag{4.2.29}$$

The resulting closed states are:

$$|k, \downarrow\rangle \,, \quad k^2 + m^2 = 0 \,, \tag{4.2.30a}$$
$$|k, \uparrow\rangle \,, \quad \text{all } k^\mu \,, \tag{4.2.30b}$$

and exact states are:

$$|k, \uparrow\rangle, \quad k^2 + m^2 \neq 0 \,. \tag{4.2.31}$$

The non-exact closed states are:

$$|k,\downarrow\rangle\,,\ \ k^2+m^2=0\,;\quad |k,\uparrow\rangle\,,\ \ k^2+m^2=0\,. \tag{4.2.32}$$

Thus, physical states must satisfy the mass-shell condition, but we have two copies of the expected spectrum. In fact, only states satisfying the extra condition:

$$b|\psi\rangle=0 \tag{4.2.33}$$

(the state $|k,\downarrow\rangle$) appear in physical amplitudes. The origin of this extra condition is kinematical. For $k^2+m^2\neq 0$, states $|k,\uparrow\rangle$ are exact—they are proportional to all physical states, and the amplitudes are identically zero. Amplitudes can only be proportional to $\delta(k^2+m^2)$. But amplitudes in field theory and string theory, while having poles and cuts, do not have δ-functions (except in $D=2$ dimensions, where kinematics is special). So they must be zero.

4.3 BRST quantization of the string

In string theory, the BRST transformations are:

$$\delta_{\mathrm{B}}X^\mu=i\epsilon(c\partial+\tilde{c}\bar{\partial})X^\mu\,, \tag{4.3.1a}$$

$$\delta_{\mathrm{B}}b=i\epsilon(T^X+T^{\mathrm{g}})\,,\quad \delta_{\mathrm{B}}\tilde{b}=i\epsilon(\tilde{T}^X+\tilde{T}^{\mathrm{g}})\,, \tag{4.3.1b}$$

$$\delta_{\mathrm{B}}c=i\epsilon c\partial c\,,\qquad \delta_{\mathrm{B}}\tilde{c}=i\epsilon\tilde{c}\bar{\partial}\tilde{c}\,. \tag{4.3.1c}$$

One can derive these from the point particle example. To the sum of the Polyakov and ghost actions, add the gauge-fixing term:

$$\frac{i}{4\pi}\int \mathrm{d}^2\sigma\, g^{1/2}B^{ab}(\delta_{ab}-g_{ab})\,. \tag{4.3.2}$$

After integrating over B_{ab} and replacing B_{ab} in the transformations with the equations of motion for g_{ab}, the transformation $\delta_{\mathrm{B}}b_{ab}=\epsilon B_{ab}$ becomes (4.3.1). The Weyl ghost is just a Lagrange multiplier making b_{ab} traceless. One can verify these transformations are nilpotent up to the equations of motion. The similarity between the BRST transformations (4.3.1) and the particle case (4.2.25) is obvious. Replacing X^μ with a more general matter CFT, matter field transformations become conformal transformations with $v(z)=c(z)$, and in the b transformation, T^{m} replaces T^X.

The Noether theorem gives the BRST current:

$$\begin{aligned} j_{\mathrm{B}}&=cT^{\mathrm{m}}+\frac{1}{2}:cT^{\mathrm{g}}:+\frac{3}{2}\partial^2c\\ &=cT^{\mathrm{m}}+:bc\partial c:+\frac{3}{2}\partial^2c \end{aligned} \tag{4.3.3}$$

and similarly $\tilde{\jmath}_B$. The last term in the current is a total derivative and does not contribute to the BRST charge; it is added manually to make the BRST current a tensor. The OPE of the current with the ghost fields and a general matter tensor field is:

$$j_{\mathrm{B}}(z)b(0)\sim\frac{3}{z^3}+\frac{1}{z^2}j^{\mathrm{g}}(0)+\frac{1}{z}T^{\mathrm{m+g}}(0)\,, \tag{4.3.4a}$$

$$j_{\mathrm{B}}(z)c(0)\sim\frac{1}{z}c\partial c(0)\,, \tag{4.3.4b}$$

$$j_{\mathrm{B}}(z)\mathcal{O}^{\mathrm{m}}(0,0)\sim\frac{h}{z^2}c\mathcal{O}^{\mathrm{m}}(0,0)+\frac{1}{z}\big[h(\partial c)\mathcal{O}^{\mathrm{m}}(0,0)+c\partial\mathcal{O}^{\mathrm{m}}(0,0)\big]\,. \tag{4.3.4c}$$

The simple poles reflect the BRST transformations of these fields.

The BRST operator is:

$$Q_{\mathrm{B}} = \frac{1}{2\pi i} \oint (\mathrm{d}z\, j_{\mathrm{B}} - \mathrm{d}\bar{z}\, \tilde{\jmath}_{\mathrm{B}}) \,. \tag{4.3.5}$$

Via standard contour discussion, the OPE implies:

$$\{Q_{\mathrm{B}}, b_m\} = L_m^{\mathrm{m}} + L_m^{\mathrm{g}} \,. \tag{4.3.6}$$

Proof. Using (2.6.18)

$$\{Q, b(z)\} = \mathrm{Res}_{z_1 \to z}\; j(z_1)\, b(z) \tag{4.3.7}$$

Extracting the mode

$$b_m = \oint \frac{dz}{2\pi i}\, z^{m+1} b(z)$$

we find

$$\begin{aligned} \{Q, b_m\} &= \oint \frac{dz}{2\pi i}\, z^{m+1}\, \mathrm{Res}_{z_1 \to z}\; j(z_1) b(z) \\ &= \oint \frac{dz}{2\pi i}\, z^{m+1}\, \mathrm{Res}_{z_1 \to z} \left[\frac{3}{(z_1 - z)^3} + \frac{j^g(z)}{(z_1 - z)^2} + \frac{(T^m + T^g)(z)}{z_1 - z} \right] \\ &= \oint \frac{dz}{2\pi i}\, z^{m+1} (T^m + T^g)(z) = L_m^m + L_m^g \end{aligned} \tag{4.3.8}$$

□

In terms of ghost field modes:

$$\begin{aligned} Q_{\mathrm{B}} = &\sum_{n=-\infty}^{\infty} (c_n L_{-n}^{\mathrm{m}} + \tilde{c}_n \tilde{L}_{-n}^{\mathrm{m}}) \\ &+ \sum_{m,n=-\infty}^{\infty} \frac{(m-n)}{2} : (c_m c_n b_{-m-n} + \tilde{c}_m \tilde{c}_n \tilde{b}_{-m-n}) : + a^{\mathrm{B}} (c_0 + \tilde{c}_0) \,. \end{aligned} \tag{4.3.9}$$

The ordering constant is $a^{\mathrm{B}} = a^{\mathrm{g}} = -1$, from the anticommutator:

$$\{Q_{\mathrm{B}}, b_0\} = L_0^{\mathrm{m}} + L_0^{\mathrm{g}} \,. \tag{4.3.10}$$

When $c^{\mathrm{m}} \neq 26$, anomalies exist in the gauge symmetry, so we expect glitches in the BRST formalism. The BRST current remains conserved: all terms in the current (4.3.3) are analytic for any value of the central charge. However, it is no longer *nilpotent*:

$$2Q_B^2 = \{Q_{\mathrm{B}}, Q_{\mathrm{B}}\} = 0 \text{ only if } c^{\mathrm{m}} = 26 \,. \tag{4.3.11}$$

The shortest derivation of this result uses the Jacobi identity. More directly, it stems from the OPE:

$$j_{\mathrm{B}}(z) j_{\mathrm{B}}(0) \sim -\frac{c^{\mathrm{m}} - 18}{2z^3} c\partial c(0) - \frac{c^{\mathrm{m}} - 18}{4z^2} c\partial^2 c(0) - \frac{c^{\mathrm{m}} - 26}{12z} c\partial^3 c(0) \,. \tag{4.3.12}$$

This requires some calculation; watch for signs from anticommutators. The simple pole implies $\{Q_{\mathrm{B}}, Q_{\mathrm{B}}\} = 0$ when $c^{\mathrm{m}} = 26$.

To show this, we use (2.6.14):

$$\{Q_B, Q_B\} = \oint \frac{dz_2}{2\pi i} \mathrm{Res}_{z_1\to z_2}\, j_B(z_1) j_B(z_2)$$
$$= \oint \frac{dz_2}{2\pi i} \mathrm{Res}_{z_1\to z_2} \left[\frac{(c^m - 18)\,\partial cc(z_2)}{2(z_1-z_2)^3} + \frac{(c^m-18)\,\partial^2 cc(z_2)}{4(z_1-z_2)^2} \right.$$
$$\left. + \frac{(c^m - 26)\,\partial^3 cc(z_2)}{12(z_1-z_2)} \right] - \frac{(c^m-26)}{12} \oint \frac{dz_2}{2\pi i}\, \partial^3 cc(z_2) \tag{4.3.13}$$

To work this out, first note that

$$\partial^3 c(z) = -\sum_m (m-1)m(m+1)c_m z^{-m-2} \tag{4.3.14}$$

and thus

$$c\,\partial^3 c(z) = -\sum_{m,n=-\infty}^{\infty} \frac{(m^3-m) : c_n c_m :}{z^{n+m+1}} \tag{4.3.15}$$

Therefore

$$\{Q_B, Q_B\} = -\frac{(c^m - 26)}{12} \oint \frac{dz_2}{2\pi i} \sum_{m,n=-\infty}^{\infty} \frac{(m^3-m) : c_n c_m :}{z^{n+m+1}} \tag{4.3.16}$$

$$= -\frac{(c^m-26)}{12} \sum_{m=-\infty}^{\infty} (m^3 - m) \; : c_{-m} c_m : \tag{4.3.17}$$

Note also from the OPE:

$$T(z) j_{\mathrm{B}}(0) \sim \frac{c^{\mathrm{m}} - 26}{2z^4} c(0) + \frac{1}{z^2} j_{\mathrm{B}}(0) + \frac{1}{z}\partial j_{\mathrm{B}}(0) \tag{4.3.18}$$

which implies j_{B} is a tensor only when $c^{\mathrm{m}} = 26$.

Let us point out some important features. The missing equations of motion are precisely the generators of conformal transformations that are zero, the part of the local symmetry group not fixed by gauge selection. This is a general result. When gauge conditions completely fix the symmetry, the missing equations of motion are proven trivial by gauge invariance. They must be zero in the sense of matrix elements (4.1.2), or more generally, in the BRST sense. Denoting the remaining symmetry generators G_I as *constraints*; for the string, these are L_m and $\tilde{L}_m$, forming the algebra:

$$[G_I, G_J] = i g^K{}_{IJ} G_K\,. \tag{4.3.19}$$

Associated with each generator is a pair of ghosts, b_I and c^I, where:

$$\{c^I, b_J\} = \delta^I{}_J\,, \quad \{c^I, c^J\} = \{b_I, b_J\} = 0\,. \tag{4.3.20}$$

The general form of the BRST operator, as exemplified by the string case (4.3.9), is:

$$\begin{aligned} Q_{\mathrm{B}} &= c^I G_I^{\mathrm{m}} - \frac{i}{2} g^K{}_{IJ} c^I c^J b_K \\ &= c^I \left(G_I^{\mathrm{m}} + \frac{1}{2} G_I^{\mathrm{g}} \right), \end{aligned} \tag{4.3.21}$$

where G_I^{m} is the matter part of G_I and:

$$G_I^{\mathrm{g}} = -i g^K{}_{IJ} c^J b_K \tag{4.3.22}$$

is the ghost part. G_I^{m} and G_I^{g} satisfy the same algebra as G_I, (4.3.19). Using commutators (4.3.19) and (4.3.20), one finds:

$$Q_{\mathrm{B}}^2 = \frac{1}{2}\{Q_{\mathrm{B}}, Q_{\mathrm{B}}\} = -\frac{1}{2} g^K{}_{IJ} g^M{}_{KL} c^I c^J c^L b_M = 0\,. \tag{4.3.23}$$

The last equality stems from the Jacobi identity for GGG, which requires $g^K{}_{IJ} g^M{}_{KL}$ to be zero when antisymmetrized in IJL. We ignored the central charge terms; they must be verified manually.

Again, the constraint algebra is the worldsheet symmetry algebra, required to be zero in physical matrix elements. When we generalize the bosonic string in Volume II, the simplest way is to do so directly in terms of the constraint algebra and determine the BRST charge in the form of (4.3.21).

String BRST cohomology

Let us now examine the BRST cohomology of the string at lowest order. Define the inner product by specifying:

$$(\alpha_m^\mu)^\dagger = \alpha_{-m}^\mu\,, \quad (\tilde\alpha_m^\mu)^\dagger = \tilde\alpha_{-m}^\mu\,, \tag{4.3.24a}$$
$$(b_m)^\dagger = b_{-m}\,, \quad (\tilde b_m)^\dagger = \tilde b_{-m}\,, \tag{4.3.24b}$$
$$(c_m)^\dagger = c_{-m}\,, \quad (\tilde c_m)^\dagger = \tilde c_{-m}\,. \tag{4.3.24c}$$

Specifically, Hermiticity of the BRST charge requires ghost fields to be Hermitian. Hermiticity of the ghost zero modes forces the inner product of ground states to be of the form:

$$\text{open string:} \quad \langle 0;k|c_0|0;k'\rangle = (2\pi)^{26}\delta^{26}(k-k')\,, \tag{4.3.25a}$$
$$\text{closed string:} \quad \langle 0;k|\tilde c_0 c_0|0;k'\rangle = i(2\pi)^{26}\delta^{26}(k-k')\,. \tag{4.3.25b}$$

Here $|0;k\rangle$ represents the matter ground state times the ghost ground state $|{\downarrow}\rangle$ with momentum k. Inserting c_0 and $\tilde c_0$ is necessary for a non-zero result. For instance, $\langle 0;k|0;k'\rangle = \langle 0;k|(c_0 b_0 + b_0 c_0)|0;k'\rangle = 0$; the last equality holds because b_0 annihilates both bras and kets. The factor i in the ghost zero-mode inner product is due to Hermiticity. The inner product of the ground states is obtained using commutation relations and the adjoints (4.3.24), as in earlier calculations (4.1.13).

We will focus on the open string; the closed string discussion is entirely analogous but twice as long. We assert (to be proven) that physical states must satisfy the extra condition:

$$b_0|\psi\rangle = 0\,. \tag{4.3.26}$$

This also implies:

$$L_0|\psi\rangle = \{Q_{\mathrm{B}}, b_0\}|\psi\rangle = 0\,, \tag{4.3.27}$$

since both Q_{B} and b_0 annihilate $|\psi\rangle$. Operator L_0 is:

$$L_0 = \alpha'(p^\mu p_\mu + m^2)\,, \tag{4.3.28}$$

where:

$$\alpha' m^2 = \sum_{n=1}^{\infty} n\left(N_{bn} + N_{cn} + \sum_{\mu=0}^{25} N_{\mu n}\right) - 1\,. \tag{4.3.29}$$

Thus, the L_0 condition (4.3.27) determines the string mass. BRST invariance and condition (4.3.26) imply each string state is on-shell. Let $\hat{\mathscr{H}}$ be the space of states satisfying (4.3.26) and (4.3.27). From the anticommutator $\{Q_{\mathrm{B}}, b_0\} = L_0$ and $[Q_{\mathrm{B}}, L_0] = 0$, it follows that Q_{B} acts on

$\hat{\mathscr{H}}$ within itself. The inner product (4.3.25a) is not quite well-defined in $\hat{\mathscr{H}}$: ghost zero modes give 0, and due to momentum constraints on-shell, $\delta^{26}(k-k')$ contains a factor $\delta(0)$. So we use a degenerate inner product $\langle \| \rangle$ in $\hat{\mathscr{H}}$, ignoring X^0 and the ghost zero mode. This is the inner product related to probability interpretation. One can verify Q_B remains Hermitian under this degenerate inner product. Note the on-shell condition determines k^0 from spatial momentum $\mathbf{k}$ (considering the incoming case, $k^0 > 0$), and we used covariant normalization for the states.

Let us now interpret the first level of the string in $D = 26$ flat spacetime. At the lowest order, $N = 0$:

$$|0;\mathbf{k}\rangle\,, \qquad -k^2 = -\frac{1}{\alpha'}\,. \tag{4.3.30}$$

This state is invariant:

$$Q_\mathrm{B}|0;\mathbf{k}\rangle = 0\,, \tag{4.3.31}$$

because every term in Q_B contains a lowering operator or L_0. This also shows no exact states exist at this level, so each invariant state corresponds to a cohomology class. These are exactly the tachyon states. The on-shell condition is the same as found in Section 1.3 for light-cone quantization and in Section 3.6 for open string vertex operators. Obtained heuristically in Chapter 1, it is here determined by the Q_B ordering constant, derived from the requirement of nilpotency.

At the next level, $N = 1$, there are $26 + 2$ states:

$$|\psi_1\rangle = (e\cdot\alpha_{-1} + \beta b_{-1} + \gamma c_{-1})|0;\mathbf{k}\rangle\,, \qquad -k^2 = 0\,, \tag{4.3.32}$$

depending on a 26-vector e_μ and two constants β and γ. The norm is:

$$\begin{aligned}\langle\psi_1\|\psi_1\rangle &= \langle 0;\mathbf{k}\|(e^*\cdot\alpha_1 + \beta^* b_1 + \gamma^* c_1)(e\cdot\alpha_{-1} + \beta b_{-1} + \gamma c_{-1})|0;\mathbf{k}'\rangle \\ &= (e^*\cdot e + \beta^*\gamma + \gamma^*\beta)\langle 0;\mathbf{k}\|0;\mathbf{k}'\rangle\,.\end{aligned} \tag{4.3.33}$$

In an orthogonal basis, there are 26 positive-norm states and 2 negative-norm states. The BRST condition is:

$$\begin{aligned}0 = Q_\mathrm{B}|\psi_1\rangle &= (2\alpha')^{1/2}(c_{-1}k\cdot\alpha_1 + c_1 k\cdot\alpha_{-1})|\psi_1\rangle \\ &= (2\alpha')^{1/2}(k\cdot e c_{-1} + \beta k\cdot\alpha_{-1})|0;\mathbf{k}\rangle\,.\end{aligned} \tag{4.3.34}$$

Terms proportional to c_0 sum to zero due to the on-shell condition and are ignored. An invariant state thus satisfies $k\cdot e = \beta = 0$. This leaves 26 linearly independent states, with 24 positive-norm and 2 zero-norm. The two zero-norm invariant states are created by c_{-1} and $k\cdot\alpha_{-1}$ and are orthogonal to all physical states including themselves. A general $|\chi\rangle$ is the same as (4.3.32) with constants e'_μ, β', γ', so a general BRST exact state at this level is:

$$Q_\mathrm{B}|\chi\rangle = (2\alpha')^{1/2}(k\cdot e' c_{-1} + \beta' k\cdot\alpha_{-1})|0;\mathbf{k}\rangle\,. \tag{4.3.35}$$

Thus, the ghost state $c_{-1}|0;\mathbf{k}\rangle$ is BRST exact, and the polarization is transverse with the equivalence relation $e_\mu \cong e_\mu + (2\alpha')^{1/2}\beta' k_\mu$. This leaves the 24 positive-norm states expected for a massless vector particle, identical to light-cone quantization and OCQ with $A = -1$.

This pattern is general: there are two additional positive-norm and two negative-norm oscillator families compared to light-cone quantization. Physical state conditions eliminate two, leaving two combinations with zero inner product. Those null oscillators are BRST exact and removed by equivalence relations.

Side note: states $b_{-1}|0;\mathbf{k}\rangle$ and $c_{-1}|0;\mathbf{k}\rangle$ are regarded as the two Faddeev-Popov ghosts in the BRST quantization of massless vector fields in field theory. The worldsheet BRST operator acts on these states in the same way the corresponding spacetime BRST operator acts in gauge field theory. Acting on the entire string Hilbert space, the open string BRST operator is some infinite-dimensional generalization of spacetime gauge theory BRST invariance, while the closed string BRST operator is a generalization of spacetime general coordinate BRST invariance in the free limit.

Generalization to closed strings is direct. We focus on the space of states satisfying:

$$b_0|\psi\rangle = \tilde{b}_0|\psi\rangle = 0\,. \tag{4.3.36}$$

This also implies:

$$L_0|\psi\rangle = \tilde{L}_0|\psi\rangle = 0\,. \tag{4.3.37}$$

In closed strings:

$$L_0 = \frac{\alpha'}{4}(p^2+m^2)\,, \qquad \tilde{L}_0 = \frac{\alpha'}{4}(p^2+\tilde{m}^2)\,, \tag{4.3.38}$$

where:

$$\frac{\alpha'}{4}m^2 = \sum_{n=1}^{\infty} n\left(N_{bn}+N_{cn}+\sum_{\mu=0}^{25} N_{\mu n}\right) - 1\,, \tag{4.3.39a}$$

$$\frac{\alpha'}{4}\tilde{m}^2 = \sum_{n=1}^{\infty} n\left(\tilde{N}_{bn}+\tilde{N}_{cn}+\sum_{\mu=0}^{25} \tilde{N}_{\mu n}\right) - 1\,. \tag{4.3.39b}$$

Repeating the exercise, at $m^2 = -4/\alpha'$ we find the tachyon, and at $m^2 = 0$, the graviton, dilaton, and antisymmetric tensor 24×24 states.

4.4 The no-ghost theorem

In this section, we prove that the BRST cohomology of the string has a positive-definite inner product and is isomorphic to the light-cone and OCQ spectra, regarding the BRST cohomology as the physical Hilbert space. We also verify that in our study of string amplitudes, the amplitudes are well-defined on the cohomology (i.e., equivalent states have identical amplitudes) and the S-matrix is unitary within the physical state space.

We operate within the general framework described at the end of Chapter 3: the worldsheet theory consists of d free fields X^μ (including $\mu = 0$), a compact unitary CFT K with central charge $26-d$, and ghosts. Virasoro generators are the sum:

$$L_m = L_m^X + L_m^K + L_m^{\mathrm{g}}\,. \tag{4.4.1}$$

Compact means L_0^K has a discrete spectrum. For example, if K corresponds to strings on a compact manifold, the term $\alpha' p^2$ in L_0 is replaced by $-\alpha'\nabla^2$, which has a discrete spectrum on compact space. In both open and closed strings, general states are labeled:

$$|N, I; k\rangle\,, \qquad |N, \tilde{N}, I; k\rangle\,, \tag{4.4.2}$$

where N (and $\tilde{N}$) refer to d-dimensional and ghost oscillators, k is the d-dimensional momentum, and I labels states of the compact CFT with given boundary conditions. Impose b_0 conditions (4.3.26) or (4.3.36), respectively implying the on-shell conditions for the open string:

$$-\sum_{\mu=0}^{d-1} k_\mu k^\mu = m^2\,, \tag{4.4.3a}$$

$$\alpha' m^2 = \sum_{n=1}^{\infty} n\left(N_{bn}+N_{cn}+\sum_{\mu=0}^{d-1} N_{\mu n}\right) + L_0^K - 1\,, \tag{4.4.3b}$$

and the closed string:

$$-\sum_{\mu=0}^{d-1} k_\mu k^\mu = m^2 = \tilde{m}^2 \, , \tag{4.4.4a}$$

$$\frac{\alpha'}{4} m^2 = \sum_{n=1}^{\infty} n \left(N_{bn} + N_{cn} + \sum_{\mu=0}^{d-1} N_{\mu n} \right) + L_0^K - 1 \, , \tag{4.4.4b}$$

$$\frac{\alpha'}{4} \tilde{m}^2 = \sum_{n=1}^{\infty} n \left(\tilde{N}_{bn} + \tilde{N}_{cn} + \sum_{\mu=0}^{d-1} \tilde{N}_{\mu n} \right) + \tilde{L}_0^K - 1 \, . \tag{4.4.4c}$$

The contributions of $d \leq \mu \leq 25$ oscillators are replaced by the eigenvalues of L_0^K or $\tilde{L}_0^K$. The only information used regarding the compact CFT is its conformal invariance at the appropriate central charge (so a nilpotent BRST operator exists) and its positive-definite inner product. The basis I can be taken as orthogonal, so the degenerate inner product is:

$$\langle 0, I; \mathbf{k} \| 0, I'; \mathbf{k}' \rangle = \langle 0, 0, I; \mathbf{k} \| 0, 0, I'; \mathbf{k}' \rangle = 2k^0 (2\pi)^{d-1} \delta^{d-1}(\mathbf{k} - \mathbf{k}') \delta_{I,I'} \, . \tag{4.4.5}$$

What do we expect for the physical Hilbert space? Define the transverse Hilbert space $\mathscr{H}^\perp$ as the states in $\hat{\mathscr{H}}$ with no longitudinal (X^0, X^1, b, c) excitations. Since these oscillators are the source of indefinite inner products, $\mathscr{H}^\perp$ has a positive-definite inner product. Light-cone gauge fixing directly eliminates longitudinal oscillators—the light-cone Hilbert space is isomorphic to $\mathscr{H}^\perp$, as seen in the flat spacetime case in Chapter 1. We will prove that in general, BRST cohomology is isomorphic to $\mathscr{H}^\perp$, meaning it has the same number of states at each mass level and a positive inner product: the no-ghost theorem.

Proof

The proof has two parts. First, discover the cohomology of the simplified BRST operator Q_1, which is a quadratic form in oscillators; second, prove that the total Q_B cohomology is identical to the Q_1 cohomology. Define light-cone oscillators:

$$\alpha_m^\pm = 2^{-1/2}(\alpha_m^0 \pm \alpha_m^1) \, , \tag{4.4.6}$$

which satisfy:

$$[\alpha_m^+, \alpha_n^-] = -m\delta_{m,-n} \, , \qquad [\alpha_m^+, \alpha_n^+] = [\alpha_m^-, \alpha_n^-] = 0 \, . \tag{4.4.7}$$

We will extensively use the quantum number:

$$N^{\mathrm{lc}} = \sum_{\substack{m=-\infty \\ m\neq 0}}^{\infty} \frac{1}{m} :\alpha_{-m}^+ \alpha_m^-: \; . \tag{4.4.8}$$

N^{lc} is the number of $-$ excitations minus the number of $+$ excitations; since the center-of-mass part is omitted, it is not a Lorentz generator. We select a coordinate system where the momentum component k^+ is non-zero. Now decompose the BRST generator using N^{lc}:

$$Q_\mathrm{B} = Q_1 + Q_0 + Q_{-1} \, , \tag{4.4.9}$$

where Q_j acting on a state changes N^{lc} by j units:

$$[N^{\mathrm{lc}}, Q_j] = jQ_j \, . \tag{4.4.10}$$

Additionally, each Q_j acting on a state increases the ghost number N^g by 1:

$$[N^\mathrm{g}, Q_j] = Q_j \, . \tag{4.4.11}$$

Expanding $Q_{\mathrm{B}}^2 = 0$ yields:

$$\left(Q_1^2\right) + \left(\{Q_1, Q_0\}\right) + \left(\{Q_1, Q_{-1}\} + Q_0^2\right) + \left(\{Q_0, Q_{-1}\}\right) + \left(Q_{-1}^2\right) = 0\,. \tag{4.4.12}$$

Each group in parentheses has a different N^{lc}, so each must be zero separately. Specifically, Q_1 itself is nilpotent and has a cohomology. Explicitly:

$$Q_1 = -(2\alpha')^{1/2} k^+ \sum_{\substack{m=-\infty \\ m\neq 0}}^{\infty} \alpha_{-m}^- c_m\,. \tag{4.4.13}$$

For $m < 0$, it annihilates a $+$ mode and creates a c; for $m > 0$, it creates a $-$ mode and annihilates a b. One can find the cohomology by investigating the action of Q_1 on an occupation number basis. This is left as an exercise; here we use a standard strategy useful for the generalization of Q_{B}. Define:

$$R = \frac{1}{(2\alpha')^{1/2} k^+} \sum_{\substack{m=-\infty \\ m\neq 0}}^{\infty} \alpha_{-m}^+ b_m \tag{4.4.14}$$

and:

$$\begin{aligned} S \equiv \{Q_1, R\} &= \sum_{m=1}^{\infty} \left(m b_{-m} c_m + m c_{-m} b_m - \alpha_{-m}^+ \alpha_m^- - \alpha_{-m}^- \alpha_m^+\right) \\ &= \sum_{m=1}^{\infty} m(N_{bm} + N_{cm} + N_m^+ + N_m^-)\,. \end{aligned} \tag{4.4.15}$$

Note that both Q_1 and R annihilate the ground state, which determines the normal-ordering constant. Note also that Q_1 commutes with S. We can then compute the cohomology in each eigenspace of S, and the total cohomology is the union of the results. If $|\psi\rangle$ is a Q_1 invariant satisfying $S|\psi\rangle = s|\psi\rangle$, then for non-zero s:

$$|\psi\rangle = \frac{1}{s}\{Q_1, R\}|\psi\rangle = \frac{1}{s} Q_1 R|\psi\rangle\,, \tag{4.4.16}$$

so $|\psi\rangle$ is actually Q_1-exact. Thus, Q_1 cohomology is non-zero only for $s = 0$. By the definition of S (4.4.15), states with $s = 0$ have no longitudinal excitations—the $s = 0$ space is exactly $\mathscr{H}^\perp$. Operator Q_1 annihilates all states in $\mathscr{H}^\perp$, so they are all Q_1-closed, and there are no Q_1 exact states in this space. Thus, the cohomology is $\mathscr{H}^\perp$ itself. We have thus proven the no-ghost theorem for Q_1. Next, we prove it for Q_{B}.

The proof consists of two steps: first, prove the cohomology comes only from $s = 0$ states (the kernel of S); second, prove $s = 0$ states are Q_1-invariant. It is useful to prove the second step abstractly, utilizing the property that all $s = 0$ states have the same ghost number, which in this case is $-\frac{1}{2}$. Suppose $S|\psi\rangle = 0$. Since S and Q_1 commute:

$$0 = Q_1 S|\psi\rangle = S Q_1|\psi\rangle\,. \tag{4.4.17}$$

The ghost number of $|\psi\rangle$ is $-\frac{1}{2}$, so the ghost number of $Q_1|\psi\rangle$ is $+\frac{1}{2}$. Since S is invertible for this ghost number, we must have $Q_1|\psi\rangle = 0$, which is what we aimed to prove.

Remaining is to prove the Q_{B} cohomology is identical to the Q_1 cohomology. The idea is to replace S with the operator:

$$S + U \equiv \{Q_{\mathrm{B}}, R\}\,. \tag{4.4.18}$$

Now, $U = \{Q_0 + Q_{-1}, R\}$ decreases N^{lc} by one or two units. In terms of N^{lc}, S is diagonal and U is lower triangular. By the general properties of lower triangular matrices, the kernel of $S + U$

cannot be larger than the kernel of its diagonal part S. In fact, they are isomorphic: if $|\psi_0\rangle$ is annihilated, then:

$$|\psi\rangle = (1 - S^{-1}U + S^{-1}US^{-1}U - \cdots)|\psi_0\rangle \tag{4.4.19}$$

is annihilated by $S+U$. The factor S^{-1} is reasonable as it always acts on $N^{\rm lc} < 0$ states, where S is invertible. For the same reason, $S+U$ is invertible except when the ghost number is $-\frac{1}{2}$. Equations (4.4.16) and (4.4.17) now apply to $Q_{\rm B}$, with $S+U$ replacing S. They imply the $Q_{\rm B}$ cohomology is isomorphic to the kernel of $S+U$, which in turn is isomorphic to the kernel of S, and thus to the Q_1 cohomology, as we set out to prove.

We must still verify the inner product is positive-definite. All terms after the first on the right side of (4.4.19) have a strictly negative $N^{\rm lc}$. By commutation relations, this inner product is non-zero only between states whose $N^{\rm lc}$ values sum to zero. Then for two states in the kernel of $S+U$ (4.4.19):

$$\langle\psi\|\psi'\rangle = \langle\psi_0\|\psi_0'\rangle\,. \tag{4.4.20}$$

The positive-definiteness of the inner product in the S kernel thus implies positive-definiteness in the $S+U$ kernel. The proof for closed strings is identical, with the addition of tilde operators for $Q_{\rm B}, N^{\rm lc}, R$, etc.

BRST-OCQ equivalence

We now prove the equivalence:

$$\mathscr{H}_{\rm OCQ} = \mathscr{H}_{\rm BRST} = \mathscr{H}_{\text{light-cone}}\,. \tag{4.4.21}$$

For a state $|\psi\rangle$ in the matter Hilbert space, adding the ghost theory gives the state:

$$|\psi, \downarrow\rangle\,. \tag{4.4.22}$$

The ghost vacuum $|\downarrow\rangle$ remains annihilated by all ghost lowering operators, i.e., b_n for $n \geq 0$ and c_n for $n > 0$. Acting with $Q_{\rm B}$:

$$Q_{\rm B}|\psi,\downarrow\rangle = \sum_{n=0}^{\infty} c_{-n}(L_n^{\rm m} - \delta_{n,0})|\psi,\downarrow\rangle = 0\,. \tag{4.4.23}$$

All terms in $Q_{\rm B}$ containing ghost lowering operators are discarded, and the constant in the $n=0$ term is known from the $Q_{\rm B}$ mode expansion (4.3.9). Thus, OCQ physical states map to BRST-closed states. The ordering constant $A=-1$ stems from the L_0 eigenvalue of the ghost vacuum.

To establish equivalence (4.4.21), we must prove more. First, we must show a well-defined mapping of equivalence classes: if ψ and ψ' are equivalent OCQ physical states, they map to the same BRST class:

$$|\psi,\downarrow\rangle - |\psi',\downarrow\rangle \tag{4.4.24}$$

must be BRST-exact. By (4.4.23), this state is BRST-closed. Furthermore, since $|\psi\rangle' - |\psi\rangle$ is an OCQ null state, the norm of state (4.4.24) is zero. From the BRST no-ghost theorem, the inner product of the cohomology is positive-definite, so a closed state with zero norm is BRST-exact, which is what we aimed to prove.

To conclude that we have an isomorphism, we must show the mapping is one-to-one and onto. One-to-one means that OCQ physical states ψ and ψ' mapping to the same BRST class must be in the same OCQ class—if:

$$|\psi,\downarrow\rangle - |\psi',\downarrow\rangle = Q_{\rm B}|\chi\rangle\,, \tag{4.4.25}$$

then $|\psi\rangle - |\psi\rangle'$ must be a null state. To see this, expand:

$$|\chi\rangle = \sum_{n=1}^{\infty} b_{-n}|\chi_n, \downarrow\rangle + \cdots ; \tag{4.4.26}$$

$|\chi\rangle$ has ghost number $-\frac{3}{2}$, so the ellipsis represents at least one c excitation and two b excitations. Substitute this into (4.4.25) and keep only terms on both sides with the ghost ground state. This gives:

$$\begin{aligned} |\psi, \downarrow\rangle - |\psi', \downarrow\rangle &= \sum_{m,n=1}^{\infty} c_m L^{\mathrm{m}}_{-m} b_{-n}|\chi_n, \downarrow\rangle \\ &= \sum_{n=1}^{\infty} L^{\mathrm{m}}_{-n}|\chi_n, \downarrow\rangle \, . \end{aligned} \tag{4.4.27}$$

Terms with ghost excitations must be zero separately and have been omitted. Thus, $|\psi\rangle - |\psi'\rangle = \sum_{n=1}^{\infty} L^{\mathrm{m}}_{-n}|\chi_n\rangle$ is an OCQ null state, making the mapping one-to-one.

Finally, we must show the mapping is onto: every Q_{B} equivalence class contains at least one state of the form (4.4.22). In fact, the specific representation (4.4.19), states annihilated by $S + U$, is of this form. To see this, consider the quantum number $N' = 2N^- + N_b + N_c$, including the total number of $-$, b and c excitations. Operator R has $N' = -1$: terms with $m > 0$ in R decrease N_c by one unit, terms with $m < 0$ decrease N^- by one unit and increase N_b by one unit. Checking $Q_0 + Q_{-1}$, one finds various terms with $N' = 1$ but nothing larger. So $U = \{R, Q_0 + Q_{-1}\}$ cannot increase N'. Examining state (4.4.19) and noting that S and $|\psi_0\rangle$ have $N' = 0$, all terms on the right have $N' \leq 0$. By definition, N' is non-negative, so $N'|\psi\rangle = 0$ must hold. This implies no excitations of $-$, b, or c, so the state is of form (4.4.22). Thus, equivalence (4.4.21) is proven.

The full power of the BRST method is needed to understand the general structure of string amplitudes. However, for most practical purposes, handling states of the form $|\psi, \downarrow\rangle$ is a major simplification, so knowing every BRST class contains at least one state without ghost mode excitations is useful; we call these OCQ-type states.

OCQ physical state condition (4.1.5) requires physical states to be highest weight states with $L_0 = 1$. The corresponding vertex operators are tensor fields $\mathscr{V}^{\mathrm{m}}$ of weight 1 constructed from matter fields. Introducing ghost state $|\downarrow\rangle$, the complete vertex operator is $c\mathscr{V}^{\mathrm{m}}$. For closed strings, the complete vertex operator is $\tilde{c}c\mathscr{V}^{\mathrm{m}}$, where $\mathscr{V}^{\mathrm{m}}$ is a $(1,1)$ type tensor. The matter part of the vertex operator is identical to that found in the Polyakov system in Section 3.6, and the ghost part will be explained simply in the next chapter.

Equation (4.4.19) defines an OCQ physical state in each cohomology class. For the specific case of flat spacetime, this state can be established more clearly using Del Giudice-Di Vecchia-Fubini (DDF) operators. To explain these, further development of vertex operator techniques is required, so this is deferred to Chapter 8.

Chapter 5

The String S-matrix

In Chapter 3, we expressed the S-matrix as a path integral over two-dimensional compact surfaces with vertex operators. In this chapter, we will reduce this path integral to a gauge-fixed form.

5.1 The Circle and the Torus

We identify a gauge slice, as shown in figure 3.7, as a choice of one configuration from each (diff $\times$ Weyl) equivalence class. Locally, we achieve this by fixing the metric. However, globally, there exists a slight mismatch between the space of metrics and the worldsheet gauge group.

Once again, the point particle serves as an excellent illustration. We wish to calculate the Euclidean path integral:

$$\int [dedX] \exp\left[-\frac{1}{2}\int d\tau \left(e^{-1}\dot{X}^\mu \dot{X}_\mu + em^2\right)\right] \tag{5.1.1}$$

Consider paths that form closed loops in spacetime, so the topology is a circle. The parameter τ takes values between 0 and 1, where the endpoints are equivalent. That is, $X^\mu(\tau)$ and $e(\tau)$ are periodic on $0 \leq \tau \leq 1$. The einbein $e(\tau)$ has one component, and there exists a local symmetry that completely determines the einbein.

The einbein transforms as $e'd\tau' = ed\tau$, so the gauge choice yields a differential equation for $\tau'(\tau)$:

$$\frac{\partial \tau'}{\partial \tau} = e(\tau) \tag{5.1.2}$$

Integrating this with the boundary condition $\tau'(0) = 0$ determines:

$$\tau'(\tau) = \int_0^\tau d\tau'' e\left(\tau''\right) \tag{5.1.3}$$

The problem is that, in general, $\tau'(1) \neq 1$, so periodicity is not preserved. In fact,

$$\tau'(1) = \int_0^1 d\tau e(\tau) = l \tag{5.1.4}$$

is the invariant length of the circle. Therefore, we cannot set $e' = 1$ while keeping the coordinate range invariant. We can either keep the coordinate range fixed and let e' be a constant value $e' = l$, or let $e' = 1$ and allow the coordinate range to vary:

$$e' = l, \quad 0 \leq \tau \leq 1 \tag{5.1.5a}$$

or

$$e' = 1, \quad 0 \leq \tau \leq l \tag{5.1.5b}$$

In either case, after fixing the gauge invariance, we are left with a universal integral over l.

In other words, not all einbeins on a circle are equivalent. There exists a single-parameter family of inequivalent einbeins, parameterized by l. The two descriptions in (5.1.5) both have analogs in string theory. Actually, we will use a description similar to (5.1.5a) to define the path integral, where the fields are functions on a fixed coordinate range, and then transform to a description like (5.1.5b), where the metric is fixed and the moduli are encoded within the coordinate range.

There is a second difficulty related to gauge fixing. The condition $e =$ constant is preserved by rigid translations:

$$\tau \rightarrow \tau + v \bmod 1 \tag{5.1.6}$$

That is, on a circle, we cannot say where the origin is. Consequently, fixing the metric leaves a small portion of unfixed local symmetry, resulting in a mismatch in both directions. The metric is not equivalent to another gauge, and the gauge transformation is not fixed by the choice of metric.

Moving to strings, we take the torus as an example, where the same difficulties arise. Starting from the coordinate region:

$$0 \leq \sigma^1 \leq 2\pi, \quad 0 \leq \sigma^2 \leq 2\pi \tag{5.1.7}$$

where $X^\mu\left(\sigma^1, \sigma^2\right)$ and $g_{ab}\left(\sigma_1, \sigma_2\right)$ are periodic in both directions. Equivalently, we can view this as point equivalence on the σ plane:

$$\left(\sigma^1, \sigma^2\right) \cong \left(\sigma^1, \sigma^2\right) + 2\pi(m, n) \tag{5.1.8}$$

where m, n are integers.

To what extent is the field space diff $\times$ Weyl redundant? The theorem is that it is impossible to transform a general metric into a unit form through a diff $\times$ Weyl transformation that keeps the periodicity (5.1.7) invariant, but it can be transformed into:

$$ds^2 = \left|d\sigma^1 + \tau d\sigma^2\right|^2 \tag{5.1.9}$$

where τ is a complex constant. When $\tau = i$, this is the unit metric δ_{ab}.

To see this, we repeat the steps used in the local discussion of section 3.3. First, we can make the metric flat through a Weyl transformation satisfying $2\nabla^2\omega = R$. By expanding the eigenmodes of ∇^2, it can be seen that under periodic boundary conditions, there is a unique solution for ω up to an additive constant. What is important here is $\int d^2\sigma g^{1/2} R$, which is 4π times the Euler number, and for a torus, this is zero. Afterwards, a transformation to new coordinates $\tilde{\sigma}^a$ turns the metric into unit form. However, just as in the point particle case, there is no guarantee that this reflects the original periodicity. Instead, we may now have:

$$\tilde{\sigma}^a \cong \tilde{\sigma}^a + 2\pi\left(m u^a + n v^a\right) \tag{5.1.10}$$

where u^a, v^a are general translations. By rotating and scaling the coordinate system, accompanied by an ω shift that keeps the metric normalized, we can always set $u = (1, 0)$. This leaves two parameters, the components of v.

Defining $w = \tilde{\sigma}^1 + i\tilde{\sigma}^2$, the metric is $dw d\bar{w}$ and the periodicity is:

$$w \cong w + 2\pi(m + n\tau) \tag{5.1.11}$$

where $\tau = v^1 + iv^2$. The torus is a parallelogram with periodic boundary conditions in the w plane, as shown in figure (5.1.1).

Alternatively, using σ^a, $w = \sigma^1 + \tau\sigma^2$ is defined. The original periodicity is preserved, but the metric takes the more general form (5.1.9). The integration over the metric reduces to two integrals over the real and imaginary parts of τ. The metric (5.1.9) is invariant under complex

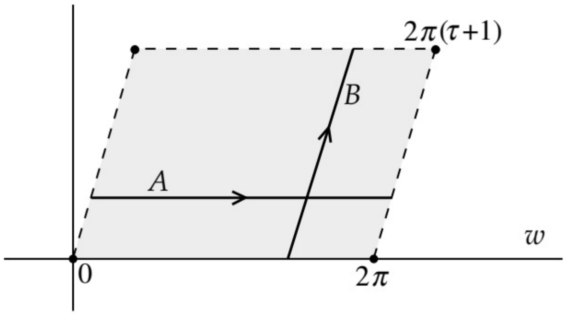


Figure 5.1.1: Torus with modulus τ, in gauge $ds^2 = dwd\bar{w}$. The upper and lower edges are identified, as are the right- and left-hand edges. Closed curves A and B are marked for later reference.

conjugation and degenerates for real τ, so we can focus our attention on $\mathrm{Im}\,\tau > 0$. Just as in the circular case (5.1.5), we can put these parameters into the metric (5.1.9) or the periodicity (5.1.11). The parameter τ is called the Teichmüller parameter or, more generally, the modulus.

Unlike the point particle, there is no simple invariant expression for τ analogous to the invariant length of a circle. There is some additional redundancy that has no analog in the point particle case. The value $\tau + 1$ generates the same equivalence set (5.1.11) as τ, replacing $(m, n) \to (m - n, n)$. The same is true for $-1/\tau$. Defining $w' = \tau w$ and replacing $(m, n) \to (n, -m)$, repeatedly using these two transformations:

$$T: \quad \tau' = \tau + 1, \quad S: \tau' = -1/\tau \tag{5.1.12}$$

generates:

$$\tau' = \frac{a\tau + b}{c\tau + d} \tag{5.1.13}$$

where a, b, c, d are integers satisfying $ad - bc = 1$.

We can also consider it as follows: the transformation

$$\begin{bmatrix} \sigma^1 \\ \sigma^2 \end{bmatrix} = \begin{bmatrix} d & b \\ c & a \end{bmatrix} \begin{bmatrix} \sigma'^1 \\ \sigma'^2 \end{bmatrix} \tag{5.1.14}$$

turns the σ metric (5.1.9) into the same form of metric in σ' but with modulus τ'. This is a diffeomorphism mapping of the torus. Since $ad - bc = 1$, it is a one-to-one mapping and preserves the periodicity (5.1.8). However, it cannot be obtained from a continuous infinitesimal transformation starting from the identity—it is a so-called large coordinate transformation. Curve A in coordinates σ maps to a curve in σ' stretched a times along the A' direction and c times along the B' direction.

These large coordinate transformations form the group $SL(2, \mathbf{Z})$. The group on the τ plane is $SL(2, \mathbf{Z})/\mathbf{Z}_2 = PSL(2, \mathbf{Z})$, because if the signs of a, b, c, d are all flipped, τ' remains invariant. The group of transformations (5.1.14) is called the modular group. Using the modular transformation (5.1.13), it can be proved that every τ is equivalent to a point in the region F_0, as shown in figure (5.1.2).

$$-\frac{1}{2} \leq \mathrm{Re}\,\tau \leq \frac{1}{2}, \quad |\tau| \geq 1 \tag{5.1.15}$$

Except for the boundaries identified in the figure, this is called the fundamental region of the upper half-plane modulo $PSL(2, \mathbf{Z})$. Due to the equivalence, F_0 can be considered rolled up, opening only when $\mathrm{Im}\,\tau \to \infty$. The fundamental region F_0 is a representation of the moduli space of (diff $\times$ Weyl) inequivalent metrics.

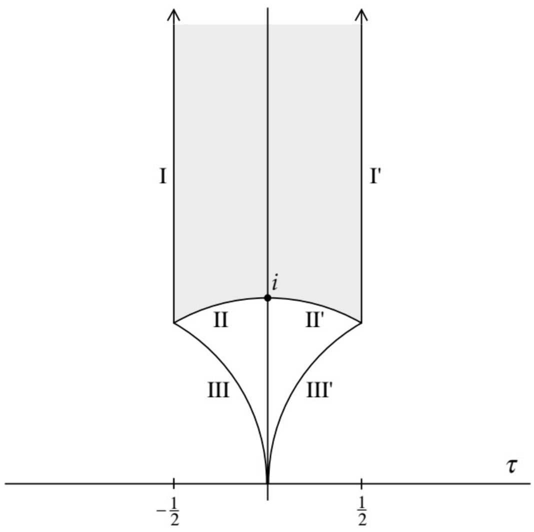


Figure 5.1.2: The standard fundamental region F_0 for the moduli space of the torus, shaded. The lines I and I' are identified, as are the arcs II and II'. A different fundamental region, mapped into F_0 by the modular transformation S, is bounded by II, II', III, and III'.

There are further difficulties, as with the point particle. Requiring the metric to take the form (5.1.9) where τ is in a given fundamental region does not fix all diff $\times$ Weyl invariance. The metric and periodicity are invariant under rigid translations:

$$\sigma^a \to \sigma^a + v^a \tag{5.1.16}$$

So this two-parameter diff $\times$ Weyl subgroup is not fixed. Additionally, the discrete transformation $\sigma^a \to -\sigma^a$ leaves the metric invariant, and in the non-oriented case, $\tau \to -\bar{\tau}$ via $\left(\sigma^1, \sigma^2\right) \to \left(-\sigma^1, \sigma^2\right)$ also leaves the metric invariant. Thus, there is again a mismatch between metric degrees of freedom and local symmetries: parameters in the metric cannot be removed by symmetry, and symmetry is not fixed by the choice of metric. Unfixed symmetries are called the Conformal Killing Group (CKG).

When there are enough vertex operators in the amplitude, we can completely fix the gauge invariance by fixing the positions of certain vertex operators. On a torus with n vertex operators, the invariance (5.1.16) can be used to fix the position of one vertex operator, leaving the integral over τ and $n-1$ other positions. The $\mathbf{Z}_2$ from $\sigma^a \to -\sigma^a$ can be used to fix a second vertex operator in the middle of the torus. In practice, for finite redundancy counts, it is often more convenient to leave some symmetries unfixed and divide by a suitable factor. Additionally, if there are too few vertex operators and some Conformal Killing invariance remains unfixed, then we must explicitly account for the redundancy count.

Note to the Reader

In the remainder of this chapter, we will discuss moduli more generally. The primary goal is the gauge-fixed expression for the string S-matrix (5.3.9). In fact, most interesting string physics can be seen in tree-level and one-loop amplitudes. For this reason, the measure can be obtained through shortcuts. For tree amplitudes, there are no metric moduli but the positions of certain vertex operators must be fixed. The measure can be reasoned out as explained below (6.4.4). For one-loop amplitudes, the measure can be understood in a relatively intuitive way by analogy with the discussion of point particles below (7.3.9).

5.2 Moduli and Riemann Surfaces

Now, we repeat the above discussion in a more general and abstract way. We integrate over all metrics on some given topology r. Call this metric space $\mathscr{G}_r$. For closed oriented surfaces, we can label r by the number of handles g, also called the genus. After accounting for diff $\times$ Weyl redundancy, we are left with the moduli space:

$$\mathscr{M}_r = \frac{\mathscr{G}_r}{(\text{ diff } \times \text{Weyl})_r} \tag{5.2.1}$$

Just as in the torus case, this space is parameterized by a finite number of moduli. For the torus, $\mathscr{M}_1$ is the upper half-plane modulo $PSL(2,\mathbf{Z})$, or equivalently any fundamental region such as F_0. There may also exist a diff $\times$ Weyl subgroup, the CKG, that leaves the metric invariant.

When vertex operators are present in the path integral, it is useful to treat their positions on an equal footing with the moduli in the metric. When we need to make a distinction, we will refer to metric moduli. One way to handle the CKG, which is applicable if vertex operators exist in the path integral, is to further specify the gauge by fixing the positions of certain vertex operators. The Polyakov path integral includes an integration over $\mathscr{G}_r$ and an integration of n vertex operators over the worldsheet $\mathscr{M}$. The moduli space with vertex operators on topology r is:

$$\mathscr{M}_{r,n} = \frac{\mathscr{G}_r \times M^n}{(\text{ diff } \times \text{Weyl})_r} \tag{5.2.2}$$

As we saw for the torus, diff_r is generally not connected. Selecting the connected component diff_{r0} containing the identity, the quotient

$$\frac{\text{diff}_r}{\text{diff}_{r0}} \tag{5.2.3}$$

is the modular group.

It is interesting to study the diff $\times$ Weyl redundancy of the metric at the infinitesimal level. That is, searching for metric variations that are not equivalent to a diff $\times$ Weyl transformation and are thus equivalent to a change in moduli. We also want to find certain diff $\times$ Weyl transformations that do not change the metric; these are the infinitesimal elements of the CKG, called Conformal Killing Vectors (CKVs).

An infinitesimal diff $\times$ Weyl transformation changes the metric by:

$$\delta g_{ab} = -2\left(P_1 \delta\sigma\right)_{ab} + (2\delta\omega - \nabla\cdot\delta\sigma) g_{ab} \tag{5.2.4}$$

where P_1 is the traceless symmetric linear combination of derivatives (3.3.17). The metric variation $\delta' g_{ab}$ corresponding to the moduli is orthogonal to all variations (5.2.4):

$$\begin{aligned} 0 &= \int d^2\sigma g^{1/2} \delta' g_{ab} \left[-2\left(P_1\delta\sigma\right)^{ab} + (2\delta\omega - \nabla\cdot\delta\sigma) g^{ab}\right] \\ &= \int d^2\sigma g^{1/2} \left[-2\left(P_1^T \delta' g\right)_a \delta\sigma^a + \delta' g_{ab} g^{ab} (2\delta\omega - \nabla\cdot\delta\sigma)\right] \end{aligned} \tag{5.2.5}$$

where the adjoint is $\left(P_1^T u\right)_a = -\nabla^b u_{ab}$.

Note:

$$\begin{aligned} (P_1,\delta\sigma)^{ab} &= \frac{1}{2}\left(\nabla^a\delta\sigma^b + \nabla^b\delta\sigma^a - g^{ab}\nabla\cdot\delta\sigma\right) \\ \delta' g_{ab}(P_1,\delta\sigma)^{ab} &= \frac{1}{2}\left(\delta' g_{ab}\nabla^a\delta\sigma^b + \delta' g_{ab}\nabla^b\delta\sigma^a - \delta' g_{ab} g^{ab}\nabla\cdot\delta\sigma\right) \end{aligned}$$

In order to make the overlap (5.2.5) zero for general $\delta\omega$ and $\delta\sigma$, we require

$$g^{ab}\delta' g_{ab} = 0 \tag{5.2.6a}$$

$$\left(P_1^T \delta' g\right)_a = 0 \tag{5.2.6b}$$

The first condition requires $\delta' g_{ab}$ to be traceless, and it is a traceless symmetric tensor in the kernel of P_1^T. For each solution to these equations, there exists a modulus.

CKVs are those $\delta\sigma$ such that $\delta g_{ab} = 0$ in (5.2.4). The trace of this equation uniquely determines $\delta\omega$, leaving the Conformal Killing Equation:

$$\left(P_1 \delta\sigma\right)_{ab} = 0 \tag{5.2.7}$$

Equations (5.2.6) and (5.2.7) become simple for variations around the conformal gauge:

$$\partial_{\bar z}\delta' g_{zz} = \partial_z \delta' g_{\bar z\bar z} = 0 \tag{5.2.8a}$$

$$\partial_{\bar z}\delta z = \partial_z \delta \bar z = 0 \tag{5.2.8b}$$

Thus, the variation of the moduli corresponds to holomorphic quadratic differentials, and CKVs correspond to holomorphic vector fields. On a torus, the only holomorphic bi-periodic functions are constants, so there are two real moduli and two real CKVs, consistent with the previous section's discussion.

The metric moduli correspond to the kernel of P_1^T (5.2.6), while CKVs correspond to the kernel of P_1 (5.2.7). The Riemann-Roch theorem relates the number of metric moduli $\mu = \dim\ker P_1^T$ and the number of CKVs $\kappa = \dim\ker P_1$ to the Euler number χ:

$$\mu - \kappa = -3\chi \tag{5.2.9}$$

We will derive it later in this chapter. For closed oriented surfaces, $-3\chi = 6g - 6$. This count is for real moduli, splitting complex moduli like τ into real and imaginary parts.

Furthermore, κ is zero for $\chi < 0$, and μ is zero for $\chi > 0$. To prove this, we cite the following property: one can always find a metric via Weyl transformation such that the scalar curvature R is constant. In this case, R has the same sign as χ. Now, $P_1^T P_1 = -\frac{1}{2}\nabla^2 - \frac{1}{4}R$, so

$$\begin{aligned}\int d^2\sigma g^{1/2}\left(P_1\delta\sigma\right)_{ab}\left(P_1\delta\sigma\right)^{ab} &= \int d^2\sigma g^{1/2}\delta\sigma_a\left(P_1^T P_1 \delta\sigma\right)^a \\ &= \int d^2\sigma g^{1/2}\left(\frac{1}{2}\nabla_a\delta\sigma_b\nabla^a\delta\sigma^b - \frac{R}{4}\delta\sigma_a\delta\sigma^a\right)\end{aligned} \tag{5.2.10}$$

For negative χ, the right side is strictly positive definite, so $P_1\delta\sigma$ cannot be zero. A similar argument shows that for positive χ, $P_1^T\delta' g$ cannot be zero. Combining these results, we have:

$$\chi > 0 : \kappa = 3\chi, \quad \mu = 0 \tag{5.2.11a}$$

$$\chi < 0 : \kappa = 0, \quad \mu = -3\chi \tag{5.2.11b}$$

Riemann Surfaces

A common way to describe the moduli space is to choose a metric from each equivalence class, which forms a family of metrics $\hat g_{ab}(t;\sigma)$ depending on moduli t^k. For the torus, (5.1.9) gives such a slice. Using the other description (5.1.11) is often more convenient, where $dw d\bar w$ is fixed and the moduli are encoded in the coordinate range. We can formalize this concept as follows.

Recall how a differentiable manifold is defined. A manifold is covered by a set of overlapping charts, where σ_m^a are the coordinates in the m-th coordinate chart. When charts m and n overlap, the coordinates in the two charts are related by:

$$\sigma_m^a = f_{mn}^a\left(\sigma_n\right) \tag{5.2.12}$$

where transition functions are required to be differentiable. For a Riemannian manifold, each chart is also given a metric $g_{m,ab}(\sigma_m)$, and the values in the overlap region are related by the usual tensor laws.

For a complex manifold, each chart has complex coordinates z_m^a. Transition functions are required to be holomorphic:

$$z_m^a = f_{mn}^a(z_n) \tag{5.2.13}$$

Since holomorphicity does not depend on the coordinates z_m^a used. Holomorphic functions on the manifold can now be defined. Just as two differentiable manifolds are equivalent if there exists a one-to-one differentiable mapping between them, two complex manifolds are equivalent if there exists a one-to-one holomorphic mapping between them.

In the two-dimensional case (one complex coordinate), the complex manifold is called a Riemann surface. In this case, there is a one-to-one correspondence:

$$\text{Riemann surfaces} \leftrightarrow \text{Riemannian manifolds mod Weyl} \tag{5.2.14}$$

We place 'mod Weyl' on the right because 'mod diff' is already implied in the definition of a Riemannian manifold. To see this isomorphism, start from the Riemannian manifold. From our discussion of conformal gauge, we know we can find z_m in each chart such that:

$$ds^2 \propto dz_m d\bar{z}_m \tag{5.2.15}$$

In adjacent charts, coordinates need not be identical, but since $ds^2 \propto dz_m d\bar{z}_m \propto dz_n d\bar{z}_n$, the transition functions are holomorphic. This is a mapping from Riemannian manifolds to two-real-dimensional complex manifolds. For the inverse mapping, one can take metric $dz_m d\bar{z}_m$ in the m-th chart and smooth it in the overlap region to produce a Riemannian manifold. The two characterizations of Riemann surfaces are thus equivalent.

The description of the torus using the complex coordinate w of the parallelogram (5.1.1) illustrates the idea of a complex manifold. Imagine taking a single coordinate chart slightly larger than the basic parallelogram of (5.1.1). The periodicity conditions

$$w \cong w + 2\pi \cong w + 2\pi\tau \tag{5.2.16}$$

are the transition functions on the overlap of the two opposite edges of the coordinate chart. These transition functions define the surface. To define the original path integral over the metric, treating the metric as fixed coordinates, such as the square (5.1.7), or at least a fixed set of coordinate charts with fixed transition functions, is the simplest starting point. To study quantum field theory on a given surface, the simplest approach is to handle unit metrics where transition function moduli are related, as in the parallelogram (5.1.11).

One can also see Riemann surfaces as follows. Define the worldsheet as a union of coordinate charts. We can use diff $\times$ Weyl degrees of freedom to reach the metric $dzd\bar{z}$ in each chart. Then the gauge choices in the chart overlap can only differ by the diff $\times$ Weyl invariance of this metric. This is exactly the conformal transformation we discussed. Thus, Riemann surfaces are the natural place for CFT to grow, because fields in CFT have clear transformation properties under conformal transformations.

5.3 The Measure for Moduli

We now review the gauge fixing of the Polyakov path integral. We already know that gauge redundancy does not completely eliminate the path integral over the metric, but leaves a finite-dimensional integral over the moduli space. To account for this, it is necessary to refine the discussion of section 3.3.

The path integral for the S-matrix is:

$$\begin{aligned} & S_{j_1 \ldots j_n}(k_1, \ldots, k_n) \\ & = \sum_{\text{topologies}} \int \frac{[d\phi dg]}{V_{\text{diff } \times \text{ Weyl}}} \exp(-S_{\mathrm{m}} - \lambda\chi) \prod_{i=1}^{n} \int d^2\sigma_i g(\sigma_i)^{1/2} \mathscr{V}_{j_i}(k_i, \sigma_i) \end{aligned} \tag{5.3.1}$$

This is the previous expression (3.5.5), but now with a universal matter field $c = \tilde{c} = 26$, where matter fields are labeled by ϕ. In gauge fixing, the integration over the metric and positions is replaced by an integration over the gauge group, moduli, and unfixed positions:

$$[dg]d^{2n}\sigma \to [d\zeta]d^{\mu}td^{2n-\kappa}\sigma \tag{5.3.2}$$

After factoring out the gauge volume, the Jacobian of this transformation becomes the measure on the moduli space.

This step is identical to section 3.3, now accounting for moduli and CKVs. Specifically, the gauge choice now fixes κ vertex coordinates, $\sigma_i^a \to \hat{\sigma}_i^a$. Denoting the set of fixed coordinates (a, i) as f, define the Faddeev-Popov measure on the moduli space as:

$$1 = \Delta_{\mathrm{FP}}(g, \sigma) \int_F d^{\mu}t \int_{\text{diff} \times \text{Weyl}} [d\zeta] \delta\left(g - \hat{g}(t)^{\zeta}\right) \prod_{(a,i) \in f} \delta\left(\sigma_i^a - \hat{\sigma}_i^{\zeta a}\right) \tag{5.3.3}$$

By definition, each metric (diff $\times$ Weyl) is equivalent to a $\hat{g}(t)$ for some t and ζ. In fact, as discussed at the end of section 5.1, there might be a residual discrete symmetry group of finite order n_R, so the delta function is non-zero at n_R points.

Substituting (5.3.3) into the path integral (5.3.1) and using the same steps as before gives:

$$\begin{aligned} S_{j_1 \ldots j_n}(k_1, \ldots, k_n) = & \sum_{\text{topologies}} \int_F d^{\mu}t \Delta_{\mathrm{FP}}(\hat{g}(t), \hat{\sigma}) \int [d\phi] \int \prod_{(a,i) \notin f} d\sigma_i^a \\ & \times \exp(-S_{\mathrm{m}}[\phi, \hat{g}(t)] - \lambda\chi) \prod_{i=1}^{n} \left[\hat{g}(\sigma_i)^{1/2} \mathscr{N}_{j_i}(k_i; \sigma_i)\right] \end{aligned} \tag{5.3.4}$$

The integration over metrics and vertex operator coordinates now reduces to an integration over the moduli space F plus unfixed coordinates with the measure given by Δ_{FP}. In the vertex operators, κ positions are fixed.

Now we calculate the Faddeev-Popov measure. The delta function is non-zero at n_R points related by symmetry, so we consider one such point and divide by n_R. Expand the definition of Δ_{FP} (5.3.3) near this point. A general metric variation equals a local symmetry variation plus a change in moduli t^k:

$$\delta g_{ab} = \sum_{k=1}^{\mu} \delta t^k \partial_{t^k} \hat{g}_{ab} - 2\left(\hat{P}_1 \delta\sigma\right)_{ab} + (2\delta\omega - \hat{\nabla} \cdot \delta\sigma)\hat{g}_{ab} \tag{5.3.5}$$

The Faddeev-Popov inverse determinant is:

$$\begin{aligned} & \Delta_{\mathrm{FP}}(\hat{g}, \hat{\sigma})^{-1} \\ & = n_R \int d^{\mu}\delta t[d\delta\omega d\delta\sigma] \delta(\delta g_{ab}) \prod_{(a,i) \in f} \delta(\delta\sigma^a(\hat{\sigma}_i)) \\ & = n_R \int d^{\mu}\delta t d^{\kappa}x \left[d\beta' d\delta\sigma\right] \\ & \quad \times \exp\left[2\pi i\left(\beta', 2\hat{P}_1\delta\sigma - \delta t^k \partial_k \hat{g}\right) + 2\pi i \sum_{(a,i) \in f} x_{ai} \delta\sigma^a(\hat{\sigma}_i)\right] \end{aligned} \tag{5.3.6}$$

The inner product definition in the second line refers to Exercise 3.2. We followed the same local discussion as in section 3.3, writing the delta function and functional as integrals over x and β_{ab}, and integrating over $\delta\omega$ to obtain the constraint that β'_{ab} is traceless.

As before, replace all bosonic variables with Grassmann variables to convert the integral (5.3.6):

$$\delta\sigma^a \to c^a \tag{5.3.7a}$$

$$\beta'_{ab} \to b_{ab} \tag{5.3.7b}$$

$$x_{ai} \to \eta_{ai} \tag{5.3.7c}$$

$$\delta t^k \to \xi^k \tag{5.3.7d}$$

Then, under the standard normalization of fields:

$$\begin{aligned}
\Delta_{\mathrm{FP}}(\hat{\mathrm{g}}, \hat{\sigma}) =& \frac{1}{n_R} \int [dbdc] d^\mu \xi d^\kappa \eta \\
& \times \exp \left[-\frac{1}{4\pi} \left(b, 2\hat{P}_1 c - \xi^k \partial_k \hat{g} \right) + \sum_{(a,i)\in f} \eta_{ai} c^a \left(\hat{\sigma}_i \right) \right] \\
=& \frac{1}{n_R} \int [dbdc] \exp \left(-S_{\mathrm{g}} \right) \prod_{k=1}^{\mu} \frac{1}{4\pi} \left(b, \partial_k \hat{g} \right) \prod_{(a,i)\in f} c^a \left(\hat{\sigma}_i \right)
\end{aligned} \tag{5.3.8}$$

In the last line, we integrated out the Grassmann variables η_{ai} and ξ^k. The appropriate integration measure on the moduli space is generated by the ghost path integral with this insertion.

We do not attempt to keep track of the overall sign in intermediate steps; since this is a Jacobian, we can implicitly choose the overall sign to yield a positive result. The complete expression for the S-matrix is:

$$\begin{aligned}
& S_{j_1 \ldots j_n} \left(k_1, \ldots, k_n \right) \\
& = \sum_{\text{topologies}} \int_F \frac{d^\mu t}{n_R} \int [d\phi dbdc] \exp \left(-S_{\mathrm{m}} - S_{\mathrm{g}} - \lambda \chi \right) \\
& \times \prod_{(a,i)\notin f} \int d\sigma_i^a \prod_{k=1}^{\mu} \frac{1}{4\pi} \left(b, \partial_k \hat{g} \right) \prod_{(a,i)\in f} c^a \left(\hat{\sigma}_i \right) \prod_{i=1}^{n} \hat{g} \left(\sigma_i \right)^{1/2} \mathscr{V}_{j_i} \left(k_i, \sigma_i \right)
\end{aligned} \tag{5.3.9}$$

This result is ready to be extended to all bosonic string theories—closed or open, oriented or non-oriented—the only difference being what topologies to include and what vertex operators to allow. Eq. (5.3.9) is a useful and beautiful result. The complexities introduced by moduli and CKVs are accounted for by inserting b and c into the path integral. For each fixed coordinate, $d\sigma_i^a$ is replaced by c_i^a, and each metric modulus gives a b insertion.

Representation in Determinant Form

In (5.3.8), we expressed the Faddeev-Popov determinant as an integral over Grassmann variables and fields. We now want to simplify it into a product of finite-dimensional functional determinants. As we continue to study the string S-matrix in the following chapters, this direct path integral approach is only one of the methods we will use.

Expand the ghost fields in a suitable complete basis:

$$c^a(\sigma) = \sum_J c_J \mathrm{C}_J^a(\sigma), \quad b_{ab}(\sigma) = \sum_K b_K \ \mathrm{B}_{Kab}(\sigma) \tag{5.3.10}$$

The complete bases C_J^a and B_{Kab} are defined as follows. The derivative P_1 appearing in the ghost action converts a vector into a traceless symmetric tensor. Its transpose performs the opposite function. The ghost action can be written as:

$$S_{\mathrm{g}} = \frac{1}{2\pi}(b, P_1 c) = \frac{1}{2\pi}\left(P_1^T b, c\right) \tag{5.3.11}$$

We cannot diagonalize P_1 in a diff-invariant way because it converts one type of field into another, but we can diagonalize $P_1^T P_1$ and $P_1 P_1^T$:

$$P_1^T P_1 \mathrm{C}_J^a = v_J'^2 \mathrm{C}_J^a, \quad P_1 P_1^T \ \mathrm{B}_{Kab} = v_K^2 \ \mathrm{B}_{Kab} \tag{5.3.12}$$

The eigenfunctions can be chosen to be real and normalized in the following inner products:

$$(\mathrm{C}_J, \mathrm{C}_{J'}) = \int d^2\sigma g^{1/2} \mathrm{C}_J^a \mathrm{C}_{J'a} = \delta_{JJ'} \tag{5.3.13a}$$

$$(\mathrm{B}_K, \ \mathrm{B}_{K'}) = \int d^2\sigma g^{1/2} \ \mathrm{B}_{Kab} \ \mathrm{B}_{K'}^{ab} = \delta_{KK'} \tag{5.3.13b}$$

Now note that:

$$\left(P_1 P_1^T\right) P_1 \mathrm{C}_J = P_1 \left(P_1^T P_1\right) \mathrm{C}_J = v_J'^2 P_1 \mathrm{C}_J \tag{5.3.14}$$

So $P_1 \mathrm{C}_J$ is an eigenfunction of $P_1 P_1^T$. In the same way, $P_1^T \ \mathrm{B}_K$ is an eigenfunction of $P_1^T P_1$. Therefore, unless $P_1 \mathrm{C}_J = 0$ or $P_1^T \ \mathrm{B}_K = 0$, there exists a one-to-one mapping between the eigenfunctions. They correspond to the eigenvalues of $P_1^T P_1$ or $P_1 P_1^T$. They are precisely the CKVs and holomorphic quadratic differentials, and their numbers are κ and μ, respectively.

We denote the zero-eigenvalue eigenfunctions as C_{0j} or B_{0k}, and non-zero eigenvalues are labeled $J, K = 1, \ldots$. For the latter, the normalization of the associated functions is:

$$\mathrm{B}_{Jab} = \frac{1}{v_J}(P_1 \mathrm{C}_J)_{ab}, \quad v_J = v_J' \neq 0 \tag{5.3.15}$$

In this basis, the ghost path integral Δ_{FP} becomes:

$$\int \prod_{k=1}^{\mu} db_{0k} \prod_{j=1}^{k} dc_{0j} \prod_J db_J dc_J \exp\left(-\frac{v_J b_J c_J}{2\pi}\right) \prod_{k'=1}^{\mu} \frac{1}{4\pi}(b, \partial_{k'}\hat{g}) \prod_{(a,i)\in f} c^a(\sigma_i) \tag{5.3.16}$$

From section A.2, we know that unless a variable appears in the integrand, the Grassmann integral is zero. c_{0j} and b_{0k} do not appear in the action, but only in the insertions. In fact, the number of each type of insertion in the integral (5.3.16), κ and μ, matches the ghost zero modes respectively. We have exactly enough insertions to give a non-zero result, and in the insertions, only the zero-mode part of the ghost fields contributes. Therefore:

$$\begin{aligned} \Delta_{\mathrm{FP}} = &\int \prod_{k=1}^{\mu} db_{0k} \prod_{k'=1}^{\mu} \left[\sum_{k''=1}^{\mu} \frac{b_{0k''}}{4\pi}(\mathrm{B}_{0k''}, \partial_{k'}\hat{g})\right] \\ &\times \int \prod_{j=1}^{\kappa} dc_{0j} \prod_{(a,i)\in f} \left[\sum_{j'=1}^{\kappa} c_{0j'} \mathrm{C}_{0j'}^a(\sigma_i)\right] \\ &\times \int \prod_J db_J dc_J \exp\left(-\frac{v_J b_J c_J}{2\pi}\right) \end{aligned} \tag{5.3.17}$$

Summing over all ways to fill the Grassmann variables in the zero-mode integrals, in each case, the integral generates a finite determinant, and the non-zero modes produce an infinite product, a functional determinant. Combined:

$$\Delta_{\mathrm{FP}} = \det \frac{(\mathrm{B}_{0k}, \partial_{k'}\hat{g})}{4\pi} \det \mathrm{C}_{0j}^a(\sigma_i) \det{}' \left(\frac{P_1^T P_1}{4\pi^2}\right)^{1/2} \tag{5.3.18}$$

Note that $\mathrm{C}_{0j}^a(\sigma_i)$ is a square matrix, where $(a, i) \in f$ ranges over κ values like j. The prime on the functional determinant indicates that the zero eigenvalues are removed.

Riemann-Roch Theorem

We now give a path integral derivation of the Riemann-Roch theorem. For the ghost current of the holomorphic bc system (no $\tilde{b}, \tilde{c}$), the current conservation anomaly derived in Exercise 3.6 is:

$$\nabla_a j^a = \frac{1-2\lambda}{4} R \tag{5.3.19}$$

Noether's theorem relates current conservation to invariance. For a non-conserved current, the discussion found:

$$\frac{\delta([d\phi]\exp(-S))}{[d\phi]\exp(-S)} = \frac{i\epsilon}{2\pi}\int d^2\sigma g^{1/2}\nabla_a j^a \to -i\epsilon\frac{2\lambda-1}{2}\chi \tag{5.3.20}$$

The ghost number symmetry acts as $\delta b = -i\epsilon b, \delta c = i\epsilon c$. When the path integral is non-zero, the transformations of the measure, action, and insertions must cancel each other. Thus, from the anomaly, we learn that the number of c insertions minus the number of b insertions is $3\chi/2$. The path integral calculation relates the number of insertions to the number of zero modes, thus giving the same difference:

$$\frac{1}{2}(\kappa-\mu) = \frac{1}{2}\left(\dim\ker P_1 - \dim\ker P_1^T\right) \tag{5.3.21}$$

The factor of $1/2$ appears because we only considered the holomorphic theory; the anti-holomorphic theory gives the same contribution. Equating the result from the anomaly with the result from the zero modes gives the Riemann-Roch theorem $\kappa - \mu = 3\chi$. For a general bc system, the same method will find:

$$\dim\ker P_n - \dim\ker P_n^T = (2n+1)\chi \tag{5.3.22}$$

where the P_n operator is defined in Exercise 3.2.

5.4 More about the measure

Here we collect some general properties of gauge-fixed string amplitudes. They play a foundational role in the more formal investigations in Chapter 9.

Gauge Invariance

The gauge-fixed amplitudes generated by the Faddeev-Popov treatment are BRST-invariant and independent of the gauge choice. However, it is useful to explicitly check whether it has all the expected properties.

First, it is independent of the choice of coordinates t on the moduli space. For new coordinates $t'^k(t)$:

$$\begin{aligned} d^\mu t' &= \left|\det\frac{\partial t'}{\partial t}\right| d^\mu t \\ \prod_{k=1}^{\mu}\frac{1}{4\pi}\left(b, \partial'_k\hat{g}\right) &= \det\left(\frac{\partial t}{\partial t'}\right)\prod_{k=1}^{\mu}\frac{1}{4\pi}\left(b, \partial_k\hat{g}\right) \end{aligned} \tag{5.4.1}$$

And these two Jacobians cancel each other out (up to a sign, which we manually fix to give a positive measure). In other words, the integral induced by the b ghost transforms like a density on the moduli space.

Second, let us see that the measure is invariant under Weyl transformations of the gauge slice. Under a Weyl transformation:

$$\delta\hat{g}'_{ab}(t;\sigma) = 2\delta\omega(t;\sigma)\hat{g}_{ab}(t;\sigma) \tag{5.4.2}$$

The variation of the action and measure produces an insertion of the local operator T_a^a, which is zero for $D = 26$ due to the equations of motion. Thus, we only need to worry about the effects on the individual insertions. Vertex operator insertions are zero due to contractions. The insertion of $c^a(\sigma_i)$ is zero according to the discussion below (3.3.24). For:

$$\begin{aligned}\left(b, \partial_k \hat{g}'\right) &= \int d^2\sigma \hat{g}'^{1/2} b_{ab} \hat{g}'^{ac} \hat{g}^{bd} \partial_k \hat{g}'_{cd} \\ &= \int d^2\sigma \hat{g}^{1/2} b_{ab} \left(\hat{g}^{ac}\hat{g}^{bd}\partial_k \hat{g}_{cd} + 2\hat{g}^{ab}\partial_k \omega\right) \\ &= (b, \partial_k \hat{g})\end{aligned} \tag{5.4.3}$$

The last equality holds due to the tracelessness of b.

Third, we check the invariance under infinitesimal diff transformations $\delta\sigma = \xi(t;\sigma)$. Extension to general diff transformations is direct. In the amplitude (5.3.9), the only terms that are not immediately invariant are the b insertions and the vertex operators with fixed coordinates. The former transforms as:

$$\begin{aligned}\delta\left(b, \partial_k \hat{g}\right) &= -2\left(b, P_1 \partial_k \xi\right) \\ &= -2\left(P_1^T b, \partial_k \xi\right) \\ &= 0\end{aligned} \tag{5.4.4}$$

where the equations of motion for b are used. The b equations of motion come from $\delta S/\delta c = 0$, so source terms will exist on the c insertions; they are needed to explain the effect of coordinate transformations on the fixed vertex operators.

BRST Invariance

We now prove the full BRST invariance of (5.3.9). From the local analysis, we know that the path integral measure and action are invariant, so we must examine the effect of BRST transformations on the insertions. In the gauge-fixed path integral (5.3.9), fixed vertex operators are accompanied by a factor of c or $\tilde{c}c$. This is our description of the BRST-invariant vertex operators corresponding to OCQ states at the end of section 4.4.

On the other hand, integrated vertex operators are not BRST-invariant; they have the BRST property:

$$\delta_{\mathrm{B}} \mathscr{V}_{\mathrm{m}} = i\epsilon \partial_a \left(c^a \mathscr{V}_{\mathrm{m}}\right) \tag{5.4.5}$$

which is zero in the integral sense. Finally, there is the variation of the b insertion:

$$\delta_{\mathrm{B}}\left(b, \partial_k \hat{g}\right) = i\epsilon\left(T, \partial_k \hat{g}\right) \tag{5.4.6}$$

The insertion of the energy-momentum tensor only produces a derivative with respect to t^k, so the BRST variation is zero except for a term that might come from the boundary of the moduli space. We will examine the boundary of the moduli space in the next chapter. In most cases, there is no surface contribution, which is due to the so-called "canceling propagator" argument. In certain situations, this argument does not apply, and we will see how to handle this.

Importantly, the gauge-fixed result (5.3.9) can be written directly from the requirement of BRST invariance without referring to the gauge-fixed form. As we saw when calculating the path integral, there exist at least μ b insertions and κ c insertions, otherwise the path integral is zero. In the amplitude (5.3.9), there are exactly the correct number of ghost insertions to give a non-zero result. Once we introduce the necessary b factors, the BRST transformation brings a convolution with the t^k derivative of the metric in the energy-momentum tensor. This is proportional to the derivative with respect to the moduli, so just as we did before, integrating over the moduli space yields an invariant result.

The result (5.3.9) can be generalized in various ways (e.g., the Conformal Killing invariant case is kept unfixed). In Chapter 7, for the torus without vertex operators, we will explicitly exemplify this point. BRST invariance implies that the amplitudes of BRST-equivalent states are the same. If we add a null part $Q_B|\chi\rangle$ to any state, the effect is to insert the variation $\delta_B \mathscr{V}_\chi$ into the path integral; this integral is zero. This is important because the physical Hilbert space is equivalent to the cohomology, so equivalent states should have the same amplitude.

The Measure for Riemann Surfaces

We derived the Faddeev-Popov measure in the framework where the gauge-fixed worldsheet is derived from a moduli-dependent metric. We now recast this result in the framework of Riemann surfaces, where the structure is encoded into moduli-dependent transition functions.

To express the value of the measure at a given point t_0 in the moduli space, we take a set of coordinate charts, where the m-th coordinate chart has complex coordinate z_m, and holomorphic transition functions, where the metric $\hat{g}\,(t_0)$ in each coordinate chart is equivalent to $dz_m d\bar{z}_m$. Now consider the change in moduli. We first describe it as a transformation in a metric with fixed transition functions, and then convert it into a transformation of transition functions with a fixed metric.

In the first description, define the Beltrami differential:

$$\mu_{ka}^b = \frac{1}{2}\hat{g}^{bc}\partial_k \hat{g}_{ac} \tag{5.4.7}$$

The b insertion for δt^k becomes:

$$\frac{1}{2\pi}\,(b,\mu_k) = \frac{1}{2\pi}\int d^2z \left(b_{zz}\mu_{k\bar{z}}^z + b_{\bar{z}\bar{z}}\mu_{kz}^{\bar{z}}\right) \tag{5.4.8}$$

In the second description, after the transformation δt^k in the moduli, there will be new coordinates in each chart:

$$z_m' = z_m + \delta t^k v_{km}^{z_m}\,(z_m,\bar{z}_m) \tag{5.4.9}$$

The superscript on v_{km}^a is a vector index; note that v_{km}^a is only defined in the m-th coordinate chart. By the definition of Riemann surfaces, $dz_m' d\bar{z}_m'$ is equivalent to the metric at $t_0 + \delta t$:

$$dz_m' d\bar{z}_m' \propto dz_m d\bar{z}_m + \delta t^k \left(\mu_{kz_m}^{\bar{z}_m} dz_m dz_m + \mu_{k\bar{z}_m}^{z_m} d\bar{z}_m d\bar{z}_m\right) \tag{5.4.10}$$

Therefore, the relation between the coordinate change v_{km}^a and the Beltrami differential is:

$$\mu_{kz_m}^{\bar{z}_m} = \partial_{z_m} v_{km}^{\bar{z}_m}, \quad \mu_{k\bar{z}_m}^{z_m} = \partial_{\bar{z}_m} v_{km}^{z_m} \tag{5.4.11}$$

This is the infinitesimal version of the Beltrami equation. It implies that $v_{km}^{z_m}$ is not holomorphic, otherwise it would not correspond to a change in moduli. Additionally, there is a holomorphic part in $v_{km}^{z_m}$ not determined by the Beltrami equation, and an anti-holomorphic part for $v_{km}^{\bar{z}_m}$; these correspond to the degrees of freedom for performing holomorphic reparameterizations in each chart.

Integrating by parts, the b insertion (5.4.8) becomes:

$$\frac{1}{2\pi}\,(b,\mu_k) = \frac{1}{2\pi i}\sum_m \oint_{C_m} \left(dz_m v_{km}^{z_m} b_{z_m z_m} - d\bar{z}_m v_{km}^{\bar{z}_m} b_{\bar{z}_m\bar{z}_m}\right) \tag{5.4.12}$$

The contour C_m surrounds the m-th coordinate chart counter-clockwise. By (5.4.9), the derivative of the coordinates with respect to the moduli at a given point is:

$$\frac{dz_m}{dt^k} = v_{km}^{z_m} \tag{5.4.13}$$

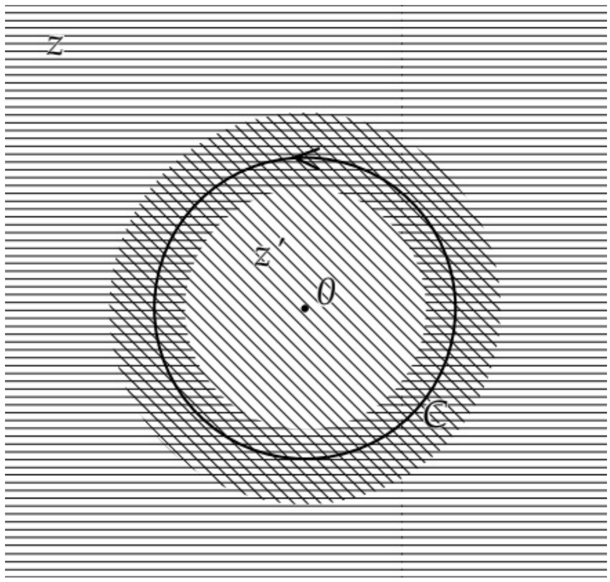


Figure 5.4.1: Coordinate patches z' (diagonal hatching) and z (horizontal hatching) with vertex operator at $z' = 0$. In the annular region, $z = z' + z_v$.

Thus, the transformation of the transition functions under a modular transformation is:

$$\left.\frac{\partial z_m}{\partial t^k}\right|_{z_n} = v_{km}^{z_m} - \left.\frac{\partial z_m}{\partial z_n}\right|_t v_{kn}^{z_n} = v_{km}^{z_m} - v_{kn}^{z_m} \tag{5.4.14}$$

Then, combining the contour integrals surrounding adjacent charts, (5.4.12) gives:

$$\frac{1}{2\pi}(b, \mu_k) = \frac{1}{2\pi i}\sum_{(mn)} \int_{C_{mn}} \left(dz_m \left.\frac{\partial z_m}{\partial t^k}\right|_{z_n} b_{z_m z_m} - d\bar{z}_m \left.\frac{\partial \bar{z}_m}{\partial t^k}\right|_{z_n} b_{\bar{z}_m \bar{z}_m} \right) \tag{5.4.15}$$

This is proportional to the derivative of the transition functions. The sum is taken over all overlapping charts. The contour C_{mn} runs between charts m and n, and is counter-clockwise from the perspective of m. The (mn) terms are symmetric with respect to m and n, though this is not obvious. Each contour is either closed or terminates at triple overlaps of charts m, n, p, where contours C_{mn}, C_{np}, and C_{pm} meet at a point. The b insertion is now expressed explicitly in terms of the data defining the Riemann surface, where the transition functions are taken modulo holomorphic equivalence.

As an illustration, we can use this to place the moduli of the metric and the moduli of the vertex operator positions on a more equal footing. Consider a vertex operator at z_v in coordinate system z. Take a small new coordinate chart z' centered at the vertex operator, as shown in figure (5.4.1).

We keep the vertex operator at $z' = 0$ and encode its position into the transition function:

$$z = z' + z_v \tag{5.4.16}$$

The measure for $z_v, \bar{z}_v$ is given by (5.4.15), where:

$$\left.\frac{\partial z'}{\partial z_v}\right|_z = -1 \tag{5.4.17}$$

Therefore, the ghost insertion is:

$$\int_C \frac{dz'}{2\pi i} b_{z'z'} \int_C \frac{d\bar{z}'_m}{-2\pi i} b_{\bar{z}'\bar{z}'} = b_{-1}\tilde{b}_{-1} \tag{5.4.18}$$

where C is any contour surrounding the vertex operator as shown. Then, the complete expression for the S-matrix can be compactly written as:

$$S(1;\ldots;n) = \sum_{\text{compact topologies}} e^{-\lambda\chi} \int_F \frac{d^m t}{n_R} \left\langle \prod_{k=1}^{m} B_k \prod_{i=1}^{n} \hat{\mathscr{V}}_i \right\rangle \tag{5.4.19}$$

Here, B_k is shorthand for the b ghost insertion (5.4.15), and $\hat{\mathscr{V}}$ represents $\tilde{c}c\mathscr{V}_{\mathrm{m}}$ for closed strings and $t_a c^a \mathscr{V}_{\mathrm{m}}$ for open strings. That is, we now treat all vertex operators as fixed and replace their coordinates with extra parameters in the transition functions. The number of moduli m is:

$$m = \mu + 2n_{\mathrm{c}} + n_{\mathrm{o}} - \kappa = -3\chi + 2n_{\mathrm{c}} + n_{\mathrm{o}} \tag{5.4.20}$$

where n_{c} and n_{o} are the numbers of closed string and open string vertex operators, respectively.

(5.4.19) shows that hatted vertex operators, i.e., those containing $\tilde{c}c$ or $t_a c^a$, are fundamental. This is exactly what the state-operator correspondence produces, and it is BRST-invariant. If the vertex operator is integrated, the ghost insertion (5.4.18) will remove $\tilde{c}c$ or $t_a c^a$. For example:

$$b_{-1}\tilde{b}_{-1} \cdot \tilde{c}c\mathscr{V}_{\mathrm{m}} = \mathscr{V}_{\mathrm{m}} \tag{5.4.21}$$

leaving the integrated form of the vertex operator. We wrote (5.4.19) from the Polyakov path integral, so vertex operators appear in the form of OCQ, but now we can use any BRST-invariant vertex operator.

Chapter 6

Tree-level amplitudes

We now study string interactions . In this chapter, we will investigate the lowest-order amplitudes, which originate from surfaces with positive Euler number . We first describe the relevant Riemann surfaces and calculate the required CFT expectation values . Next, we will study scattering amplitudes, first for open strings and then for closed strings . Along the way, we will introduce an important generalization in open string theory: Chan-Paton factors . At the end of this section, we will return to CFT and discuss the general properties of expectation values .

6.1 Riemann surfaces

There are three types of Riemann surfaces with positive Euler number: the sphere, the disk, and the projective plane .

The Sphere

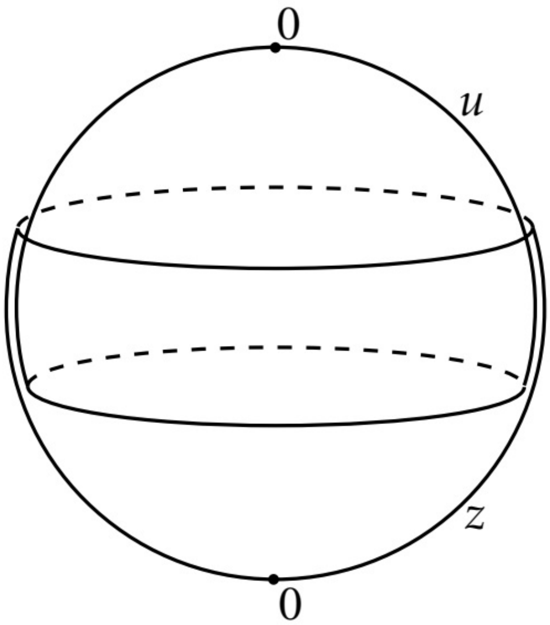


Figure 6.1.1: A sphere constructed from the z and u coordinate charts .

The sphere S_2 can be covered by two coordinate charts, as shown in figure 6.1.1 . We take disks $|z| < \rho$ and $|u| < \rho$, where $\rho > 1$, and glue them via :

$$u = 1/z \tag{6.1.1}$$

In fact, we can take $\rho \to \infty$. In this way, the coordinate z is well-defined everywhere except at the "North Pole" $u = 0$. We can view the sphere as a Riemann surface by taking flat metrics on both coordinate charts and connecting them with a conformal (coordinate plus Weyl) transformation . Alternatively, we can view it as a Riemannian manifold with a globally defined metric . The general conformal gauge metric is :

$$\mathrm{d}s^2 = \exp(2\omega(z,\bar{z}))\mathrm{d}z\mathrm{d}\bar{z}\,. \tag{6.1.2}$$

Since $\mathrm{d}z\mathrm{d}\bar{z} = |z|^4\mathrm{d}u\mathrm{d}\bar{u}$, the condition for this metric to be non-singular at $u = 0$ is that $\exp(2\omega(z,\bar{z}))$ decays as $|z|^{-4}$ as $z \to \infty$. For example :

$$\mathrm{d}s^2 = \frac{4r^2\,\mathrm{d}z\mathrm{d}\bar{z}}{(1+z\bar{z})^2} = \frac{4r^2\,\mathrm{d}u\mathrm{d}\bar{u}}{(1+u\bar{u})^2} \tag{6.1.3}$$

describes a sphere with radius r and curvature $R = 2/r^2$.

According to the general discussion in section **??**, the sphere has no moduli but has 6 CKVs, making every metric (diff $\times$ Weyl) equivalent to (6.1.3) . We observe this at the infinitesimal level . As in (**??**), we look for holomorphic tensor fields $\delta g_{zz}(z)$ and holomorphic vector fields $\delta z(z)$. These must be defined on the entire sphere, so we examine the transformation to the u coordinate chart :

$$\delta u = \frac{\partial u}{\partial z}\delta z = -z^{-2}\delta z\,, \tag{6.1.4a}$$

$$\delta g_{uu} = \left(\frac{\partial u}{\partial z}\right)^{-2}\delta g_{zz} = z^4\delta g_{zz}\,. \tag{6.1.4b}$$

Any holomorphic quadratic differential δg_{zz} must be holomorphic with respect to z and vanish as z^{-4} at infinity; therefore, it must be identically zero . On the other hand, the CKV δz being holomorphic at $u = 0$ requires it to grow no faster than z^2 as $z \to \infty$. Thus, the general CKV is :

$$\delta z = a_0 + a_1 z + a_2 z^2\,, \tag{6.1.5a}$$

$$\delta\bar{z} = a_0^* + a_1^*\bar{z} + a_2^*\bar{z}^2\,, \tag{6.1.5b}$$

which has 3 complex parameters or 6 real parameters, exactly as expected from the Riemann-Roch theorem . Exponentiating these infinitesimal transformations gives the *Möbius group* :

$$z' = \frac{\alpha z + \beta}{\gamma z + \delta}\,. \tag{6.1.6}$$

Rescaling $\alpha, \beta, \gamma, \delta$ does not change the transformation, so we can fix $\alpha\delta - \beta\gamma = 1$ and treat the set under global sign inversion of $\alpha, \beta, \gamma, \delta$ as equivalent, which defines the $PSL(2, \mathbb{C})$ group . This is the most general coordinate transformation that is holomorphic on the entire S_2 . it is a one-to-one mapping including the point at infinity . Three of the six parameters correspond to ordinary rotations, forming the subgroup $SO(3)$ of $PSL(2, \mathbb{C})$.

The Disk

The disk D_2 can be constructed from the sphere by identifying two points under reflection . For example, identify z and z' such that :

$$z' = 1/\bar{z}\,. \tag{6.1.7}$$

In polar coordinates $z = re^{i\phi}$, this takes the reciprocal of the radius but preserves the angle, so the unit disk $|z| \leq 1$ is the fundamental domain after gluing . Points on the unit circle are fixed by the reflection, so it becomes a boundary . It is often more convenient to use the conformally equivalent inversion :

$$z' = \bar{z} \tag{6.1.8}$$

Now the upper half-plane is the fundamental domain, and the real axis is the boundary .

The CKG of the disk is the subgroup of $PSL(2,\mathbb{C})$ that keeps the boundary of the disk invariant . For the inversion (6.1.8), this is the subgroup where $\alpha, \beta, \gamma, \delta$ in (6.1.6) are all real, namely $PSL(2,\mathbb{R})$, the Möbius group of real parameters . One CKV is the ordinary rotational symmetry on the disk . Once again, all metrics are equivalent, and there are no moduli .

The Projective Plane

The projective plane RP_2 can also be obtained as a $\mathbb{Z}_2$ identification of the sphere . Glue points z and z' such that :

$$z' = -1/\bar{z} \tag{6.1.9}$$

These points are antipodal points in the round metric (6.1.3) . There are no fixed points, so there is no boundary in the resulting space, but the space is non-orientable . A fundamental domain for this gluing is the unit disk $|z| \leq 1$, but points $e^{i\phi}$ and $-e^{i\phi}$ are identified . Another choice is the upper z-plane, where there are no moduli . The CKG is the subgroup of $PSL(2,\mathbb{C})$ reflecting the (6.1.9) identification; this is the ordinary rotation group $SO(3)$. Both the disk and the projective plane can be represented as spheres with points identified under a $\mathbb{Z}_2$ transformation or *involution* . In fact, every worldsheet can be obtained from a closed oriented surface by identification under one or two $\mathbb{Z}_2$ involutions . Thus, Green's functions can be obtained using the method of images .

6.2 Expectation values of scalars

The basic quantities we need are the expectation values of vertex operators . To develop several useful techniques and viewpoints, we will calculate them in three different ways: direct path integrals, using holomorphy, and the operator method later in this chapter .

The path integral method has already been used for the Faddeev-Popov determinant in section **??** and for the harmonic oscillator in the appendix . Starting from a general functional :

$$Z[J] = \left\langle \exp\left(i \int \mathrm{d}^2\sigma\, J(\sigma)\cdot X(\sigma) \right) \right\rangle \tag{6.2.1}$$

where $J_\mu(\sigma)$ is arbitrary . From now on, we work on an arbitrary compact two-dimensional surface M, and the spacetime dimension is d . We expand $X^\mu(\sigma)$ on a complete basis $\mathsf{X}_I(\sigma)$:

$$X^\mu(\sigma) = \sum_I x_I^\mu \mathsf{X}_I(\sigma)\,, \tag{6.2.2a}$$

$$\nabla^2 \mathsf{X}_I = -\omega_I^2 \mathsf{X}_I\,, \tag{6.2.2b}$$

$$\int_M \mathrm{d}^2\sigma\, g^{1/2} \mathsf{X}_I \mathsf{X}_{I'} = \delta_{II'}\,. \tag{6.2.2c}$$

Then :

$$Z[J] = \prod_{I,\mu} \int \mathrm{d}x_I^\mu \exp\left(-\frac{\omega_I^2 x_I^\mu x_{I\mu}}{4\pi\alpha'} + i x_I^\mu J_{I\mu} \right)\,, \tag{6.2.3}$$

where :

$$J_I^\mu = \int \mathrm{d}^2\sigma J^\mu(\sigma) \mathsf{X}_I(\sigma)\,. \tag{6.2.4}$$

Except for the constant mode :

$$\mathsf{X}_0 = \left(\int \mathrm{d}^2\sigma\, g^{1/2}\right)^{-1/2} , \tag{6.2.5}$$

this integral is Gaussian . The action for the constant mode is zero, so it yields a δ-function .

> **Remark.** *For $I = 0$, (6.2.3) gives :*
>
> $$\int \mathrm{d}x_0^\mu \; \exp(ix_0^\mu J_{0\mu}) \propto \delta^d(J_0) \,.$$

Evaluating the integral yields :

$$\begin{aligned} Z[J] &= i(2\pi)^d \delta^d(J_0) \prod_{I\neq 0} \left(\frac{4\pi^2\alpha'}{\omega_I^2}\right)^{d/2} \exp\left(-\frac{\pi\alpha' J_I\cdot J_I}{\omega_I^2}\right) \\ &= i(2\pi)^d \delta^d\left(J_0\right) \left(\det{}' \frac{-\nabla^2}{4\pi^2\alpha'}\right)^{-d/2} \\ &\quad \times \exp\left(-\frac{1}{2}\int \mathrm{d}^2\sigma\, \mathrm{d}^2\sigma'\; J(\sigma)\cdot J(\sigma')G'(\sigma,\sigma')\right) \,. \end{aligned} \tag{6.2.6}$$

As discussed in sections 2.1 and 3.2, the Gaussian integral generated by the timelike mode x_I^0 has the wrong sign, so it is defined via contour rotation $x_I^0 \to -ix_I^d, I \neq 0$. The primed Green's function removes the zero-mode contribution :

$$G'(\sigma_1,\sigma_2) = \sum_{I\neq 0} \frac{2\pi\alpha'}{\omega_I^2} \mathsf{X}_I(\sigma_1)\mathsf{X}_I(\sigma_2) \,. \tag{6.2.7}$$

It satisfies the differential equation :

$$\begin{aligned} -\frac{1}{2\pi\alpha'}\nabla^2 G'(\sigma_1,\sigma_2) &= \sum_{I\neq 0} \mathsf{X}_I(\sigma_1)\mathsf{X}_I(\sigma_2) \\ &= g^{-1/2}\delta^2(\sigma_1-\sigma_2) - \mathsf{X}_0^2 \,, \end{aligned} \tag{6.2.8}$$

where the completeness of X_I is used . An ordinary Green's function with a δ-function source does not exist . It would correspond to the electrostatic potential of a single charge, but on a compact surface, field lines flowing out of the source have nowhere to go . The X_0^2 term can be viewed as a background charge distribution used to neutralize the system .

The Sphere

Specifically on the sphere, the solution to the differential equation (6.2.8) is :

$$G'(\sigma_1,\sigma_2) = -\frac{\alpha'}{2}\ln|z_{12}|^2 + f(z_1,\bar z_1) + f(z_2,\bar z_2) \,. \tag{6.2.9}$$

> **Remark.** *On the sphere $g = \mathrm{e}^{4\omega}$, $\nabla^2 = \mathrm{e}^{-2\omega}\partial\bar\partial$, so (6.2.8) becomes :*
>
> $$\partial\bar\partial G' = -2\pi\alpha'\delta^2(\sigma_1-\sigma_2) + 2\pi\alpha'\mathrm{e}^{2\omega}\,\mathsf{X}_0^2$$
>
> *where $\partial\bar\partial\ln|z_{12}|^2 = 2\pi\delta^2(z,\bar z) = 4\pi\delta^2(\sigma_1-\sigma_2)$.*

where :

$$f(z,\bar{z}) = \frac{\alpha' \mathsf{X}_0^2}{4} \int \mathrm{d}^2 z' \ \exp\left[2\omega(z',\bar{z}')\right] \ln|z - z'|^2 + k \,. \tag{6.2.10}$$

The condition that G' is orthogonal to X_0 determines the constant k, but in any case, we will see that the function f is independent of all expectation values . It originates from the background charge, but the δ-function from the zero-mode integration forces overall neutrality, $J_0^\mu = 0$, so the background field has no net contribution .

Now consider the path integral containing a product of tachyon vertex operators :

$$A^n_{S_2}(k,\sigma) = \left\langle \left[\mathrm{e}^{ik_1\cdot X(\sigma_1)}\right]_\mathrm{r} \left[\mathrm{e}^{ik_2\cdot X(\sigma_2)}\right]_\mathrm{r} \cdots \left[\mathrm{e}^{ik_n\cdot X(\sigma_n)}\right]_\mathrm{r} \right\rangle_{S_2} . \tag{6.2.11}$$

This corresponds to :

$$J(\sigma) = \sum_{i=1}^n k_i \delta^2(\sigma - \sigma_i) \,. \tag{6.2.12}$$

Then the amplitude (6.2.6) becomes :

$$\begin{aligned} A^n_{S_2}(k,\sigma) = {} & i C^X_{S_2} (2\pi)^d \delta^d(\textstyle\sum_i k_i) \\ & \times \exp\left(- \sum_{i<j}^n k_i \cdot k_j G'(\sigma_i,\sigma_j) - \frac{1}{2} \sum_{i=1}^n k_i^2 G'_\mathrm{r}(\sigma_i,\sigma_i) \right) . \end{aligned} \tag{6.2.13}$$

The constant here is :

$$C^X_{S_2} = \mathsf{X}_0^{-d} \left(\det{}' \frac{-\nabla^2}{4\pi^2\alpha'} \right)^{-d/2}_{S_2} . \tag{6.2.14}$$

Remark. *According to* (6.2.4) *and* (6.2.5)*,* $J_0 = \int \mathrm{d}^2\sigma\, \mathsf{X}_0 J(\sigma) = \mathsf{X}_0 \sum_i k_i$*, so* $\delta^d(J_0) = \mathsf{X}_0^{-d} \delta^d(\sum_i k_i)$ *.*

The determinant can be regularized and calculated, but we do not need to calculate it explicitly . We adopt the renormalization from section 3.6 such that self-contractions include :

$$G'_\mathrm{r}(\sigma,\sigma') = G'(\sigma,\sigma') + \frac{\alpha'}{2} \ln d^2(\sigma,\sigma') \,. \tag{6.2.15}$$

Note that :

$$G'_\mathrm{r}(\sigma,\sigma) = 2f(z,\bar{z}) + \alpha' \omega(z,\bar{z}) \tag{6.2.16}$$

is finite . Then the path integral on the sphere is :

$$A^n_{S_2}(k,\sigma) = i C^X_{S_2} (2\pi)^d \delta^d(\textstyle\sum_i k_i) \exp\left(-\dfrac{\alpha'}{2} \displaystyle\sum_i k_i^2 \omega\left(\sigma_i\right) \right) \prod_{i<j} |z_{ij}|^{\alpha' k_i \cdot k_j} \,. \tag{6.2.17}$$

The dependence on the conformal factor $\omega(\sigma)$ is obtained from the Weyl anomaly of the vertex operators . For on-shell operators, it will cancel the variation of $g^{1/2}$. We will take the metric to be flat (pushing the curvature to infinity) in a large region containing all vertex operators; thus, terms in $\omega(\sigma_i)$ are dropped .

Notes on (6.2.16):

$$\ln d^2(\sigma,\sigma') - \ln|z_{12}|^2 = \ln\left(\frac{\int \mathrm{e}^{2\omega}\,\mathrm{d}z\mathrm{d}\bar{z}}{\int \mathrm{d}z\mathrm{d}\bar{z}}\right) = 2\omega$$

Notes on (6.2.17): Taking $n=2$ as an example, the part depending on f is

$$k_1k_2(f_1+f_2) + \frac{1}{2}k_1^2\cdot 2f_1 + \frac{1}{2}k_2^2\cdot 2f_2 = (k_1f_1+k_2f_2)(k_1+k_2)$$

Since the path integral contains $\delta(\sum_i k_i) = k_1+k_2$, it is independent of f .

Higher-order vertex operators are exponentials multiplied by partial derivatives of X^μ, so we further need :

$$\left\langle \prod_{i=1}^{n}\left[\mathrm{e}^{ik_i\cdot X(z_i,\bar{z}_i)}\right]_{\mathrm{r}} \prod_{j=1}^{p}\partial X^{\mu_j}(z'_j)\prod_{k=1}^{q}\bar{\partial}X^{\nu_k}(\bar{z}''_k)\right\rangle_{S_2} . \tag{6.2.18}$$

This is given by summing over all contractions, where each ∂X or $\bar{\partial}X$ must contract with an exponential or other ∂X and $\bar{\partial}X$. The XX contraction is $-\frac{1}{2}\alpha'\ln|z|^2$, and f is again dropped in the final expression . This result is summarized as :

$$\begin{aligned} &iC_{S_2}^X(2\pi)\delta^d(\textstyle\sum_i k_i)\exp\left(-\frac{\alpha'}{2}\displaystyle\sum_i k_i^2\omega(\sigma_i)\right)\prod_{i<j}^{n}|z_{ij}|^{\alpha' k_i\cdot k_j} \\ &\qquad\times\left\langle \prod_{j=1}^{p}\left[v^{\mu_j}(z'_j)+q^{\mu_j}(z'_j)\right]\prod_{k=1}^{q}\left[\tilde{v}^{\nu_k}(\bar{z}''_k)+\tilde{q}^{\nu_k}(\bar{z}''_k)\right]\right\rangle \end{aligned} \tag{6.2.19}$$

where :

$$v^\mu(z) = -i\frac{\alpha'}{2}\sum_{i=1}^{n}\frac{k_i^\mu}{z-z_i}, \qquad \tilde{v}^\mu(\bar{z}) = -i\frac{\alpha'}{2}\sum_{i=1}^{n}\frac{k_i^\mu}{\bar{z}-\bar{z}_i} \tag{6.2.20}$$

originate from contractions with the exponentials . The expectation value of $q^\mu = \partial X^\mu - v^\mu$ is given by contraction with $-\eta^{\mu\nu}(z-z')^{-2}\alpha'/2$ and summing, while the expectation value of $\tilde{q}^\nu$ is given by conjugation .

Now we re-calculate in a different way, using holomorphy . As an example, consider :

$$\langle \partial X^\mu(z_1)\partial X^\nu(z_2)\rangle_{S_2} . \tag{6.2.21}$$

The OPE determines it to be :

$$-\frac{\alpha'\eta^{\mu\nu}}{2z_{12}^2}\langle 1\rangle_{S_2} + g(z_1,z_2)\,, \tag{6.2.22}$$

where $g(z_1,z_2)$ is holomorphic in both variables . In the u coordinate chart :

$$\partial_u X^\mu = -z^2\partial_z X^\mu\,, \tag{6.2.23}$$

so the condition for holomorphy at $u=0$ is that the expectation value of $\partial_z X^\mu$ decays as z^{-2} as z goes to infinity . More generally, the behavior of a tensor with weight $(h,0)$ at infinity is z^{-2h} . Focusing on the correlation of (6.2.22) with respect to z_1 at fixed z_2, this implies $g(z_1,z_2)$ decays as z_1^{-2} at infinity, hence the holomorphic factor is zero . Compared with the path integral result (6.2.19), this is consistent . Although $\langle 1\rangle_{S_2}$ seems to diverge like $\delta(0)$, this is just the zero-mode divergence from the infinite spacetime volume . Regarding the expectation value of a product of vertex operators :

$$A_{S_2}^n(k,\sigma) = \left\langle \prod_{i=1}^{n} :\mathrm{e}^{ik_i\cdot X(z_i,\bar{z}_i)}: \right\rangle_{S_2} \tag{6.2.24}$$

calculating with this method is less direct because they are not holomorphic . In fact, they factorize into holomorphic and anti-holomorphic parts, but this is subtle and is best introduced in the language of operators, which we will do in chapter 8 . We use the holomorphy of the translation current . Consider the expectation value with an additional current $\partial X^\mu(z)$. The OPE of the current and vertex operators determines the singularity at z :

$$\left\langle \partial X^\mu(z) \prod_{i=1}^n :\mathrm{e}^{ik_i\cdot X(z_i,\bar{z}_i)}: \right\rangle_{S_2} = -\frac{i\alpha'}{2} A^n_{S_2}(k,\sigma) \sum_{i=1}^n \frac{k_i^\mu}{z-z_i} + \text{terms holomorphic at } z\,. \tag{6.2.25}$$

Now observe $z\to\infty$. The condition that $\partial_u X^\mu$ is holomorphic at $u=0$ again requires this expectation value to vanish as z^{-2} as $z\to\infty$. Then the holomorphic terms in (6.2.25) are zero, and the term at order z^{-1} is zero . We find momentum conservation :

$$A^n_{S_2}(k,\sigma)\sum_{i=1}^n k_i^\mu = 0\,. \tag{6.2.26}$$

We state this in a slightly different way . Consider the contour integral of the spacetime translation current :

$$p^\mu = \frac{1}{2\pi i}\oint_C (\mathrm{d}z\, j_z^\mu - \mathrm{d}\bar{z}\, j_{\bar{z}}^\mu)\,, \tag{6.2.27}$$

where the contour C encloses all exponential operators . There are two ways to calculate it . One is to shrink the contour C until it becomes a small circle around each vertex operator, which picks out the $(z-z_i)^{-1}$ term from each OPE and gives $\sum_{i=1}^n k_i^\mu$. The other is to expand it until it is a small circle on the u coordinate chart: due to holomorphy at $u=0$, it must be zero . This discussion of deforming contours is widely used in CFT . We have used the OPE to determine singularities and then used holomorphy to completely determine the dependence on z . Now let us look at the second term of the OPE as $z\to z_1$. Expanding (6.2.25) gives :

$$-\frac{i\alpha'}{2}A^n_{S_2}(k,\sigma)\left(\frac{k_1^\mu}{z-z_1}+\sum_{i=2}^n\frac{k_i^\mu}{z_1-z_i}+O(z-z_1)\right)\,. \tag{6.2.28}$$

Contracting with ik_1^μ, the $(z-z_1)^0$ order term must be consistent with the OPE :

$$ik_1\cdot\partial X(z)\,:\mathrm{e}^{ik_1\cdot X(z_1,\bar{z}_1)}:=\frac{\alpha' k_1^2}{2(z-z_1)}\,:\mathrm{e}^{ik_1\cdot X(z_1,\bar{z}_1)}:+\,\partial_{z_1}:\mathrm{e}^{ik_1\cdot X(z_1,\bar{z}_1)}:+\,O(z-z_1)\,. \tag{6.2.29}$$

This implies :

$$\partial_{z_1}A^n_{S_2}(k,\sigma)=\frac{\alpha'}{2}A^n_{S_2}(k,\sigma)\sum_{i=2}^n\frac{k_1\cdot k_i}{z_{1i}}\,. \tag{6.2.30}$$

Integrating and using conjugate equations as well as momentum conservation . In the sense of normalization difference, this determines the path integral :

$$A^n_{S_2}(k,\sigma)\propto\delta^d(\textstyle\sum_i k_i)\displaystyle\prod_{i<j}^n|z_{ij}|^{\alpha' k_i\cdot k_j}\,, \tag{6.2.31}$$

which is consistent with the first method . Note that for comparison, we must push the curvature to infinity ($\omega(\sigma_i)=0$), which makes $[]_\mathrm{r}=::$.

It is worth noting that the intermediate steps of the path method depend on the specific choice of the Riemann metric . This metric is needed to preserve coordinate invariance in these steps . Two different metrics, if Weyl equivalent, still give different Laplacians and eigenfunctions . But in the end, in string theory, these correlations must be dropped . The second method uses only the fundamental properties of the Riemann surface, its holomorphy .

The Disk

The extension to the disk is direct . By taking the representation of the sphere above and constraining z to the upper complex plane, we obtain the representation of the disk . The Neumann boundary term is interpreted as an image charge :

$$G'(\sigma_1, \sigma_2) = -\frac{\alpha'}{2} \ln |z_1 - z_2|^2 - \frac{\alpha'}{2} \ln |z_1 - \bar{z}_2|^2 \,, \tag{6.2.32}$$

differing by a term that is dropped due to momentum conservation . Then :

$$\begin{aligned} \left\langle \prod_{i=1}^{n} :\mathrm{e}^{ik_i \cdot X(z_i, \bar{z}_i)}: \right\rangle_{D_2} =& iC^X_{D_2} (2\pi)^d \delta^d(\textstyle\sum_i k_i) \displaystyle\prod_{i=1}^{n} |z_i - \bar{z}_i|^{\alpha' k_i^2/2} \\ & \times \prod_{i<j}^{n} |z_i - z_j|^{\alpha' k_i \cdot k_j} |z_i - \bar{z}_j|^{\alpha' k_i \cdot k_j} \,. \end{aligned} \tag{6.2.33}$$

For expectation values containing $\partial_a X^\mu$, one again sums over contractions, but now using the Green's function (6.2.32) . Note that $\partial X^\mu(z) \bar{\partial} X^\nu(\bar{z}')$ (or $q^\mu(z) \tilde{q}^\nu(\bar{z}')$) has a non-zero contraction .

For two points on the boundary, the two terms in the Green's function (6.2.32) are equal; even after subtracting the first term in normal ordering, the Green's function still diverges at zero interval . For this reason, boundary operators must be defined using boundary normal ordering, where the subtraction is doubled :

$$\circ X^\mu(y_1) X^\nu(y_2) \circ = X^\mu(y_1) X^\nu(y_2) + 2\alpha' \eta^{\mu\nu} \ln |y_1 - y_2| \,, \tag{6.2.34}$$

where y represents coordinates on the real axis . Combinatorics is the same as for other forms of normal ordering . The expectation values of boundary normal ordered boundary operators have the same well-behaved properties (no singularities) as internal conformal normal ordered operators .

If an expectation value contains both boundary exponentials and internal exponentials, each with appropriate normal ordering, drop the $|z_i - \bar{z}_i|^{\alpha' k_i^2/2}$ factor for the boundary operators and take the appropriate limit of the internal result (6.2.33) . Explicitly, for exponentials all on the boundary :

$$\left\langle \prod_{i=1}^{n} \circ \mathrm{e}^{ik_i \cdot X(y_i)} \circ \right\rangle_{D_2} = iC^X_{D_2} (2\pi)^d \delta^d(\textstyle\sum_i k_i) \prod_{i<j}^{n} |y_i - y_j|^{2\alpha' k_i \cdot k_j} \,. \tag{6.2.35}$$

More generally :

$$\begin{aligned} \left\langle \prod_{i=1}^{n} \circ \mathrm{e}^{ik_i \cdot X(y_i)} \circ \prod_{j=1}^{p} \partial_y X^{\mu_j}(y'_j) \right\rangle_{D_2} &= iC^X_{D_2} (2\pi)^d \delta^d(\textstyle\sum_i k_i) \\ & \times \prod_{i<j}^{n} |y_{ij}|^{2\alpha' k_i \cdot k_j} \left\langle \prod_{j=1}^{p} \Big[v^{\mu_j}(y'_j) + q^{\mu_j}(y'_j) \Big] \right\rangle_{D_2} \,, \end{aligned} \tag{6.2.36}$$

where :

$$v^\mu(y) = -2i\alpha' \sum_{i=1}^{n} \frac{k_i^\mu}{y - y_i} \tag{6.2.37}$$

and q is contracted with $-2\alpha'(y - y')^{-2} \eta^{\mu\nu}$.

The Projective Plane

The Green's function given by the method of images is :

$$G'(\sigma_1, \sigma_2) = -\frac{\alpha'}{2} \ln |z_1 - z_2|^2 - \frac{\alpha'}{2} \ln |1 + z_1 \bar{z}_2|^2 \,. \tag{6.2.38}$$

Now :

$$\begin{aligned} \left\langle \prod_{i=1}^{n} :\mathrm{e}^{ik_i X(z_i, \bar{z}_i)}: \right\rangle_{RP_2} &= i C^X_{RP_2} (2\pi)^d \delta^d(\textstyle\sum_i k_i) \prod_{i=1}^{n} |1 + z_i \bar{z}_i|^{\alpha' k_i^2/2} \\ &\times \prod_{i<j}^{n} |z_i - z_j|^{\alpha' k_i \cdot k_j} |1 + z_i \bar{z}_j|^{\alpha' k_i k_j} \,, \end{aligned} \tag{6.2.39}$$

and similarly for more general expectation values . Since there are no fixed points, it is impossible for point z to touch its image $-\bar{z}^{-1}$, so there is no boundary .

6.3 bc Conformal Field Theory

The Sphere

The path integral for ghost fields has already been established in section ?? . According to the Riemann-Roch theorem, the simplest non-zero expectation value is :

$$\langle c(z_1)c(z_2)c(z_3)\tilde{c}(\bar{z}_4)\tilde{c}(\bar{z}_5)\tilde{c}(\bar{z}_6)\rangle_{S_2} \,. \tag{6.3.1}$$

Up to normalization (i.e., a functional determinant), the result (??) is the zero-mode determinant :

$$\det \mathsf{C}^a_{0j}(\sigma_i) \,. \tag{6.3.2}$$

Six CKVs were found in section 6.1 . In the complex basis, they are :

$$\mathsf{C}^z = 1, z, z^2 \,, \quad \mathsf{C}^{\bar{z}} = 0 \,, \tag{6.3.3a}$$

$$\mathsf{C}^z = 0 \,, \quad \mathsf{C}^{\bar{z}} = 1, \bar{z}, \bar{z}^2 \,. \tag{6.3.3b}$$

In this basis, the determinant splits into two 3×3 blocks, and the expectation value becomes :

$$C^{\mathrm{g}}_{S_2} \det \begin{vmatrix} 1 & 1 & 1 \\ z_1 & z_2 & z_3 \\ z_1^2 & z_2^2 & z_3^2 \end{vmatrix} \det \begin{vmatrix} 1 & 1 & 1 \\ \bar{z}_4 & \bar{z}_5 & \bar{z}_6 \\ \bar{z}_4^2 & \bar{z}_5^2 & \bar{z}_6^2 \end{vmatrix} = C^{\mathrm{g}}_{S_2} z_{12}z_{13}z_{23}\bar{z}_{45}\bar{z}_{46}\bar{z}_{56} \,. \tag{6.3.4}$$

The constant $C^{\mathrm{g}}_{S_2}$ contains the functional determinant and the finite-dimensional Jacobian (independent of position); the Jacobian arises because the basis (6.3.3) is not orthogonal .

For tree-level amplitudes, this is the only bc path integral we need, but for completeness, we give :

$$\left\langle \prod_{i=1}^{p+3} c(z_i) \prod_{j=1}^{p} b(z'_j) \cdot (\text{anti-holomorphic}) \right\rangle_{S_2} = C^{\mathrm{g}}_{S_2} \frac{z_{p+1,p+2}z_{p+1,p+3}z_{p+2,p+3}}{(z_1 - z'_1)\cdots(z_p - z'_p)} \cdot (\text{anti-holomorphic}) \pm \text{permutations} \,, \tag{6.3.5}$$

which is completed by contracting b and c . The anti-holomorphic part has the same form . We do not care about the overall sign; for the Faddeev-Popov determinant in any case, we will take the absolute value .

To provide another derivation using holomorphy, once again, let us first examine the conservation law . The conservation law can be obtained by inserting the ghost number contour integral $\oint_C \mathrm{d}z\, j(z)/2\pi i$ into the amplitude (6.3.5), where C encloses all vertex operators and the ghost number current is $j = -:bc:$. From the OPE of b and c, this merely counts the number of c fields minus the number of b fields, giving $n_c - n_b$ times the original amplitude . Now use the conformal transformation (2.5.17) to pull the contour back to the u coordinate chart :

$$\oint_C \frac{\mathrm{d}z}{2\pi i} j_z = -\oint_C \frac{\mathrm{d}u}{2\pi i} j_u + 3 \to 3 \,. \tag{6.3.6}$$

The offset $+3$ is because j is not a tensor, and its OPE with T contains a z^{-3} term . In the last step, we used holomorphy in the u coordinate chart . The non-zero amplitude thus has $n_c - n_b = 3$ and similarly $n_{\tilde{c}} - n_{\tilde{b}} = 3$, which is consistent with the Riemann-Roch theorem .

From the OPE, the ghost expectation value (6.3.1) is holomorphic with respect to each value, and it must be zero when two identical anticommuting fields come together . Therefore, it must be of the form :

$$z_{12}z_{13}z_{23}\bar{z}_{45}\bar{z}_{46}\bar{z}_{56} F(z_1,z_2,z_3)\tilde{F}(\bar{z}_4,\bar{z}_5,\bar{z}_6) \,, \tag{6.3.7}$$

where F and $\tilde{F}$ are holomorphic and anti-holomorphic functions of position, respectively . As $z_1 \to \infty$, it tends to $z_1^2 F$. However, c is a tensor with weight -1, so the amplitude cannot be greater than z_1^2 at infinity . Thus $F(z_i)$ must be independent of z_1, and similarly for z_2 and z_3 . The discussion for $\tilde{F}$ is similar, and we again obtain the result (6.3.4) . For the general case (6.3.5), the same discussion gives the result :

$$C^{\mathrm{g}}_{S_2} \prod_{i<i'}^{p+3} z_{ii'} \prod_{j<j'}^{p} z'_{jj'} \prod_{i=1}^{p+3} \prod_{j=1}^{p} (z_i - z'_j)^{-1} \cdot (\text{anti-holomorphic}) \,, \tag{6.3.8}$$

which has the correct poles, zeros, and correct behavior at infinity . Summing the permutations in (6.3.5) obviously gives this term . We considered only $\lambda = 2$, which is related to the ghosts of the bosonic string, but both methods can be readily generalized to any λ .

The Disk

The simplest way to obtain the bc amplitude on the disk is the doubling trick . As in (2.7.30), we can represent the holomorphic and anti-holomorphic fields on the upper half-plane using holomorphic fields on the entire plane . Using :

$$\tilde{b}(\bar{z}) = b(z'), \quad \tilde{c}(\bar{z}) = c(z'), \quad z' = \bar{z}, \;\; \mathrm{Im}\, z > 0 \,. \tag{6.3.9}$$

The expectation value of holomorphic fields is obtained as for the sphere :

$$\langle c(z_1)c(z_2)c(z_3)\rangle_{D_2} = C^{\mathrm{g}}_{D_2} z_{12} z_{13} z_{23} \,. \tag{6.3.10}$$

For example :

$$\begin{aligned} \langle c(z_1)c(z_2)\tilde{c}(\bar{z}_3)\rangle_{D_2} &= \langle c(z_1)c(z_2)c(z'_3)\rangle_{D_2} \\ &= C^{\mathrm{g}}_{D_2} z_{12}(z_1 - \bar{z}_3)(z_2 - \bar{z}_3) \,. \end{aligned} \tag{6.3.11}$$

More general correlation functions are obtained in the same way .

The Projective Plane

The doubling trick can be used again . The inversion $z' = -\bar{z}^{-1}$ implies :

$$\tilde{b}(\bar{z}) = \left(\frac{\partial z'}{\partial \bar{z}}\right)^2 b(z') = z'^4 b(z') \,, \quad \tilde{c}(\bar{z}) = \left(\frac{\partial z'}{\partial \bar{z}}\right)^{-1} c(z') = z'^{-2} c(z') \,. \tag{6.3.12}$$

Once again :

$$\langle c(z_1)c(z_2)c(z_3)\rangle_{RP_2} = C^{\mathrm{g}}_{RP_2} z_{12} z_{13} z_{23} \tag{6.3.13}$$

and :

$$\begin{aligned} \langle c(z_1)c(z_2)\tilde{c}(\bar{z}_3)\rangle_{RP_2} &= z_3'^{-2} \langle c(z_1)c(z_2)c(z'_3)\rangle_{RP_2} \\ &= C^{\mathrm{g}}_{RP_2} z_{12}(1 + z_1\bar{z}_3)(1 + z_2\bar{z}_3) \,. \end{aligned} \tag{6.3.14}$$

Remark. *The last step of the above equation :*

$$\begin{aligned} z_3'^{-2} \cdot C^{g}_{RP_2}(z_1 - z_2)(z_1 - z'_3)(z_2 - z'_3) &= \bar{z}_3^2 C^{g}_{RP_2}(z_1 - z_2)(z_1 + \bar{z}_3^{-1})(z_2 + \bar{z}_3^{-1}) \\ &= C^{g}_{RP_2}(z_1 - z_2)(1 + z_1\bar{z}_3)(1 + z_2\bar{z}_3) \,. \end{aligned}$$

6.4 The Veneziano amplitude

Open string amplitudes are simpler than closed string amplitudes, so we begin with them .

We represent the disk as the upper half-plane, so the boundary coordinate y is real . There are no moduli . After fixing the metric, the CKG $PSL(2,\mathbb{R})$ can be used to fix three vertex operators to arbitrary positions y_1, y_2, y_3 on the boundary, except that the group does not change the cyclic order of the vertex operators, so we sum over two orders . For three open string tachyons on the disk, the general expression for the string S-matrix (**??**) reduces to :

$$S_{D_2}(k_1;k_2;k_3) = g_o^3 \mathrm{e}^{-\lambda}\left\langle \circ c^1 \mathrm{e}^{ik_1\cdot X}(y_1)\circ \circ c^1\mathrm{e}^{ik_2\cdot X}(y_2)\circ\circ c^1\mathrm{e}^{ik_3\cdot X}(y_3)\circ\right\rangle_{D_2} + (k_2\leftrightarrow k_3), \quad (6.4.1)$$

where each fixed coordinate integral is replaced by the corresponding c ghost . Each open string introduces a factor g_o, which is the coupling constant of the open string . The factor $\mathrm{e}^{-\lambda}$ comes from the Euler number term in the action . Of course, g_o is related to e^{λ}, $g_o^2 \propto e^{\lambda}$, but we will determine the proportionality constant as we go further . The expectation value in (6.4.1) was found in the previous section (again, taking the absolute value of the ghost fields), giving :

$$\begin{aligned} S_{D_2}(k_1;k_2;k_3) &= ig_o^3 C_{D_2}(2\pi)^{26}\delta^{26}(\textstyle\sum_i k_i) \\ &\quad\times |y_{12}|^{1+2\alpha' k_1\cdot k_2}|y_{13}|^{1+2\alpha' k_1\cdot k_3}|y_{23}|^{1+2\alpha' k_2\cdot k_3} + (k_2\leftrightarrow k_3)\,, \end{aligned} \quad (6.4.2)$$

where $C_{D_2} = \mathrm{e}^{-\lambda}C^X_{D_2}C^{\mathrm{g}}_{D_2}$. Momentum conservation and the mass-shell condition $k_i^2 = 1/\alpha'$ imply :

$$2\alpha' k_1\cdot k_2 = \alpha'(k_3^2 - k_1^2 - k_2^2) = -1 \quad (6.4.3)$$

The same applies to other $k_i\cdot k_j$, so this reduces to :

$$S_{D_2}(k_1;k_2;k_3) = 2ig_o^3 C_{D_2}(2\pi)^{26}\delta^{26}(\textstyle\sum_i k_i) \quad (6.4.4)$$

This is independent of the choice of gauge y_i, which is a general property of the Faddeev-Popov scheme . Weyl invariance is key—if the vertex operators were not on-shell, the amplitude would depend on the choice of y_i .

We can also use the independence of y_i to give the Faddeev-Popov determinant without calculation . Using the mass-shell condition and momentum conservation, the expectation value of X^μ is proportional to $|y_{12}y_{13}y_{23}|^{-1}$, so the measure must be its reciprocal . When the number of vertex operators $n > 3$, the same measure still applies because in all these cases, three positions are fixed . The 4-tachyon amplitude is obtained in the same way :

$$\begin{aligned} &S_{D_2}(k_1;k_2;k_3;k_4) \\ &\quad= g_o^4\mathrm{e}^{-\lambda}\int_{-\infty}^{\infty}\mathrm{d}y_4\,\left\langle \prod_{i=1}^{3}\circ c^1(y_i)\mathrm{e}^{ik_i\cdot X(y_i)}\circ\circ\mathrm{e}^{ik_4\cdot X(y_4)}\circ\right\rangle + (k_2\leftrightarrow k_3) \\ &\quad= ig_o^4 C_{D_2}(2\pi)^{26}\delta^{26}(\textstyle\sum_i k_i)|y_{12}y_{13}y_{23}|\displaystyle\int_{-\infty}^{\infty}\mathrm{d}y_4\prod_{i<j}|y_{ij}|^{2\alpha' k_i\cdot k_j} + (k_2\leftrightarrow k_3)\,. \end{aligned} \quad (6.4.5)$$

After a variable substitution (Möbius transformation) for y_4, this is again independent of $y_{1,2,3}$. It is customary to take $y_1 = 0, y_2 = 1$, and $y_3\to\infty$. Amplitudes are usually expressed in terms of Mandelstam variables :

$$s = -(k_1+k_2)^2,\quad t = -(k_1+k_3)^2,\quad u = -(k_1+k_4)^2 \quad (6.4.6)$$

These are not independent: momentum conservation and the mass-shell condition imply :

$$s+t+u = \sum_i m_i^2 = -\frac{4}{\alpha'}\,. \quad (6.4.7)$$

Using $2\alpha' k_i \cdot k_j = -2 + \alpha'(k_i + k_j)^2$, the amplitude becomes :

$$S_{D_2}(k_1;k_2;k_3;k_4) = ig_o^4 C_{D_2}(2\pi)^{26}\delta^{26}(\textstyle\sum_i k_i) \times \left[\int_{-\infty}^{\infty} \mathrm{d}y_4\, |y_4|^{-\alpha' u-2}|1-y_4|^{-\alpha' t-2} + (t \to s)\right] . \quad (6.4.8)$$

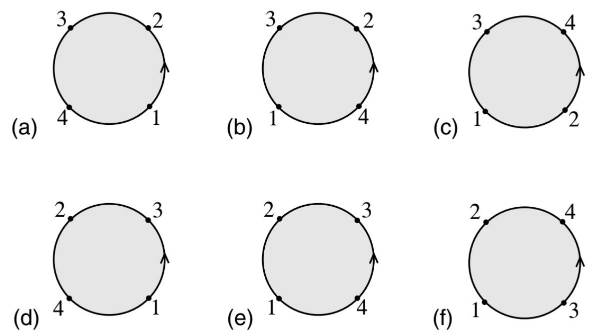


Figure 6.4.1: Six cyclically inequivalent orderings of four open string vertex operators on a circle . Except for a jump from ∞ to $-\infty$ at point 3, the coordinate y increases in the direction of the arrow .

The integral splits into three parts, $-\infty < y_4 < 0$, $0 < y_4 < 1$, and $1 < y_4 < \infty$. For these three ranges, the ordering of vertex operators is as shown in figure 6.4.1(a)(b)(c) . Möbius invariance can be used to transform these ranges to any other, so the contributions they give are equivalent to permutations of vertex operators . The $(t \leftrightarrow s)$ term gives figure 6.4.1(d)(e)(f) . Taken together :

$$S_{D_2}(k_1;k_2;k_3;k_4 = 2ig_o^4 C_{D_2}(2\pi)^{26}\delta^{26}(\textstyle\sum_i k_i)[I(s,t) + I(t,u) + I(u,s)] , \quad (6.4.9)$$

where :

$$I(s,t) = \int_0^1 \mathrm{d}y\, y^{-\alpha' s-2}(1-y)^{-\alpha' t-2} . \quad (6.4.10)$$

These three terms come from figures 6.4.1 (c)(f), (b)(d), and (a)(e), respectively .

If $\alpha' s < -1$ and $\alpha' t < -1$, the integral $I(s,t)$ converges . As $\alpha' s \to -1$, the integral diverges at $y = 0$. To study this divergence, take a neighborhood of $y = 0$ and approximate the integrand :

$$\begin{aligned} I(s,t) &= \int_0^r dy y^{-\alpha' s-2} + \text{terms analytic at } \alpha' s = -1 \\ &= -\frac{r^{-\alpha' s-1}}{\alpha' s+1} + \text{terms analytic at } \alpha' s = -1 \\ &= -\frac{1}{\alpha' s+1} + \text{terms analytic at } \alpha' s = -1 . \end{aligned} \quad (6.4.11)$$

In (6.4.11), we calculated the integral in the convergence region . We see that this divergence is a pole at $s = -1/\alpha'$, the mass square of the open string tachyon . The variable s is exactly the center-of-mass energy squared for the scattering $1 + 2 \to 3 + 4$, so this pole is a resonance produced by intermediate tachyon states . Once again, this is an artifact of the bosonic string, i.e., the lightest string state is a tachyon, which is irrelevant to the discussion . This pole is due to the process in figure 6.4.2(a) . There, tachyons 1 and 2 merge into a single tachyon, then split back into tachyons 3 and 4 .

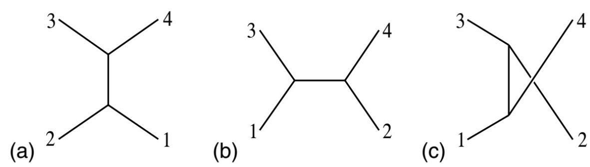


Figure 6.4.2: Processes giving poles in the (a) s-, (b) t-, and (c) u-channels .

Since the singularity at $\alpha' s = -1$ is only a pole, $I(s,t)$ can be analytically continued around this point and into the region $\alpha' s > -1$. The amplitude is defined via analytic continuation . The divergence of the amplitude at this pole is a significant physical feature: the resonance corresponding to an intermediate string state propagating over long spacetime distances . Divergence bypassing this pole is not; it is merely an artifact of this specific integral representation of the amplitude . Continuation poses no problem . In fact, we will see that every string divergence is of this basic form, so this analytic continuation removes all divergences—except, of course, the divergence of the pole itself . The pole is on the real axis, and we need to define it more precisely . For Minkowski processes, the correct ϵ treatment is :

$$\frac{1}{\alpha' s + 1} \equiv \frac{1}{\alpha' s + 1 + i\epsilon} \equiv \mathrm{P}\frac{1}{\alpha' s + 1} - i\pi\delta(\alpha' s + 1) \,, \tag{6.4.12}$$

where P represents the principal value . Unitarity (which we will systematically develop in chapter **??**) requires this pole to appear, and determines the coefficient from the scattering amplitude of two tachyons to one string :

$$S_{D_2}(k_1; k_2; k_3; k_4) = i \int \frac{\mathrm{d}^{26}k}{(2\pi)^{26}} \frac{S_{D_2}(k_1; k_2; k) S_{D_2}(-k; k_3; k_4)}{-k^2 + \alpha'^{-1} + i\epsilon} + \text{ terms analytic at } k^2 = \frac{1}{\alpha'} \,. \tag{6.4.13}$$

Gathering factors in the 4-tachyon amplitude, including the equivalent contribution of $I(u,s)$ to this pole, and using the 3-tachyon result (6.4.4), condition (6.4.13) gives :

$$C_{D_2} = \mathrm{e}^{-\lambda} C^X_{D_2} C^{\mathrm{g}}_{D_2} = \frac{1}{\alpha' g_{\mathrm{o}}^2} \,. \tag{6.4.14}$$

Remark. *Singular term given by* (6.4.13) *:*

$$i \cdot (2ig_{\mathrm{o}}^3 C_{D_2})^2 (2\pi)^{26} \alpha' \delta^{2b}(\textstyle\sum_i k_i) \frac{1}{\alpha' s + 1} \,,$$

Singular term given by (6.4.9) *:*

$$2 \cdot 2ig_{\mathrm{o}}^4 C_{D_2} (2\pi)^{26} \delta^{26}(\textstyle\sum_i k_i) \frac{-1}{\alpha' s + 1} \,.$$

The two together give (6.4.14) *.*

Then the 3-tachyon amplitude is :

$$S_{D_2}(k_1; k_2; k_3) = \frac{2ig_{\mathrm{o}}}{\alpha'} (2\pi)^{26} \delta^{26}(\textstyle\sum_i k_i) \,. \tag{6.4.15}$$

Various functional determinants are dropped . Using unitarity, all normalizations can be expressed in terms of the coupling g_o appearing in vertex operators . In fact, the determinants can also be calculated through careful regularization and renormalization, and the relative normalization of different topologies is consistent with results derived from unitarity .

Continuing analytically around the pole, we encounter further singularities . Expanding the integrand at $y = 0$:

$$I(s,t) = \int_0^r \mathrm{d}y \left[y^{-\alpha' s - 2} + (\alpha' t + 2) y^{-\alpha' s - 1} + \cdots \right] . \tag{6.4.16}$$

The second term gives the pole at $\alpha' s = 0$:

$$I(s,t) = \frac{u-t}{2s} + \text{terms analytic at } \alpha' s = 0 \, . \tag{6.4.17}$$

From subsequent terms in the Taylor expansion, poles of the amplitude are at :

$$\alpha' s = -1, 0, 1, 2, \ldots \, . \tag{6.4.18}$$

These are exactly the positions of open string states . The integral $I(s,t)$ has poles in variable t at the same positions (6.4.18), which come from endpoint $y = 1$. This is due to the process in figure 6.4.2(b) . Other two terms in the amplitude (6.4.9) give poles in the s-, t-, and u-channels (figure 6.4.2c) . Because the residue of (6.4.17) at $s = 0$ is odd with respect to $u - t$, this pole will actually cancel with the pole in $I(s,u)$. Singularities at even multiples of $1/\alpha'$ all exhibit this behavior . But for the more general open string theory introduced in the next section, this no longer holds . Defining the Euler beta function :

$$B(a,b) = \int_0^1 \mathrm{d}y \, y^{a-1} (1-y)^{b-1} \, , \tag{6.4.19}$$

makes :

$$I(s,t) = B(-\alpha_\mathrm{o}(s), -\alpha_\mathrm{o}(t)), \qquad \alpha_\mathrm{o}(x) = 1 + \alpha' x \, . \tag{6.4.20}$$

This can be expressed as gamma functions . For fixed w, define $y = v/w$, which gives :

$$w^{a+b-1} B(a,b) = \int_0^w \mathrm{d}v \, v^{a-1} (w-v)^{b-1} \, . \tag{6.4.21}$$

Multiplying both sides by e^{-w} and integrating $\int_0^\infty \mathrm{d}w$, reassembling gives :

$$\begin{aligned} \Gamma(a+b) B(a,b) &= \int_0^\infty \mathrm{d}v \, v^{a-1} \mathrm{e}^{-v} \int_0^\infty \mathrm{d}(w-v) \, (w-v)^{b-1} \mathrm{e}^{-(w-v)} \\ &= \Gamma(a) \Gamma(b) \, . \end{aligned} \tag{6.4.22}$$

Then the 4-tachyon amplitude is :

$$\begin{aligned} S_{D_2}(k_1; k_2; k_3; k_4) &= \frac{2 i g_\mathrm{o}^2}{\alpha'} (2\pi)^{26} \delta^{26}(\textstyle\sum_i k_i) \\ &\times \Big[B(-\alpha_\mathrm{o}(s), -\alpha_\mathrm{o}(t)) + B(-\alpha_\mathrm{o}(s), -\alpha_\mathrm{o}(u)) + B(-\alpha_\mathrm{o}(t), -\alpha_\mathrm{o}(u)) \Big] \, , \end{aligned} \tag{6.4.23}$$

where :

$$B(-\alpha_\mathrm{o}(x), -\alpha_\mathrm{o}(y)) = \frac{\Gamma(-\alpha' x - 1) \Gamma(-\alpha' y - 1)}{\Gamma(-\alpha' x - \alpha' y - 2)} \, . \tag{6.4.24}$$

This is the Veneziano amplitude, originally written down to model certain features of strong interactions .

The high-energy behavior of the Veneziano amplitude is important . There are two interesting regions: the Regge limit ,

$$s \to \infty, \quad t \text{ fixed} \, , \tag{6.4.25}$$

and the hard scattering limit :

$$s \to \infty, \quad t/s \text{ fixed} \, . \tag{6.4.26}$$

If we consider the scattering process $1+2 \to 3+4$ (so that k_1^0 and k_2^0 are positive, and k_3^0 and k_4^0 are negative) , then in the center-of-mass frame of 1-2 :

$$s = E^2, \quad t = (4m^2 - E^2)\sin^2\frac{\theta}{2}, \quad u = (4m^2 - E^2)\cos^2\frac{\theta}{2}\,, \tag{6.4.27}$$

where E is the center-of-mass frame energy, and θ is the angle between particles 1 and 3 . The Regge limit is high energy at a small angle, while the hard scattering limit is high energy at a fixed angle . Using Stirling's approximation, $\Gamma(x+1) \approx x^x \mathrm{e}^{-x}(2\pi x)^{1/2}$, behavior in the Regge region is :

$$S_{D_2}(k_1; k_2; k_3; k_4) \propto s^{\alpha_\mathrm{o}(t)}\Gamma(-\alpha_0(t))\,, \tag{6.4.28}$$

where $\alpha_\mathrm{o}(t) = \alpha' t + 1$, i.e., the amplitude varies as a power of s, where the power depends on t . This is Regge behavior . At the poles of the gamma function, the amplitude is an integer power of s, corresponding to exchanging a string with integer spin $\alpha_o(t)$.

In the hard scattering limit :

$$S_{D_2}(k_1; k_2; k_3; k_4) \approx \exp\left[-\alpha'(s \ln s\alpha' + t \ln t\alpha' + u \ln u\alpha')\right] = \exp\left[-\alpha' s f(\theta)\right]\,, \tag{6.4.29}$$

where :

$$f(\theta) \approx -\sin^2\frac{\theta}{2}\ln\sin^2\frac{\theta}{2} - \cos^2\frac{\theta}{2}\ln\cos^2\frac{\theta}{2} \tag{6.4.30}$$

is positive . (6.4.29) is noteworthy . High-energy, fixed-angle scattering probes the internal structure of the scattered objects . Rutherford discovered the atomic nucleus using hard α atom scattering . Hard electron-nucleon scattering at SLAC revealed the quark components of nucleons . In quantum field theory, hard scattering amplitudes decay according to a power law in s . Even for a composite object like a nucleon, if its components are point-like, its amplitude follows a power law . The exponential decay (6.4.29) is much softer, reflecting an object with a size of $\alpha^{1/2}$, which is what we expect .

We started from 3-point amplitudes, skipping zero-, one-, and two-point amplitudes . We will discuss these amplitudes and their interpretations in section 6.6 .

6.5 Chan-Paton factors and gauge interactions

In this section, we will examine the interactions of massless vector states of the open string . To make this discussion more interesting, we first generalize the open string theory .

At the end of chapter 3, we introduced a broad class of bosonic string theories, but in our preliminary examination of interactions, we focused on the simple case of 26-dimensional flat spacetime . We can think about it in terms of symmetry: this theory has a massive 26-dimensional Poincaré invariance . In the closed bosonic string, this is the only theory with this symmetry . The proof is summarized as follows: worldsheet Noether currents of spacetime translations have components with weights (1,0) and (0,1) . By the discussion given in section 2.9, these currents are holomorphic with respect to z or $\bar{z}$. In calculations in this chapter, this is sufficient to determine all expectation values .

However, there exists a generalization in open string theory . Open strings have boundaries, endpoints . In quantum systems with distinguishable endpoints, it is natural to contain degrees of freedom at these points in addition to fields propagating in the bulk . At each endpoint of the open string, we add new degrees of freedom called Chan-Paton degrees of freedom, which can be one of n states . Then the basis of string states is :

$$|N; k; ij\rangle\,, \tag{6.5.1}$$

where i and j label the states of the left and right endpoints, ranging from 1 to n . The energy-momentum tensor is defined as usual, with no dependence on the new degrees of freedom

. Thus, conformal invariance is automatic . Since the Chan-Paton degrees of freedom are invariant, Poincaré invariance is also automatic . Although the worldsheet dynamics of these new degrees of freedom are trivial, they have profound implications for spacetime physics .

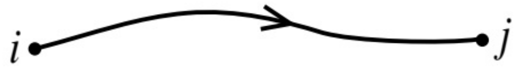


Figure 6.5.1: An open string with Chan-Paton degrees of freedom .

In the string theory of strong interactions, the motivation for introducing Chan-Paton factors was to introduce $SU(3)$ flavor quantum numbers: endpoints are like quarks and anti-quarks connected by a color-current tube . We now introduce it in a broader framework: consider all possible symmetries . In chapter 8, we will give a new interpretation of Chan-Paton degrees of freedom, and in chapter 14, we will add a possible further improvement . Now there are n^2 scalar tachyons, n^2 massless vector bosons, and so on . Introduce n^2 Hermitian matrices λ^a_{ij} and normalize as :

$$\mathrm{Tr}(\lambda^a \lambda^b) = \delta^{ab} \,, \tag{6.5.2}$$

These matrices are a complete basis for the states of the two endpoints . They are representation matrices of $U(n)$, so one can guess that the massless vector bosons are associated with $U(n)$ symmetry; we will soon see that this is indeed the case .

Define the basis :

$$|N;k;a\rangle = \sum_{i,j=1}^{n} |N;k;ij\rangle \lambda^a_{ij} \,. \tag{6.5.3}$$

Now consider the 4-tachyon amplitude shown in figure 6.4.1(a), in which vertex operators are arranged in the cyclic order 1234 . Because Chan-Paton degrees of freedom do not appear in the Hamiltonian, their states cannot evolve between vertex operators: the state of the right endpoint of tachyon 1 must be the same as the state of the left endpoint of tachyon 2, and so on . Therefore, the amplitude in figure 6.4.1(a) will contain the factor :

$$\mathrm{Tr}(\lambda^{a_1}\lambda^{a_2}\lambda^{a_3}\lambda^{a_4}) \,, \tag{6.5.4}$$

which originates from the overlap of the Chan-Paton wavefunctions of each tachyon . This rule can be generalized to arbitrary amplitudes: each vertex operator now contains the Chan-Paton factor λ^a_{ij} from the endpoint wavefunction, and the amplitude for each worldsheet is multiplied by the trace of Chan-Paton factors along the boundary . The 3-tachyon amplitude becomes :

$$S_{D_2}(k_1,a_1;k_2,a_2;k_3,a_3) = \frac{ig_{\mathrm{o}}}{\alpha'}(2\pi)^{26}\delta^{26}(\textstyle\sum_i k_i)\,\mathrm{Tr}(\lambda^{a_1}\lambda^{a_2}\lambda^{a_3} + \lambda^{a_1}\lambda^{a_3}\lambda^{a_2}) \,, \tag{6.5.5}$$

The two cyclic orders now have different Chan-Paton traces . The 4-tachyon amplitude is :

$$\begin{aligned} S_{D_2}(k_1,a_1;k_2,a_2;k_3,a_3;k_4,a_4) &= \frac{ig_{\mathrm{o}}^2}{\alpha'}(2\pi)^{26}\delta^{26}(\textstyle\sum_i k_i) \\ \times \Big[&\mathrm{Tr}(\lambda^{a_1}\lambda^{a_2}\lambda^{a_4}\lambda^{a_3} + \lambda^{a_1}\lambda^{a_3}\lambda^{a_4}\lambda^{a_2})B(-\alpha_{\mathrm{o}}(s),-\alpha_{\mathrm{o}}(t)) \\ &+ \mathrm{Tr}(\lambda^{a_1}\lambda^{a_3}\lambda^{a_2}\lambda^{a_4} + \lambda^{a_1}\lambda^{a_4}\lambda^{a_2}\lambda^{a_3})B(-\alpha_{\mathrm{o}}(t),-\alpha_{\mathrm{o}}(u)) \\ &+ \mathrm{Tr}(\lambda^{a_1}\lambda^{a_2}\lambda^{a_3}\lambda^{a_4} + \lambda^{a_1}\lambda^{a_4}\lambda^{a_3}\lambda^{a_2})B(-\alpha_{\mathrm{o}}(s),-\alpha_{\mathrm{o}}(u))\Big] \,. \end{aligned} \tag{6.5.6}$$

Again considering the unitarity relation (6.4.13), the pole at $s = -1/\alpha'$ gives the left side the factor :

$$\frac{1}{4}\,\mathrm{Tr}\Big(\{\lambda^{a_1},\lambda^{a_2}\}\{\lambda^{a_3},\lambda^{a_4}\}\Big)\,, \tag{6.5.7}$$

while the right side gains the factor :

$$\frac{1}{4}\sum_a \mathrm{Tr}\Big(\{\lambda^{a_1},\lambda^{a_2}\}\lambda^a\Big)\,\mathrm{Tr}\Big(\{\lambda^{a_3},\lambda^{a_4}\}\lambda^a\Big)\,, \tag{6.5.8}$$

namely summing over the Chan-Paton wavefunctions of the intermediate states . The completeness and normalization (6.5.2) of λ^a_{ij} imply that for any A and B, one has :

$$\mathrm{Tr}(A\lambda^a)\,\mathrm{Tr}(B\lambda^a) = \mathrm{Tr}(AB)\,, \tag{6.5.9}$$

thus the amplitude remains unitary .

Proof. Let $\mathrm{Tr}(A\lambda^a) = \langle A|\lambda^a\rangle = \langle\lambda^a|A\rangle$, then :

$$\mathrm{Tr}(A\lambda^a)\,\mathrm{Tr}(B\lambda^a) = \langle A|\lambda^a\rangle\langle\lambda^\alpha|B\rangle = \langle A|B\rangle\,,$$

where $\sum_a |\lambda^a\rangle\langle\lambda^a| = 1$ comes from $\langle\lambda^a|\lambda^b\rangle = \delta^{ab}$. □

Gauge Interactions

The amplitude of a gauge boson with two tachyons is :

$$\begin{aligned} S_{D_2}(k_1,a_1,e_1;k_2,a_2;k_3,a_3) &= -ig'_o g_o^2 \mathrm{e}^{-\lambda} e_{1\mu} \\ &\times \Big\langle \circ c^1 \dot{X}^\mu \mathrm{e}^{ik_1\cdot X}(y_1)\circ\circ c^1\mathrm{e}^{ik_2\cdot X}(y_2)\circ\circ c^1\mathrm{e}^{ik_3\cdot X}(y_3)\circ\Big\rangle_{D_2}\mathrm{Tr}(\lambda^{a_1}\lambda^{a_2}\lambda^{a_3}) \\ &+ (k_2,a_2)\leftrightarrow(k_3,a_3)\,. \end{aligned} \tag{6.5.10}$$

We use the bosonic vertex operator (**??**), but temporarily allow an independent normalization constant g'_o . Using the results from section 6.2, the path integral of X is :

$$\begin{aligned} \Big\langle\circ\dot{X}^\mu\mathrm{e}^{ik_1\cdot X}(y_1)\circ\circ\mathrm{e}^{ik_2\cdot X}(y_2)\circ\circ\mathrm{e}^{ik_3\cdot X}(y_3)\circ\Big\rangle_{D_2} &= -2i\alpha'\left(\frac{k_2^\mu}{y_{12}}+\frac{k_3^\mu}{y_{13}}\right) \\ \times\, iC^X_{D_2}(2\pi)^{26}\delta^{26}(\textstyle\sum_i k_i)|y_{12}|^{2\alpha' k_1\cdot k_2}|y_{13}|^{2\alpha' k_1\cdot k_3}|y_{23}|^{2\alpha' k_2\cdot k_3}\,. \end{aligned} \tag{6.5.11}$$

Using momentum conservation, mass-shell condition, and physical state condition $k_1\cdot e_1 = 0$, the amplitude becomes :

$$\begin{aligned} &S_{D_2}(k_1,a_1,e_1;k_2,a_2;k_3,a_3) \\ &= -ig'_o e_1\cdot k_{23}(2\pi)^{26}\delta^{26}(\textstyle\sum_i k_i)\,\mathrm{Tr}\Big(\lambda^{a_1}[\lambda^{a_2},\lambda^{a_3}]\Big)\,, \end{aligned} \tag{6.5.12}$$

where $k_{ij}\equiv k_i - k_j$. This is again independent of the position of vertex operators .

The $s = 0$ pole in the 4-tachyon amplitude is no longer zero . The term that was canceled now has Chan-Paton factors in a different order, so the pole is proportional to :

$$\mathrm{Tr}\Big([\lambda^{a_1},\lambda^{a_2}][\lambda^{a_3},\lambda^{a_4}]\Big)\,. \tag{6.5.13}$$

Relating the coefficient of this pole to the amplitude (6.5.12) through unitarity gives :

$$g'_o = (2\alpha')^{-1/2} g_o\,. \tag{6.5.14}$$

This is the same as the relative normalization derived from the state-operator map: there is only one independent coupling constant . For the coupling of 3 gauge bosons, a similar calculation gives :

$$\begin{aligned}
&S_{D_2}(k_1, a_1, e_1; k_2, a_2, e_2; k_3, a_3, e_3) \\
&\quad = ig'_o(2\pi)^{26}\delta^{26}(\textstyle\sum_i k_i)\Big(e_1 \cdot k_{23} e_2 \cdot e_3 + e_2 \cdot k_{31} e_3 \cdot e_1 + e_3 \cdot k_{12} e_1 \cdot e_2 \\
&\qquad + \frac{\alpha'}{2} e_1 \cdot k_{23} e_2 \cdot k_{31} e_3 \cdot k_{12}\Big) \operatorname{Tr}\Big(\lambda^{a_1}[\lambda^{a_2}, \lambda^{a_3}]\Big) . \qquad (6.5.15)
\end{aligned}$$

To first order in momentum, the amplitudes we found can be re-produced from the following spacetime action :

$$\boldsymbol{S} = \frac{1}{g_o'^2} \int \mathrm{d}^{26}x \left[-\frac{1}{2} \operatorname{Tr}(D_\mu \varphi D^\mu \varphi) + \frac{1}{2\alpha'} \operatorname{Tr}(\varphi^2) + \frac{2^{1/2}}{3\alpha'^{1/2}} \operatorname{Tr}(\varphi^3) - \frac{1}{4} \operatorname{Tr}(F_{\mu\nu} F^{\mu\nu}) \right], \quad (6.5.16)$$

where the tachyon field φ and the Yang-Mills vector potential A_μ are written as $n \times n$ matrices, i.e., $A_\mu = A_\mu^a \lambda^a$. Furthermore, $D_\mu \varphi = \partial_\mu \varphi - i[A_\mu, \varphi]$, and $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu - i[A_\mu, A_\nu]$.

This is the action for a $U(n)$ gauge field coupled to a scalar field in the adjoint representation . The addition of the Chan-Paton factor is exactly what is needed for the gauge-invariant representation . Since decoupling of non-physical states is guaranteed in string perturbation theory, gauge invariance is automatic, and we will study it further .

At momentum k much smaller than the string scale, the only open string state is the massless gauge boson . As discussed in section **??**, in this limit, the physics should reduce to an effective field theory of massless states . Therefore, introducing the tachyon into the action (6.5.16) must be illogical in some places; we did so for heuristics . But now let us focus on the gauge boson . The form of the 4-gauge boson amplitude is similar to the Veneziano amplitude, but with extra structure from the polarization tensor . Expanding it as a power series in $\alpha' k^2$, the first term survives in the zero-slope limit; it is the sum of pole terms in s, t, u plus a constant . Consistency will guarantee that it is exactly the same as the 4-gauge boson amplitude obtained from the Yang-Mills Lagrangian $F_{\mu\nu}F^{\mu\nu}$ in field theory . Note, however, the term at order $\alpha' k^3$ in the 3-gauge boson amplitude (6.5.15) . This implies higher-order derivative terms in the Lagrangian :

$$-\frac{2i\alpha'}{3g_o'^2} \operatorname{Tr}(F_\mu{}^\nu F_\nu{}^\omega F_\omega{}^\mu) . \qquad (6.5.17)$$

Similarly, expanding the 4-point amplitude will reveal an infinite sum of high-order interaction terms (besides (6.5.17), they indeed have no kinematic contribution to the 3-gauge boson amplitude) . In low-energy regions, string loop amplitudes also reduce to loops obtained from the effective Lagrangian . By the general logic of effective Lagrangians, higher-order derivative terms are not very important at low energies ; they become important at the scale $\alpha' k^2 \approx 1$, which is where new physics (massive string states) appears and the effective action is no longer applicable .

If we have a cutoff, why do we still need renormalization? Renormalization still has meaning . And in fact, this is its true interpretation: it means that low-energy physics is independent of the details of high-energy physics, except for the parameters in the effective Lagrangian . This is bittersweet: it means we can use ordinary quantum field theory to make predictions at accelerator energy levels without knowing the theory at the Planck scale, but it also means we cannot use physics at particle accelerators to probe the theory at the Planck scale .

String spectra and amplitudes have an obvious global $U(n)$ symmetry ,

$$\lambda^a \to U \lambda^a U^\dagger , \qquad (6.5.18)$$

which preserves the Chan-Paton trace and the norm of the states . From the details of the amplitudes, we see that this is actually local symmetry of spacetime . We will see that the elevation from global worldsheet symmetry to local spacetime symmetry is a very common phenomenon in string theory .

All open string states transform according to the $n \times n$ adjoint representation under $U(n)$ symmetry . Incidentally, $U(n)$ is not a simple Lie algebra: $U(n) = SU(n) \times U(1)$. The $U(1)$ gauge boson, $\lambda_{ij} = \delta_{ij}/n^{1/2}$, decouples from amplitudes (6.5.12) and (6.5.15) . The adjoint representation of $U(1)$ is trivial, so all string states are neutral under $U(1)$ symmetry .

Non-orientable Strings

The extension to non-orientable strings is interesting . First consider a theory without Chan-Paton factors . Besides Möbius invariance, the X^μ CFT on a sphere or disk is invariant under reflection symmetry in the direction of the open string $\sigma^1 \to \pi - \sigma^1$ or on a closed string $\sigma^1 \to 2\pi - \sigma^1$. Worldsheet parity is generated by the operator Ω . From the mode expansion, in open strings :

$$\Omega \alpha_n^\mu \Omega^{-1} = (-1)^n \alpha_n^\mu \,, \tag{6.5.19}$$

and in closed strings :

$$\Omega \alpha_n^\mu \Omega^{-1} = \tilde{\alpha}_n^\mu \,. \tag{6.5.20}$$

This symmetry also extends to ghost fields . But for concentration, we do not discuss them here; their contribution to tree-level diagrams is merely a fixed factor .

Regardless of whether it is an open or closed string, tachyon vertex operators are even under worldsheet parity (this is obvious for ghost-free integrated vertex operators; fixed operators must transform in the same way), which determines the sign of the operator . Then all states can be categorized according to their parity eigenvalues $\omega = \pm 1$. The relation (6.5.19) in open strings implies :

$$\Omega |N; k\rangle = \omega_N |N; k\rangle, \quad \omega_N = (-1)^{1+\alpha' m^2} \,. \tag{6.5.21}$$

Worldsheet parity is multiplicative conservation . For example, the 3-tachyon amplitude is non-zero, consistent with $(+1)^3 = 1$. On the other hand, massless vectors have $\omega = -1$, so we can expect the vector-tachyon-tachyon amplitude (6.5.10) and the 3-vector amplitude (6.5.15) to be zero in the absence of Chan-Paton factors, and this is indeed the case (λ^a is replaced by 1, and the commutator is zero) . Different cyclic orders are also related to each other via worldsheet parity and thus cancel out .

Given a consistent oriented string theory, by restricting the spectrum to states with $\omega = +1$, we give a new non-orientable string theory . States where $\alpha' m^2$ is odd are retained, while states where $\alpha' m^2$ is even, including the photon, are forbidden . ω conservation guarantees that if all external states have $\omega = +1$, then intermediate states of tree-level diagrams also have $\omega = +1$. Therefore, at least at the tree level, the unitarity of non-orientable strings follows from the unitarity of oriented strings .

The primary interest in non-orientable theory lies in the treatment of Chan-Paton factors . Since we can identify opposite endpoints of an open string, worldsheet parity must reverse them :

$$\Omega |N; k; ij\rangle = \omega_N |N; k; ji\rangle \,. \tag{6.5.22}$$

Once again, this is a symmetry of all amplitudes in the oriented theory . To form a non-orientable theory, we again constrain the spectrum to worldsheet parity eigenvalues $\omega = +1$. Take a set of basis for λ^a_{ij} such that each matrix is either symmetric $s^a = +1$ or antisymmetric $s^a = -1$. Then :

$$\Omega |N; k; a\rangle = \omega_N s^a |N; k; a\rangle \,. \tag{6.5.23}$$

The worldsheet parity eigenvalue is $\omega = \omega_N s^a$, so the non-orientable spectrum is :

$$\text{even multiples of } \alpha' m^2 : \qquad \text{antisymmetric } \lambda^a \,, \tag{6.5.24a}$$

$$\text{odd multiples of } \alpha' m^2 \ : \qquad \text{symmetric } \lambda^a \, . \tag{6.5.24b}$$

For massless gauge bosons, Chan-Paton factors are $n \times n$ antisymmetric matrices, so the gauge group is $SO(n)$. States at even mass levels transform according to the adjoint representation of the orthogonal group $SO(n)$, while states at odd mass levels transform according to a traceless symmetric tensor plus a singlet representation . The oriented theory has a large class of direction-reversing symmetries obtained as a combination of Ω and $U(n)$ rotations :

$$\Omega_\gamma |N;k;ij\rangle = \omega_N \gamma_{jj'} \left|N;k;j'i'\right\rangle \gamma^{-1}_{i'i} \, . \tag{6.5.25}$$

By restricting the spectrum to $\omega_y = +1$, we can construct more general non-orientable theories . This is again compatible with interactions; applying Ω_γ twice gives :

$$\Omega_\gamma^2 |N;k;ij\rangle = [(\gamma^T)^{-1}\gamma]_{ii'} |N;k;i'j'\rangle (\gamma^{-1}\gamma^T)_{j'j} \, . \tag{6.5.26}$$

We assume $\Omega_\gamma^2 = 1$ for reasons explained below . This implies :

$$\gamma^T = \pm\gamma \, . \tag{6.5.27}$$

That is, γ is symmetric or antisymmetric .

The Chan-Paton basis transformation is :

$$|N;k;ij\rangle' = U^{-1}_{ii'} |N;k;i'j'\rangle U_{j'j} \, , \tag{6.5.28}$$

which changes γ to :

$$\gamma' = U^T \gamma U \, . \tag{6.5.29}$$

In the symmetric case, one can always find a basis such that $\gamma = 1$, which gives the case already examined above . In the antisymmetric case, there exists a basis such that :

$$\gamma = M \equiv i \begin{bmatrix} 0 & I \\ -I & 0 \end{bmatrix} \, . \tag{6.5.30}$$

Here I is a $k \times k$ identity matrix; since γ is an invertible antisymmetric matrix, $n = 2k$ must be even . We take a basis for the Chan-Paton wavefunctions such that $M(\lambda^a)^T M = s^{a'} \lambda^a$, where $s^{a'} = \pm 1$. Then the worldsheet eigenvalue is $\omega_\gamma = \omega_N s^{a'}$, and the non-orientable spectrum is :

$$\text{even multiples of } \alpha' m^2 \ : \qquad M(\lambda^a)^T M = -\lambda^a \, , \tag{6.5.31a}$$

$$\text{odd multiples of } \alpha' m^2 \ : \qquad M(\lambda^a)^T M = +\lambda^a \, . \tag{6.5.31b}$$

At even mass levels, including gauge bosons, this defines the adjoint representation of the symplectic group $Sp(k)$.

To construct a non-orientable theory, we must have $\Omega_\gamma^2 = 1$ for the following reasons . Since $\Omega^2 = 1$, Ω_γ^2 must only act on the Chan-Paton factors . In fact, from (6.5.26), its action on the Chan-Paton wavefunction is :

$$\lambda \to (\gamma^T)^{-1} \gamma \lambda \gamma^{-1} \gamma^T = \lambda \, , \tag{6.5.32}$$

where, since all states are invariant under Ω_γ, the last equal sign must hold in non-orientable theory . Now we assert that the allowed Chan-Paton wavefunctions must constitute a complete basis . The key is that two open strings can exchange endpoints through the split-aggregate interaction in figure 3.4(c) . In this way, a complete basis, namely any Chan-Paton state $|ij\rangle$, can be reached . By Schur's Lemma, if (6.5.32) holds for the complete basis, then $\gamma^{-1}\gamma^T = 1$, hence $\Omega_\gamma^2 = 1$. By taking different sets of λ^a, we can try to obtain different gauge theories . In fact, the oriented $U(n)$ theory and non-orientable $SO(n)$ and $Sp(k)$ theories constructed above are the only possibilities . The extension of the completeness discussion in (6.5.7) – (6.5.9) shows they are the most general solutions for non-orientable theories . In particular, exceptional Lie algebras cannot be obtained from Chan-Paton factors . In closed string theory, there are other mechanisms giving gauge bosons and allowing other groups .

6.6 Closed string tree amplitudes

The discussion of closed string amplitudes is similar to the above . The amplitude of 3 closed string tachyons is :

$$S_{S_2}(k_1;k_2;k_3) = g_c^3 e^{-2\lambda} \left\langle \prod_{i=1}^{3} :\tilde{c}ce^{ik_i\cdot X}(z_i,\bar{z}_i): \right\rangle_{S_2} . \tag{6.6.1}$$

In this case, the CKG $PSL(2,\mathbb{C})$ (Möbius group) can fix three vertex operators at arbitrary positions $z_{1,2,3}$. Taking the expectation value from section 6.2, the result is again independent of the vertex operators :

$$S_{S_2}(k_1;k_2;k_3) = ig_c^3 C_{S_2}(2\pi)^{26}\delta^{26}(\textstyle\sum_i k_i) , \tag{6.6.2}$$

where $C_{S_2} = e^{-2\lambda} C_{S_2}^X C_{S_2}^g$.

For 4 closed string tachyons :

$$S_{S_2}(k_1;k_2;k_3;k_4) = g_c^4 e^{-2\lambda} \int_{\mathbb{C}} d^2 z_4 \left\langle \prod_{i=1}^{3} :\tilde{c}ce^{ik_i\cdot X}(z_i,\bar{z}_i)::e^{ik_4\cdot X}(z_4,\bar{z}_4): \right\rangle_{S_2} , \tag{6.6.3}$$

where the integral range is the entire complex plane $\mathbb{C}$. Calculating the expectation value and letting $z_1 = 0, z_2 = 1, z_3 = \infty$, this becomes :

$$S_{S_2}(k_1;k_2;k_3;k_4) = ig_c^4 C_{S_2}(2\pi)^{26}\delta^{26}(\textstyle\sum_i k_i) J(s,t,u) , \tag{6.6.4}$$

where :

$$J(s,t,u) = \int_{\mathbb{C}} d^2 z_4 \, |z_4|^{-\alpha' u/2-4} |1-z_4|^{-\alpha' t/2-4} . \tag{6.6.5}$$

Here $s+t+u = -16/\alpha'$; we have labeled here the correlation of J with all 3 variables to emphasize their symmetry . When $s,t,u < -4/\alpha'$, the amplitude converges . As $z_4 \to 0$, it has a pole with respect to u; as $z_4 \to 1$, it has a pole with respect to t; as $z_4 \to \infty$, it has a pole with respect to s . These poles are at :

$$\alpha' s\,,\ \alpha' t\,,\ \alpha' u = -4, 0, 4, 8, \ldots , \tag{6.6.6}$$

which are the mass squared of closed string states . The pole at $\alpha' s = -4$ is :

$$ig_c^4 C_{S_2} \int_{|z_4|>1/\epsilon} d^2 z_4 \, |z_4|^{\alpha' s/2} \sim -\frac{8\pi i g_c^4 C_{S_2}}{\alpha' s + 4} . \tag{6.6.7}$$

Unitarity gives :

$$C_{S_2} = \frac{8\pi}{\alpha' g_c^2} , \tag{6.6.8}$$

therefore :

$$S_{S_2}(k_1;k_2;k_3) = \frac{8\pi i g_c}{\alpha'}(2\pi)^{26}\delta^{26}(\textstyle\sum_i k_i) . \tag{6.6.9}$$

Similar to the Veneziano amplitude, the 4-closed-string-tachyon amplitude can be expressed as gamma functions :

$$S_{S_2}(k_1;k_2;k_3;k_4) = \frac{8\pi i g_c^2}{\alpha'}(2\pi)^{26}\delta^{26}(\textstyle\sum_i k_i) C(-\alpha_c(t), -\alpha_c(u)) , \tag{6.6.10}$$

where $\alpha_c(x) = 1 + \alpha' x/4$ and :

$$C(a,b) = \int_{\mathbb{C}} d^2 z \, |z|^{2a-2} |1-z|^{2b-2}$$

$$= 2\pi \frac{\Gamma(a)\Gamma(b)\Gamma(c)}{\Gamma(a+b)\Gamma(a+c)\Gamma(b+c)} \,, \quad a+b+c=1\,. \tag{6.6.11}$$

This is the Virasoro-Shapiro amplitude . There is only one term, and the poles in the s, t, u channels come from the gamma functions in the numerator . Similar to the Veneziano amplitude, the Virasoro-Shapiro amplitude exhibits Regge behavior in the Regge limit :

$$S_{S_2}(k_1; k_2; k_3; k_4) \propto s^{2\alpha_c(t)} \frac{\Gamma(-\alpha_c(t))}{\Gamma(1+\alpha_c(t))} \,, \tag{6.6.12}$$

and exponential behavior in the hard scattering limit :

$$S_{S_2}(k_1; k_2; k_3; k_4) \propto \exp\left[-\frac{\alpha'}{2}(s \ln s\alpha' + t \ln t\alpha' + u \ln u\alpha')\right] . \tag{6.6.13}$$

On the sphere, the amplitude for one massless closed string and two closed string tachyons is :

$$\begin{aligned} S_{S_2}(k_1, e_1; k_2; k_3) &= g_c^2 g_c' \mathrm{e}^{-2\lambda} e_{1\mu\nu} \Big\langle :\tilde{c}c\partial X^\mu \bar{\partial} X^\nu \mathrm{e}^{ik_1\cdot X}(z_1, \bar{z}_1): \\ &\qquad :\tilde{c}c\mathrm{e}^{ik_2\cdot X}(z_2, \bar{z}_2)::\tilde{c}c\mathrm{e}^{ik_3\cdot X}(z_3, \bar{z}_3): \Big\rangle_{S_2} \\ &= -\frac{\pi i\alpha'}{2} g_c' e_{1\mu\nu} k_{23}^\mu k_{23}^\nu (2\pi)^{26} \delta^{26}(\textstyle\sum_i k_i) \,, \end{aligned} \tag{6.6.14}$$

where $e_{1\mu\nu}e_1^{\mu\nu} = 1$. Expanding the Virasoro-Shapiro amplitude (6.6.10) at the $s = 0$ pole, unitarity can give :

$$g_c' = \frac{2}{\alpha'} g_c \,, \tag{6.6.15}$$

which is again consistent with the state-operator map containing the overall constant g_c . The amplitude (6.6.14) can be obtained from field theory via the action $\boldsymbol{S} + \boldsymbol{S}_T$, where $\boldsymbol{S}$ is the action for massless fields (**??**), and :

$$\boldsymbol{S}_T = -\frac{1}{2} \int \mathrm{d}^{26}x\, (-G)^{1/2} \mathrm{e}^{-2\tilde{\Phi}} \left(G^{\mu\nu}\partial_\mu T \partial_\nu T - \frac{4}{\alpha'} T^2 \right) , \tag{6.6.16}$$

which is the action for a closed string tachyon T . For example, the graviton amplitude with polarization $e_{\mu\nu}$ can be obtained by expanding :

$$\tilde{G}_{\mu\nu} = \eta_{\mu\nu} - 2\kappa e_{\mu\nu} \mathrm{e}^{ik\cdot x} \,, \tag{6.6.17}$$

from this action . Note that this is the Einstein metric, and its action (**??**) does not contain the dilaton . The normalization of the fluctuation is determined by the fluctuation of the graviton kinetic energy term in the spacetime action . Specifically, if taking $\tilde{G}_{\mu\nu} - \eta_{\mu\nu} = -2\kappa e_{\mu\nu} f(x)$ where $e_{\mu\nu}e^{\mu\nu} = 1$, the effective action of f has the canonical normalization $\frac{1}{2}$ for real scalars . The field theory amplitude matches the string theory result (6.6.14) and relates to the normalization of the gravitational coupling vertex operator :

$$\kappa = \pi\alpha' g_c' = 2\pi g_c \,. \tag{6.6.18}$$

The 3-massless-closed-string amplitude is :

$$S_{S_2}(k_1, e_1; k_2, e_2; k_3, e_3) = \frac{i\kappa}{2} (2\pi)^{26} \delta^{26}(\textstyle\sum_i k_i) e_{1\mu\nu} e_{2\alpha\beta} e_{3\gamma\delta} T^{\mu\alpha\gamma} T^{\nu\beta\delta} \,, \tag{6.6.19}$$

where :

$$T^{\mu\alpha\gamma} = k_{23}^\mu \eta^{\alpha\gamma} + k_{31}^\alpha \eta^{\gamma\mu} + k_{12}^\gamma \eta^{\mu\alpha} + \frac{\alpha'}{8} k_{23}^\mu k_{31}^\alpha k_{12}^\gamma \,. \tag{6.6.20}$$

The k^2 term of this amplitude corresponds to the spacetime action (??), while the k^4 and k^6 terms come from several high-order derivative interactions, including terms of order 2 and 3 in the spacetime curvature . Higher-order corrections to the action can also be obtained by calculating high-loop corrections of the worldsheet β function (3.7.11) .

If we set $\alpha' = \frac{1}{2}$ for open strings and $\alpha' = 2$ for closed strings, the tensor structure of the closed string amplitude (6.6.19) is just two copies of the open string amplitude (6.5.15) . The same result holds for the amplitude (6.6.14) . This is the result of the free-field expectation values on the sphere factorizing into holomorphic and anti-holomorphic parts . Similar factorization occurs for four or more closed strings before integrating over the vertex operator positions . Furthermore, through careful treatment of the integration contours, it is possible to find relations between the integrated amplitudes . For 4-tachyon amplitudes, after using $\Gamma(x)\Gamma(1-x)\sin(\pi x) = \pi$, the above integrals have the following relationship :

$$J(s,t,u,\alpha') = -2\sin\pi\alpha_{\mathrm{c}}(t) I(s,t,4\alpha') I(t,u,4\alpha') \, ; \tag{6.6.21}$$

we have now indicated the explicit correlation of the integral for α' . For the general integral appearing in 4-point closed string amplitudes, there is the following relationship :

$$\begin{aligned} &\int_{\mathbb{C}} \mathrm{d}^2 z \, z^{a-1+m_1} \bar{z}^{a-1+n_1} (1-z)^{b-1+m_2} (1-\bar{z})^{b-1+n_2} \\ &\qquad = 2\sin[\pi(b+n_2)] B(a+m_1, b+m_2) B(b+n_2, 1-a-b-n_1-n_2) \, . \end{aligned} \tag{6.6.22}$$

This implies the relationship between 4-point open string amplitudes and 4-point closed string amplitudes :

$$A_{\mathrm{c}}(s,t,u,\alpha',g_{\mathrm{c}}) = \frac{\pi i g_{\mathrm{c}}^2 \alpha'}{g_{\mathrm{o}}^4} \sin[\pi\alpha_{\mathrm{c}}(t)] A_{\mathrm{o}}(s,t,\tfrac{1}{4}\alpha',g_{\mathrm{o}}) A_{\mathrm{o}}(t,u,\tfrac{1}{4}\alpha',g_{\mathrm{o}})^* \, , \tag{6.6.23}$$

where the open string amplitude contains only one of the 6 cyclic orderings, and the pole is in the stated channel .

Consistency

In chapter ??, we will discuss convergence and gauge invariance of tree-level amplitudes in a general way, but as an introduction, we now first use OPE to see how this works at the lowest order . Examine the operator product :

$$\begin{aligned} :\mathrm{e}^{ik_1\cdot X(z_1,\bar{z}_1)}::\mathrm{e}^{ik_4\cdot X(z_4,\bar{z}_4)}:= |z_{14}|^{\alpha' k_1\cdot k_4} :\Big(& 1 + i z_{14} k_1\cdot\partial X + i\bar{z}_{14} k_1\cdot\bar{\partial} X \\ & - z_{14}\bar{z}_{14} k_1\cdot\partial X k_1\cdot\bar{\partial} X + \cdots \Big) \mathrm{e}^{i(k_1+k_4)\cdot X}(z_4,\bar{z}_4): \, . \end{aligned} \tag{6.6.24}$$

It appears in amplitudes where z_{14} is integrated out . When :

$$\alpha' k_1\cdot k_4 = \frac{\alpha'}{2}(k_1+k_4)^2 - 4 > -2 \, , \tag{6.6.25}$$

the integral converges at $z_{14}\to 0$, and it has a pole at -2 . The coefficient of this pole is the tachyon vertex operator . Therefore, in any amplitude, if a pair of tachyons has total momentum $(k_1+k_4)^2 = 4/\alpha'$, there will be a pole, and unitarity requires this pole to be proportional to the amplitude with one fewer tachyon . Poles in momentum space correspond to long distances in spacetime, so this is a process where two tachyons scatter into one tachyon, which then interacts with other particles . Further doing OPE, the $O(z_{14})$ and $O(\bar{z}_{14})$ terms do not produce tachyons because the angular integration gives zero residue, and $O(z_{14}\bar{z}_{14})$ gives massless poles, and so on . Now we study more closely how local spacetime symmetry is maintained in string amplitudes . If any polarization vector is of the form $e_\mu = k_\mu$, or $e_{\mu\nu} = \zeta_\mu k_\nu + k_\mu\tilde{\zeta}_\mu$, where $k\cdot\zeta = k\cdot\tilde{\zeta} = 0$, the scattering amplitudes we calculated are all zero . This corresponds to the action being invariant under Yang-Mills, coordinate, and antisymmetric tensor symmetries . As discussed in section 3.6, the vertex operator of longitudinal polarization is the sum of a full derivative and another term, which is zero due to the equations of motion . After integration, the full derivative term is zero, but the equations of motion term may have sources at other vertex locations in the path integral . When $k_1 \cdot k_4$ is large, the operator product (6.6.24) decreases rapidly to zero . Using this property, for any pair of operators, there will always be a kinematic region where all possible connected terms (contact) are suppressed . Then the amplitude of null polarization is identically zero in this region because all amplitudes except at poles are analytic; the amplitude of null polarization must be zero everywhere . We see that the poles required by unitarity and the destruction of all possible divergences and spacetime gauge invariance arise in the limits $z \to 0, 1, \infty$ and when two vertex operators approach each other . For a sphere with 4 marked points, they are the boundaries of the moduli space . For historical reasons, the analytic continuation discussion here is called the canceled propagator argument .

Closed Strings on D_2 and RP_2

Lowest-order closed-string-open-string interactions come from disks having both closed string vertex operators and open string vertex operators . Their low-energy effective actions can be derived from a general discussion . In a trivial closed string background, we find the usual gauge kinetic term ,

$$-\frac{1}{4g_{\mathrm{o}}'^2}\int \mathrm{d}^{26}x\ \mathrm{Tr}(F_{\mu\nu}F^{\mu\nu})\,. \tag{6.6.26}$$

Clearly the metric must be coupled to it in a covariant way . In addition, coupling with the dilaton can be derived . Recall $g_{\mathrm{o}}'^2 \propto \mathrm{e}^{\Phi_0}$, where Φ_0 is the expectation value of the dilaton . So we replace it again (inside the integral) with $\Phi = \Phi_0 + \tilde{\Phi}$. The action :

$$-\frac{1}{4g_{\mathrm{o}}'^2}\int \mathrm{d}^{26}x\,(-G)^{1/2}\mathrm{e}^{-\tilde{\Phi}}\,\mathrm{Tr}(F_{\mu\nu}F^{\mu\nu}) \tag{6.6.27}$$

contains all interactions except for derivatives of the closed string field . Indices are now raised and lowered with $G_{\mu\nu}$. This action reflects the property that the effective action from Euler number χ contains weight $g_{\mathrm{c}}^{-\chi} \propto \mathrm{e}^{-\chi\Phi}$.

The disk and projective plane also contribute to pure closed string interactions . The amplitude of n closed strings is of order $g_{\mathrm{o}}^{-2}g_{\mathrm{c}}^{n} \sim g_{\mathrm{c}}^{n-1}$, one g_{c} more than the sphere . For closed string loop amplitudes considered in the next section, since emitting or absorbing a closed string adds 2 g_{c} factors, it is g_{c}^2 times the sphere . Therefore the sphere and the projective plane are at the "half-loop order" .

One of the most interesting amplitudes on the sphere and the projective plane is the one containing a single closed string vertex operator . Fixing the position of the vertex operator removes only two of the three CKVs; the residual gauge transformation consists of a rotation about the vertex operator position . Therefore, the scattering amplitude must be divided by the volume of the residual CKG . We have not yet explicitly proven it, but the torus in the next chapter will solve this problem . For a disk with one closed string, this is a finite and non-zero factor . The amplitude is a numerical factor times $g_{\mathrm{c}}g_{\mathrm{o}}^{-2}$, which is a pure number, multiplied by powers of α' (from dimensional analysis) . We will not solve for these numerical factors here, but will obtain them indirectly in chapter 8 .

Therefore, for closed strings emerging from a vacuum (momentum zero), there is an amplitude from both the disk and the projective plane; this type of amplitude is called a tadpole amplitude . In other words, the background closed string fields would be corrected to order g_{c} from their initial values . We can also write the effective action :

$$-\Lambda\int \mathrm{d}^{26}x\,(-G)^{1/2}\mathrm{e}^{-\tilde{\Phi}}\,, \tag{6.6.28}$$

which is the potential for the dilaton, which we will further examine in the next chapter .

On the other hand, for a single closed string on the sphere, the amplitude is zero . The residual CKG is the non-compact subgroup of $PSL(2,\mathbb{C})$, and thus this amplitude is divided by an infinite volume . A non-zero result would have logical inconsistency, i.e., zero-order correction to the background field . Similarly, the two-closed-string amplitude on the sphere (zero-order correction to mass) is also zero . For the corresponding disk amplitude, when there are one or two open strings, it is also zero . Amplitudes without vertex operators are also meaningful—they happen to calculate the $\tilde{\Phi}^0$ order term in the Taylor expansion of the action (6.6.28) . A disk without vertex operators is therefore not zero, which requires a formal treatment of the conformal Killing volume .

6.7 General results

In this section, we will obtain some general results for CFT on spheres and disks .

Möbius Invariance

We have seen that there is a globally defined group of conformal transformations on the sphere, the Möbius group $PSL(2,\mathbb{C})$:

$$z' = \frac{\alpha z+\beta}{\gamma z+\delta}\,, \tag{6.7.1}$$

where $\alpha,\beta,\gamma,\delta$ are complex numbers satisfying $\alpha\delta-\beta\gamma=1$. This is the most general one-to-one conformal transformation on the sphere S_2 . Expectation values must be invariant under any Möbius transformation :

$$\left\langle \mathscr{A}_i(z_1,\bar{z}_1)\ldots\mathscr{A}_k(z_n,\bar{z}_n)\right\rangle_{S_2} = \left\langle \mathscr{A}'_i(z_1,\bar{z}_1)\ldots\mathscr{A}'_k(z_n,\bar{z}_n)\right\rangle_{S_2}. \tag{6.7.2}$$

We will examine the consequences of this symmetry for 1, 2, 3, and 4 local operators .

For a single operator with weight $(h_i,\tilde{h}_i)$, rescaling plus rotation $z'=\gamma z$ gives :

$$\left\langle \mathscr{A}_i(0,0)\right\rangle_{S_2} = \left\langle \mathscr{A}'_i(0,0)\right\rangle_{S_2} = \gamma^{-h_i}\bar{\gamma}^{-\tilde{h}_i}\left\langle \mathscr{A}_i(0,0)\right\rangle_{S_2}. \tag{6.7.3}$$

Therefore, unless $h_i=\tilde{h}_i=0$, the one-point function is zero . Since the matter factor is the expectation value of the $(1,1)$ operator, this is another way to see that the one-point amplitude is zero on the sphere .

When $n=2$, we can use translation plus $z'=\gamma z$ to move any pair of operators to positions 0 and 1, giving :

$$\left\langle \mathscr{A}_i(z_1,\bar{z}_1)\mathscr{A}_j(z_2,\bar{z}_2)\right\rangle_{S_2} = z_{12}^{-h_i-h_j}\bar{z}_{12}^{-\tilde{h}_i-\tilde{h}_j}\left\langle \mathscr{A}_i(1,1)\mathscr{A}_j(0,0)\right\rangle_{S_2}, \tag{6.7.4}$$

so the position dependence is completely determined . Singularity implies $J_i+J_j\in\mathbb{Z}$, where $J_i=h_i-\tilde{h}_i$.

Remark. *$J_i+J_j=(h_i+h_j)-(\tilde{h}_i+\tilde{h}_j)$, so $z_{12}\to \mathrm{e}^{i\theta}z_{12}$, $\bar{z}_{12}\to\mathrm{e}^{-i\theta}\bar{z}_{12}$.*

Conformal transformation $z'=z+\epsilon(z-z_1)(z-z_2)+O(\epsilon^2)$ puts further constraints on the two-point function, keeping z_1 and z_2 fixed . For general operators this is complex, but for tensor fields $\mathcal{O}_p$ and $\mathcal{O}_q$, it implies :

$$\text{unless } h_p=h_q,\quad \tilde{h}_p=\tilde{h}_q\,,\qquad \left\langle \mathcal{O}_p(z_1,\bar{z}_1)\mathcal{O}_q(z_2,\bar{z}_2)\right\rangle_{S_2}=0\,. \tag{6.7.5}$$

By Möbius transformations, any three points $z_{1,2,3}$ can be brought to specific positions . When $n\geq 3$, Möbius invariance thus reduces expectation values from functions of n complex variables to functions of $n-3$ variables . Again, this result takes a simple form only for tensor fields . For example, for three tensor fields :

$$\left\langle \prod_{i=1}^{3} \mathcal{O}_{p_i}(z_i, \bar{z}_i) \right\rangle_{S_2} = C_{p_1p_2p_3} \prod_{i<j}^{3} z_{ij}^{h-2(h_i+h_j)} \bar{z}_{ij}^{\tilde{h}-2(\tilde{h}_i+\tilde{h}_j)} , \tag{6.7.6}$$

where $C_{p_1p_2p_3}$ is independent of position, and $h = h_1 + h_2 + h_3$. For four primary fields :

$$\left\langle \prod_{i=1}^{4} \mathcal{O}_{p_i}(z_i, \bar{z}_i) \right\rangle_{S_2} = C_{p_1p_2p_3p_4}(z_c, \bar{z}_c) \, (z_{12}z_{34})^h (\bar{z}_{12}\bar{z}_{34})^{\tilde{h}} \times \prod_{i<j}^{4} z_{ij}^{-h_i-h_j} \bar{z}_{ij}^{-\tilde{h}_i-\tilde{h}_j} , \tag{6.7.7}$$

where $h = \sum_i h_i$, $\tilde{h} = \sum_i \tilde{h}_i$, and $z_c = z_{12}z_{34}/z_{13}z_{24}$ is the Möbius-invariant cross ratio . The function $C_{p_1p_2p_3p_4}(z_c, \bar{z}_c)$ is not determined by conformal invariance, so we reduced the function of four variables to a function of a single variable .

On a disk represented as a half-plane, only Möbius transformations with real $\alpha, \beta, \gamma, \delta$ are retained, constituting the group $PSL(2, \mathbb{R})$. By examining the complete conformal algebra, we will gain more information . We will see in chapter 15 that it will determine all expectation values in terms of these tensor fields .

Path Integrals and Matrix Elements

Path integrals we have considered can be linked with operator representations . Consider the path integral for two operators on a sphere, one at the origin and one at infinity :

$$\left\langle \mathscr{A}_i'(\infty, \infty) \mathscr{A}_j(0,0) \right\rangle_{S_2} . \tag{6.7.8}$$

Primed labels denote the u coordinate system, which must be taken for operators at infinity; with slight ambiguity of notation, we can still represent position in terms of z . Using the state-operator map, one can replace the disk $|z| < 1$ containing $\mathscr{A}_j$ with the state $\Psi_{\mathscr{A}_j}$ on circle $|z| = 1$. We can also replace the disk $|z| > 1$ containing $\mathscr{A}_i$ with the state $\Psi_{\mathscr{A}_i}$ on circle $|z| = 1$. So the expectation value (6.7.8) becomes :

$$\int [\mathrm{d}\phi_\mathrm{b}] \; \Psi_{\mathscr{A}_i}[\phi_\mathrm{b}^\Omega] \Psi_{\mathscr{A}_j}[\phi_\mathrm{b}] \, ; \tag{6.7.9}$$

where $\phi_\mathrm{b}^\Omega(\sigma) = \phi_\mathrm{b}(2\pi - \sigma)$, from mapping $zu = 1$.

This inner product of wavefunctions is similar to an inner product, so we define :

$$\langle\!\langle i|j \rangle = \left\langle \mathscr{A}_i'(\infty, \infty) \mathscr{A}_j(0,0) \right\rangle_{S_2} . \tag{6.7.10}$$

This inner product was introduced by Zamolodchikov . For a unitary operator in a unitary theory, Möbius transformations show it is the same as the coefficient of 1 in the OPE :

$$\mathcal{O}_i(z, \bar{z}) \mathcal{O}_j(0,0) = \frac{\langle\!\langle i|j \rangle}{z^{h_i+h_j} \bar{z}^{\tilde{h}_i+\tilde{h}_j} \langle 1 \rangle_{S_2}} + \cdots . \tag{6.7.11}$$

We abbreviate $|\mathscr{A}_i\rangle$ as $|i\rangle$. The $\langle\!\langle \, | \, \rangle$ inner product is not the same as the quantum mechanical inner product $\langle \, | \, \rangle$. The latter is Hermitian, while the former is not; it is bilinear and does not contain complex conjugation, and there would be a sign difference if i and j anticommuted . i.e. :

$$\langle\!\langle i|j \rangle = \pm \langle\!\langle j|i \rangle , \tag{6.7.12}$$

because all we do is swap two operators and rename $z \leftrightarrow u$. Sometimes we write $\langle\!\langle i|j \rangle$ as $\mathscr{G}_{ij}$, and $\mathscr{G}^{ij}$ is the inverse metric; they can be used to raise and lower indices : $\mathscr{A}^i = \mathscr{G}^{ij} \mathscr{A}_j$, $\mathscr{A}_i = \mathscr{G}_{ij} \mathscr{A}^j$.

Operators in path integrals are translated into Hilbert space in the usual way . For example :

$$\left\langle \mathscr{A}_i'(\infty,\infty)\mathscr{A}_k(1,1)\mathscr{A}_j(0,0)\right\rangle_{S_2} = \langle\!\langle i|\hat{\mathscr{A}_k}(1,1)|j\rangle\ , \tag{6.7.13}$$

the " ˆ " is to emphasize that we are in Hilbert space . Using OPE, the left side becomes :

$$\sum_l c^l{}_{kj}\left\langle \mathscr{A}_i'(\infty,\infty)\mathscr{A}_l(0,0)\right\rangle_{S_2} = c_{ikj}\ . \tag{6.7.14}$$

Thus, three-point expectation values on a sphere, matrix elements of general local operators, and OPE coefficients where all indices are subscripts, are actually the same objects . By Möbius transformation, we also have :

$$\left\langle \mathscr{A}_i'(\infty,\infty)\mathscr{A}_k(z_1,\bar z_1)\mathscr{A}_j(0,0)\right\rangle_{S_2} = z_1^{h_i-h_k-h_j}\,\bar z_1^{\tilde h_i-\tilde h_k-\tilde h_j}c_{ikj}\ . \tag{6.7.15}$$

Four-point functions translated into operator representation :

$$\left\langle \mathscr{A}_i'(\infty,\infty)\mathscr{A}_k(z_1,\bar z_1)\mathscr{A}_l(z_2,\bar z_2)\mathscr{A}_j(0,0)\right\rangle_{S_2} = \langle\!\langle i|\mathrm{T}[\hat{\mathscr{A}_k}(z_1,\bar z_1)\hat{\mathscr{A}_l}(z_2,\bar z_2)]|j\rangle\!\rangle\ . \tag{6.7.16}$$

where T represents radial ordering . Let $|z_1| > |z_2|$ and insert the complete basis :

$$1 = |m\rangle\mathscr{G}^{mn}\langle\!\langle n|\ . \tag{6.7.17}$$

The 4-point amplitude (6.7.16) becomes :

$$\sum_m z_1^{h_i-h_k-h_m}\,\bar z_1^{\tilde h_i-\tilde h_k-\tilde h_m}\,z_2^{h_m-h_l-h_j}\,\bar z_2^{\tilde h_m-\tilde h_l-\tilde h_j}c_{ikm}c^m{}_{lj}\ . \tag{6.7.18}$$

Thus the operator product coefficients not only determine three-point expectation values, but also determine four-point expectation values . When $|z_1 - z_2| > |z_1|$, the expansion (6.7.18) holds in its overlap with the region $|z_1| > |z_2|$, and we can translate $\mathscr{A}_k$ to the origin and give a similar expression in the form of $c_{ilm}c^m{}_{kj}$. The equivalence of these two expansions shows the associativity of OPE .

Operator Calculation

The Hilbert space representation gives us another way to calculate expectation values . Let us take the 4-exponential operator as an example :

$$\begin{aligned}\Big\langle :\!\mathrm{e}^{ik_4\cdot X(\infty,\infty)}\!:'\,:\!\mathrm{e}^{ik_1\cdot X(z_1,\bar z_1)}\!::\!\mathrm{e}^{ik_2\cdot X(z_2,\bar z_2)}\!::\!\mathrm{e}^{ik_3\cdot X(0,0)}\!:\Big\rangle_{S_2}&\\ = \langle\!\langle 0;k_4|\mathrm{T}[:\mathrm{e}^{ik_1\cdot X_1} :: \mathrm{e}^{ik_2\cdot X_2}:]|0;k_3\rangle\ .&\end{aligned} \tag{6.7.19}$$

(2.7.11) proves that for this CFT, : : ordered operators are identical to : : ordered ones . By definition :

$$:\mathrm{e}^{ik\cdot X} := \mathrm{e}^{ik\cdot X_C}\mathrm{e}^{ik\cdot X_A}\ , \tag{6.7.20}$$

where :

$$X^\mu_C(z,\bar z) = x^\mu - i\left(\frac{\alpha'}{2}\right)^{1/2}\sum_{m=1}^\infty\frac{1}{m}(\alpha^\mu_{-m}z^m + \tilde\alpha^\mu_{-m}\bar z^m)\ , \tag{6.7.21a}$$

$$X^\mu_A(z,\bar z) = -i\frac{\alpha'}{2}p^\mu\ln|z|^2 + i\left(\frac{\alpha'}{2}\right)^{1/2}\sum_{m=1}^\infty\frac{1}{m}\left(\frac{\alpha^\mu_m}{z^m}+\frac{\tilde\alpha^\mu_m}{\bar z^m}\right)\ . \tag{6.7.21b}$$

For $|z_1| > |z_2|$, the matrix element (6.7.19) becomes :

$$\langle\!\langle 0;k_4|\mathrm{e}^{ik_1\cdot X_{1C}}\,\mathrm{e}^{ik_1\cdot X_{1A}}\,\mathrm{e}^{ik_2\cdot X_{2C}}\,\mathrm{e}^{ik_2\cdot X_{2A}}|0;k_3\rangle\ . \tag{6.7.22}$$

Using the Campbell-Baker-Hausdorff (CBH) formula :

$$\begin{aligned} \mathrm{e}^{ik_1\cdot X_{1A}}\,\mathrm{e}^{ik_2\cdot X_{2C}} &= \mathrm{e}^{ik_2\cdot X_{2C}}\mathrm{e}^{ik_1\cdot X_{1A}}\mathrm{e}^{-[k_1\cdot X_{1A},k_2\cdot X_{2C}]} \\ &= \mathrm{e}^{ik_2\cdot X_{2C}}\mathrm{e}^{ik_1\cdot X_{1A}}|z_{12}|^{\alpha' k_1\cdot k_2}\,. \end{aligned} \tag{6.7.23}$$

Remark. *CBH:* $\mathrm{e}^X\mathrm{e}^Y = \mathrm{e}^{X+Y+\frac{1}{2}[X,Y]}$, *where* $[X,Y]$ *commutes with* X *and* Y, *then :*

$$\mathrm{e}^Y\mathrm{e}^X\mathrm{e}_{[X,Y]} = \mathrm{e}^{Y+X-\frac{1}{2}[X,Y]}\mathrm{e}^{[X,Y]} = \mathrm{e}^{Y+X+\frac{1}{2}[X,Y]} = \mathrm{e}^X\mathrm{e}^Y\,.$$

Thus (6.7.22) becomes :

$$\begin{aligned} &|z_{12}|^{\alpha' k_1\cdot k_2}\langle\!\langle 0;k_4|\mathrm{e}^{ik_1\cdot X_{1C}+ik_2\cdot X_{2C}}\,\mathrm{e}^{ik_1\cdot X_{1A}+ik_2\cdot X_{2A}}|0;k_3\rangle \\ &\quad= |z_{12}|^{\alpha' k_1\cdot k_2}\langle\!\langle 0;k_4|\mathrm{e}^{i(k_1+k_2)\cdot x}\,\mathrm{e}^{\alpha'(k_1\ln|z_1|+k_2\ln|z_2|)\cdot p}|0;k_3\rangle \\ &\quad= |z_{12}|^{\alpha' k_1\cdot k_2}|z_1|^{\alpha' k_1\cdot k_3}|z_2|^{\alpha' k_2\cdot k_3}\langle\!\langle 0;k_1+k_2+k_4|0;k_3\rangle \\ &\quad= iC^X_{S_2}(2\pi)^d\delta^d(\textstyle\sum_i k_i)|z_{12}|^{\alpha' k_1\cdot k_2}|z_1|^{\alpha' k_1\cdot k_3}|z_2|^{\alpha' k_2\cdot k_3}\,, \end{aligned} \tag{6.7.24}$$

In the last line, we used the normalization for the two-point expectation value . This is similar to (6.2.31), the latter would gain the factor $|z_4|^{\alpha' k_4^2}$ from coordinate system transformation before letting $z_4 \to \infty$. All other results can be obtained through the oscillator method .

Remark. (6.2.31) *here is :*

$$|z_{12}|^{\alpha' k_1\cdot k_2}|z_1|^{\alpha' k_1\cdot k_3}|z_2|^{\alpha' k_2\cdot k_3}|z_1-z_4|^{\alpha' k_1\cdot k_4}|z_2-z_4|^{\alpha' k_2\cdot k_4}|z_3-z_4|^{\alpha' k_3\cdot k_4}\,,$$

where the last three factors related to z_4 *tend to* $|z_4|^{-\alpha' k_4\cdot k_4}$ *as* $z_4\to\infty$, *but* $\mathrm{e}^{ik_4\cdot x}$ *was taken in the* u *chart, so it must be multiplied by factor* $|z_4|^{\alpha' k_4\cdot k_4}$ *.*

Relations between Inner Products

By acting on the left vector with a suitable antilinear operator, non-degenerate bilinear inner products and Hermitian inner products can always be related to each other . Let's look at an example . From free field expectation values, using the fact that $|0;k\rangle$ maps to $\mathrm{e}^{ik\cdot X}$, we have :

$$\langle\!\langle 0;k|0;l\rangle = iC^X_{S_2}(2\pi)^d\delta^d(k+l)\,. \tag{6.7.25}$$

Compare it with the inner product of X^μCFT ,

$$\langle 0;k|0;l\rangle = (2\pi)^d\delta^d(l-k)\,. \tag{6.7.26}$$

They differ only by $k\to -k$ and normalization :

$$\langle\!\langle 0;k| = iC^X_{S_2}\langle\!\langle 0;-k|\,. \tag{6.7.27}$$

For more general operators, there is a natural notation for conjugation in CFT . In Euclidean quantum mechanics, Hermitian conjugation reverses Euclidean time, so the natural operator for Euclidean conjugation is conjugation × time reversal: Hermitian operators in Minkowski space undergoing this composite operation are also Hermitian . We do the same definition in CFT, but must simultaneously introduce time reversal on a conformal frame :

$$\overline{\mathscr{A}(p)} = \mathscr{A}'(p')^\dagger\,. \tag{6.7.28}$$

Here p and p' are related through radial spacetime inversion $z' = \bar{z}^{-1}$; primed operators are in the u coordinate system, and unprimed ones are in the z coordinate system . To see how it works, examine a holomorphic operator with weight h, whose Laurent expansion is :

$$\mathcal{O}(z) = i^h \sum_{n=-\infty}^{\infty} \frac{\mathcal{O}_n}{z^{n+h}} \,. \tag{6.7.29}$$

Its adjoint operator :

$$\mathcal{O}(z)^\dagger = i^{-h} \sum_{n=-\infty}^{\infty} \frac{\mathcal{O}_n^\dagger}{\bar{z}^{n+h}} \,. \tag{6.7.30}$$

Then Euclidean adjoint is :

$$\overline{\mathcal{O}(z)} = i^{-h}(-z^{-2})^h \sum_{n=-\infty}^{\infty} \frac{\mathcal{O}_n^\dagger}{z^{-n-h}} = i^h \sum_{n=-\infty}^{\infty} \frac{\mathcal{O}_{-n}^\dagger}{z^{n+h}} \,. \tag{6.7.31}$$

If an operator is Hermitian in Minkowski space, $\mathcal{O}_{-n}^\dagger = \mathcal{O}_n$, this operator is also Hermitian under $\overline{\mathcal{O}}$. The Euclidean conjugation (6.7.28) conjugates all i that appear, but keeps z and $\bar{z}$ unchanged . If the operator contains N_a anticommuting fields, then a factor of $(-1)^{N_\mathrm{a}(N_\mathrm{a}-1)/2}$ will be produced due to reversing the order of these fields .

This is the natural conformal invariant conjugation operation, so it must be :

$$\langle\!\langle \overline{\mathscr{A}}_i | = K \langle\!\langle \mathscr{A}_i | \tag{6.7.32}$$

where K is some constant . For a real illustration :

$$\begin{aligned} \langle\!\langle \bar{i} | j \rangle &\equiv \langle\!\langle 1 | \overline{\mathscr{A}_i'(\infty,\infty)} | j \rangle = K \langle 1 | \overline{\mathscr{A}_i'(\infty,\infty)} | j \rangle \\ &= K \langle 1 | \mathscr{A}_i(0,0)^\dagger | j \rangle = K \langle j | \mathscr{A}_i(0,0) | 1 \rangle^* \\ &= K \langle j | i \rangle^* = K \langle i | j \rangle \,. \end{aligned} \tag{6.7.33}$$

Here $\langle\!\langle 1| = K\langle 1|$ is neither a definition nor an assumption about the proportionality coefficient; it must hold because $|1\rangle$ is the unique $SL(2,\mathbb{C})$ invariant state .

For the X CFT, we see $K^X = iC^X_{S_2}$. For ghost field CFT, the Laurent expansion of amplitude (6.3.4) gives $\langle\!\langle 0|\tilde{c}_0 c_0|0\rangle = -C^\mathrm{g}_{S_2}$, where $|0\rangle = \tilde{c}_1 c_1 |1\rangle$. The Hermitian inner product is defined in (4.3.25) as $\langle 0|\tilde{c}_0 c_0|0\rangle = i$, so $K^\mathrm{g} = iC^\mathrm{g}_{S_2}$.

Besides the imaginary factor i prohibited by unitary CFT, $k = 1$ can be made by adding an Euler number term to the action . Thus, in the Hermitian basis of operators, the two inner products are equivalent, and differences between them can generally be ignored . However, note that vertex operators with non-zero momentum cannot be Hermitian .

The above discussion can be fully generalized to open strings, where the sphere is replaced by the disk .

Chapter 7

One-loop Amplitudes

After discussing several relevant surfaces, we will focus on the torus—considering first CFT on the torus and then scattering amplitudes. Subsequently, we generalize to open and unoriented string theories. The most important topic is understanding how short-distance ultraviolet divergences are forbidden.

7.1 Riemann Surfaces

There are four types of Riemann surfaces with zero Euler number.

Torus

The torus T^2 discussed in section **??** is the unique oriented closed surface with zero Euler number. We describe it as a complex plane with the metric $\mathrm{d}s^2 = \mathrm{d}w\mathrm{d}\bar{w}$ and the equivalence relations

$$w \cong w + 2\pi \cong w + 2\pi\tau \, . \tag{7.1.1}$$

There are two moduli, the real and imaginary parts of $\tau = \tau_1 + i\tau_2$, and two CKVs: translations. In terms of real coordinates $w = \sigma^1 + i\sigma^2$, we have

$$(\sigma^1, \sigma^2) \cong (\sigma^1 + 2\pi, \sigma^2) \cong (\sigma^1 + 2\pi\tau_1, \sigma^2 + 2\pi\tau_2) \, , \tag{7.1.2}$$

which allows us to view the torus as a cylinder with base circumference 2π and height $2\pi\tau_2$, but with the ends rotated by $2\pi\tau_1$ and then glued together.

In coordinates $z = \exp(-iw)$, the equivalence $w \cong w + 2\pi$ is automatic, while $w \cong w + 2\pi\tau$ becomes

$$z \cong z \exp(-2\pi i\tau) \, . \tag{7.1.3}$$

The fundamental region is the annulus

$$1 \leq |z| \leq \exp(2\pi\tau_2) \, . \tag{7.1.4}$$

The torus is obtained by rotating the outer circle by $2\pi\tau_1$ and then gluing it to the inner circle. Unless otherwise stated, we will use w coordinates.

Cylinder (Annulus)

The cylinder C_2 is defined by

$$0 \leq \operatorname{Re} w \leq \pi, \quad w \cong w + 2\pi i t \, . \tag{7.1.5}$$

This is a strip of width π and length $2\pi t$, with the ends of the strip glued together. There is only one modulus t in the range $0 < t < \infty$. Unlike the torus, there is no modular group; the long cylinder limit $t \to 0$ is quite different from the long strip limit $t \to \infty$. There is only one CKV: translation parallel to the boundary.

The cylinder can be obtained from a torus with $\tau = it$ via the following involution equivalence

$$w' = -\bar{w}\,, \tag{7.1.6}$$

which is a reflection about the imaginary axis. The lines $\sigma^1 = 0, \pi$ are fixed by this reflection and thus become boundaries. The other coordinate has periodic equivalence

$$(\sigma^1, 0) \cong (\sigma^1, 2\pi t)\,. \tag{7.1.7}$$

Klein Bottle

The Klein bottle K_2 can be viewed as the complex plane with the equivalences

$$w \cong w + 2\pi \cong -\bar{w} + 2\pi i t \tag{7.1.8}$$

or

$$(\sigma^1, \sigma^2) \cong (\sigma^1 + 2\pi, \sigma^2) \cong (-\sigma^1, \sigma^2 + 2\pi t)\,. \tag{7.1.9}$$

This is a cylinder of circumference 2π and height $2\pi t$, but with the ends glued via a parity transformation Ω. The single modulus t spans $0 < t < \infty$, and there is no modular group. Translation in the σ^2 direction is the unique CKV.

The Klein bottle can be obtained from a torus with $\tau = 2it$, subject to the equivalence

$$w' = -\bar{w} + 2\pi i t\,. \tag{7.1.10}$$

The Klein bottle can also be viewed as a sphere with two crosscaps.

Möbius Strip

The Möbius strip M_2 can be obtained from a long strip via Ω

$$0 \le \mathrm{Re}\, w \le \pi, \qquad w \cong -\bar{w} + \pi + 2\pi i t\,. \tag{7.1.11}$$

The modulus t spans $0 < t < \infty$, and the CKV is translation in σ^2. It can be obtained from a torus with $\tau = 2it$ through two involution equivalences

$$w' = -\bar{w} \qquad \text{and} \qquad w' = w + \pi(2it + 1)\,. \tag{7.1.12}$$

The Möbius strip can be viewed as a disk with one crosscap.

7.2 CFT on the Torus

Scalar Correlators

As in the case of the sphere, we start from the Green function (6.2.7), which satisfies

$$\frac{2}{\alpha'}\bar{\partial}\partial G'(w, \bar{w}; w', \bar{w}') = -2\pi\delta^2(w - w') + \frac{1}{4\pi\tau_2}\,. \tag{7.2.1}$$

Remark. *In (6.2.8), we now have* $g^{-1/2} = 1$ *and* $\delta^2(\sigma_1 - \sigma_2) = 2\delta^2(w - w')$*, so*

$$\int \mathrm{d}^2 z\, x_0^2 = 1 \quad \Rightarrow \quad (2\pi)^2 \tau_2 x_0^2 = 1 \quad \Rightarrow \quad x_0^2 = \frac{1}{(2\pi)^2 \tau_2}$$

Then (6.2.8) becomes

$$\frac{1}{2\pi\alpha'} \nabla^2 G' = -2\delta^2(w - w') + \frac{1}{4\pi^2 \tau_2} \,,$$

and using $\nabla^2 = 4\bar{\partial}\partial$*, we obtain (7.2.1).*

The Green function is periodic in both directions of the torus. After removing the source and background charge terms, it should be the sum of a holomorphic and an anti-holomorphic function. These properties relate it to theta functions. We conjecture

$$G'(w, \bar{w}; w', \bar{w}') \sim -\frac{\alpha'}{2} \ln \left| \vartheta_1\!\left(\frac{w - w'}{2\pi} \middle| \tau \right) \right|^2 . \tag{7.2.2}$$

As the variable of ϑ_1 approaches zero, it goes to zero linearly, providing the correct behavior for $w \to w'$. However, due to the quasi-periodicity (7.2.32b), it is not perfectly bi-periodic—it changes under $w \to w + 2\pi\tau$ by $-\alpha'[\mathrm{Im}(w - w') + \pi\tau_2]$. Additionally, the background charge is missing. Both points can be easily corrected:

$$G'(w, \bar{w}; w', \bar{w}') = -\frac{\alpha'}{2} \ln \left| \vartheta_1\!\left(\frac{w - w'}{2\pi} \middle| \tau \right) \right|^2 + \alpha' \frac{[\mathrm{Im}(w - w')]^2}{4\pi\tau_2} + k(\tau, \bar{\tau}) \,. \tag{7.2.3}$$

The function $k(\tau, \bar{\tau})$ is determined by the orthogonality of X_0, but as with the sphere, it is dropped due to spacetime momentum conservation.

Parallel to the previous result (6.2.13), the expectation value of vertex operators is

$$\begin{aligned} \left\langle \prod_{i=1}^{n} :\mathrm{e}^{ik_i \cdot X(z_i, \bar{z}_i)}: \right\rangle_{T^2} &= iC^X_{T^2}(\tau)(2\pi)^d \delta^d(\textstyle\sum_i k_i) \\ &\times \prod_{i<j} \left| \frac{2\pi}{\partial_\nu \vartheta_1(0|\tau)} \vartheta_1\!\left(\frac{w_{ij}}{2\pi} \middle| \tau \right) \exp\!\left[-\frac{(\mathrm{Im}\, w_{ij})^2}{4\pi\tau_2} \right] \right|^{\alpha' k_i \cdot k_j} . \end{aligned} \tag{7.2.4}$$

The factor $2\pi/\partial_\nu \vartheta_1$ arises from renormalization of self-contractions.

Scalar Partition Function

The total normalization for the sphere was absorbed into the string coupling constant, but we cannot do the same for the torus because it has a non-trivial τ dependence. In fact, we will see significant physics from amplitudes without vertex operators. Consider the path integral without vertex operators $\langle 1 \rangle_{T^2(\tau)} \equiv Z(\tau)$. We can view the torus with modulus τ as a field theory on a circle evolving for Euclidean time $2\pi\tau_2$, translated in the σ^1 direction by $2\pi\tau_1$, and then identifying the ends. In the language of operators, this gives the trace

$$\begin{aligned} Z(\tau) &= \mathrm{Tr}\Big[\exp(2\pi i \tau_1 P - 2\pi\tau_2 H) \Big] \\ &= (q\bar{q})^{-d/24}\, \mathrm{Tr}\Big(q^{L_0} \bar{q}^{\tilde{L}_0} \Big) \,. \end{aligned} \tag{7.2.5}$$

Here $q = \exp(2\pi i \tau)$. The momentum $P = L_0 - \tilde{L}_0$ generates translations in σ^1, while the Hamiltonian $H = L_0 + \tilde{L}_0 - \frac{1}{24}(c + \tilde{c})$ generates translations in σ^2. As in statistical mechanics, this trace weighted by the exponential of the Hamiltonian and other conserved quantities is called the partition function.

This trace can be split into a sum over occupation numbers $N_{\mu n}$ and $\tilde{N}_{\mu n}$ and an integral over momentum k^μ. $Z(\tau)$ becomes

$$V_d(q\bar{q})^{-d/24}\int \frac{\mathrm{d}^d k}{(2\pi)^d}\exp\bigl(-\pi\tau_2\alpha' k^2\bigr)\prod_{\mu,n}\sum_{N_{\mu n},\tilde{N}_{\mu n}=0}^{\infty} q^{nN_{\mu n}}\bar{q}^{n\tilde{N}_{\mu n}} . \tag{7.2.6}$$

The spacetime volume factor V_d comes from the continuous normalization of momentum, where $\sum_k$ becomes $V_d(2\pi)^{-d}\int \mathrm{d}^d k$. The sum is a geometric series

$$\sum_{N=0}^{\infty} q^{nN} = (1-q^n)^{-1} , \tag{7.2.7}$$

yielding

$$Z(\tau) = iV_d Z_X(\tau)^d , \tag{7.2.8}$$

where

$$Z_X(\tau) = (4\pi^2\alpha'\tau_2)^{-1/2}|\eta(\tau)|^{-2} . \tag{7.2.9}$$

Here

$$\eta(\tau) = q^{1/24}\prod_{n=1}^{\infty}(1-q^n) \tag{7.2.10}$$

is the Dedekind eta function. The factor of i comes from the rotation $k^0 \to ik^d$. Also, previously $C^X_{T^2}(\tau) = Z_X(\tau)^d$.

(7.2.8) is clearly invariant under $\tau \to \tau+1$. The invariance of Z_X under $\tau \to -1/\tau$ is given by (7.2.44), $\eta(-1/\tau) = (-i\tau)^{1/2}\eta(\tau)$. This generates the entire modular group. Similarly, expectation values with vertex operators are modular covariant.

This calculation can be performed using holomorphy. The main idea is to obtain a differential equation for τ. Consider a torus with modulus τ and make a small change to the metric $\delta g_{ww} = \epsilon^*$. The new metric is

$$\begin{aligned}\mathrm{d}s^2 &= \mathrm{d}w\mathrm{d}\bar{w} + \epsilon^*\mathrm{d}w^2 + \epsilon\mathrm{d}\bar{w}^2 \\ &= (1+\epsilon^*+\epsilon)\mathrm{d}[w+\epsilon(\bar{w}-w)]\,\mathrm{d}[\bar{w}+\epsilon^*(w-\bar{w})] + O(\epsilon^2) .\end{aligned} \tag{7.2.11}$$

That is, this metric is Weyl equivalent to $\mathrm{d}w'\mathrm{d}\bar{w}'$, where $w' = w+\epsilon(\bar{w}-w)$. It has periodicity

$$w' \cong w' + 2\pi \cong w' + 2\pi(\tau - 2i\tau_2\epsilon) . \tag{7.2.12}$$

The change in the metric is thus equivalent to a change in the modulus

$$\delta\tau = -2i\tau_2\epsilon . \tag{7.2.13}$$

After changing the metric, the change in the path integral is

$$\begin{aligned}\delta Z(\tau) &= -\frac{1}{2\pi}\int \mathrm{d}^2 w\,\Bigl[\delta g_{\bar{w}\bar{w}}\bigl\langle T_{ww}(w)\bigr\rangle + \delta g_{ww}\bigl\langle T_{\bar{w}\bar{w}}(\bar{w})\bigr\rangle\Bigr] \\ &= -2\pi i\Bigl[\delta\tau\bigl\langle T_{ww}(0)\bigr\rangle - \delta\bar{\tau}\bigl\langle T_{\bar{w}\bar{w}}(0)\bigr\rangle\Bigr] .\end{aligned} \tag{7.2.14}$$

In the second line, we used translation invariance to perform the integration. To find the expectation value of the energy-momentum tensor, we use the OPE

$$\partial_w X^\mu(w)\partial_w X_\mu(0) = -\frac{\alpha' d}{2w^2} - \alpha' T_{ww}(0) + O(w^2) . \tag{7.2.15}$$

Now,

$$\begin{aligned} Z(\tau)^{-1}\langle\partial_w X^\mu(w)\partial_w X_\mu(0)\rangle &= \partial_w\partial_{w'}G'(w,\bar{w};w',\bar{w}')|_{w'=0} \\ &= \frac{\alpha' d}{2}\frac{\vartheta_1\partial_w^2\vartheta_1 - \partial_w\vartheta_1\partial_w\vartheta_1}{\vartheta_1^2} + \frac{\alpha' d}{8\pi\tau_2}\,, \end{aligned} \tag{7.2.16}$$

where all variables of the theta functions are $(w/2\pi, \tau)$. It indeed has a double pole at $w = 0$. Carefully expanding the numerator and denominator to order w^2, the w^0 term in (7.2.16) is

$$\frac{\alpha' d}{6}\frac{\partial_w^3\vartheta_1}{\partial_w\vartheta_1} + \frac{\alpha' d}{8\pi\tau_2}\,. \tag{7.2.17}$$

Then the OPE (7.2.15) gives

$$\langle T_{ww}(0)\rangle = \left(-\frac{d}{6}\frac{\partial_w^3\vartheta_1(0|\tau)}{\partial_w\vartheta_1(0|\tau)} - \frac{d}{8\pi\tau_2}\right)Z(\tau)\,, \tag{7.2.18}$$

and the variation (7.2.14) gives the differential equation

$$\partial_\tau \ln Z(\tau) = \frac{\pi i d}{3}\frac{\partial_w^3\vartheta_1(0|\tau)}{\partial_w\vartheta_1(0|\tau)} + \frac{id}{4\tau_2}\,. \tag{7.2.19}$$

To go further, using

$$\partial_w^2\vartheta_1\left(\frac{w}{2\pi}\middle|\tau\right) = \frac{i}{\pi}\partial_\tau\vartheta_1\left(\frac{w}{2\pi}\middle|\tau\right)\,. \tag{7.2.20}$$

Rewrite (7.2.19) as

$$\partial_\tau \ln Z(\tau) = -\frac{d}{3}\partial_\tau \ln \partial_w\vartheta_1(0|\tau) + \frac{id}{4\tau_2}\,. \tag{7.2.21}$$

Together with the conjugate equation, this determines $Z(\tau)$ up to a numerical coefficient as

$$Z(\tau) = |\partial_w\vartheta_1(0|\tau)|^{-2d/3}\tau_2^{-d/2}\,, \tag{7.2.22}$$

The numerical coefficient is determined by other methods. Using (7.2.43), we can see it matches (7.2.8).

bc CFT

For the ghost CFT, the partition function is also determined by the trace of states. On both the left and right sides, there are two sets of creation operators b_{-n} and c_{-n}, but they are anticommutative. Thus, occupation numbers can only be 0 or 1. Therefore, an oscillator of mode n contributes $(1 + q^n)$. Additionally, there are four ground states, which give

$$\begin{aligned} \mathrm{Tr}\Big[\exp(2\pi i\tau_1 P - 2\pi\tau_2 H)\Big] &= (q\bar{q})^{13/12}\,\mathrm{Tr}\Big(q^{L_0}\bar{q}^{\tilde{L}_0}\Big) \\ &= 4(q\bar{q})^{1/12}\prod_{n=1}^{\infty}|1+q^n|^4\,. \end{aligned} \tag{7.2.23}$$

However, for anticommutative fields, this trace corresponds to a path integral with antiperiodic boundaries in the time direction. To calculate the Faddeev-Popov determinant, the ghost fields must have the same periodicity as the original coordinate transformation, which is periodic. Therefore, we need

$$\begin{aligned} Z(\tau) &= \mathrm{Tr}\Big[(-1)^F\exp(2\pi i\tau_1 P - 2\pi\tau_2 H)\Big] \\ &= 0\,, \end{aligned} \tag{7.2.24}$$

where $(-1)^F$ anticommutes with all ghost fields. The trace is zero because the ghost numbers of $|{\downarrow\downarrow}\rangle$ and $|{\uparrow\uparrow}\rangle$ are opposite to those of $|{\uparrow\downarrow}\rangle$ and $|{\downarrow\uparrow}\rangle$ modulo 2. From the path integral viewpoint, $Z(\tau)$ is zero due to ghost zero modes arising from the moduli and CKVs. They must be saturated by suitable insertions. The simplest non-zero amplitude is

$$\left\langle c(w_1)b(w_2)\tilde{c}(\bar{w}_3)\tilde{b}(\bar{w}_4)\right\rangle . \tag{7.2.25}$$

In the operator calculation, we write each field in its mode expansion. Only the $n = 0$ terms contribute. Then (7.2.25) becomes

$$\mathrm{Tr}\Big[(-1)^F c_0 b_0 \tilde{c}_0 \tilde{b}_0 \exp(2\pi i\tau_1 P - 2\pi\tau_2 H)\Big] . \tag{7.2.26}$$

The operator $c_0 b_0 \tilde{c}_0 \tilde{b}_0$ projects onto the ground state $|{\uparrow\uparrow}\rangle$, so the result is

$$(q\bar{q})^{1/12} \prod_{n=1}^{\infty} |1 - q^n|^4 = |\eta(\tau)|^4 . \tag{7.2.27}$$

Note the sign change in the infinite product comes from $(-1)^F$ in the trace. Since the CKVs and quadratic differentials on the torus are constant, (7.2.27) is independent of the position of the ghost fields.

General CFT

For a general CFT

$$Z(\tau) = \sum_i q^{h_i - c/24} \bar{q}^{\tilde{h}_i - \tilde{c}/24} (-1)^{F_i} , \tag{7.2.28}$$

where i runs over all states of the CFT, and F_i is the worldsheet fermion number. Invariance under $\tau \to \tau + 1$ requires $h_i - \tilde{h}_i - \frac{1}{24}(c - \tilde{c})$ to be an integer. Considering the identity operator, this requires $c - \tilde{c}$ to be a multiple of 24. For general operators, this then requires integer spin

$$h_i - \tilde{h}_i \in \mathbb{Z} . \tag{7.2.29}$$

Incidentally, CFTs where $c - \tilde{c}$ is not a multiple of 24 must also be interesting. They cannot be modular invariant alone, but with suitable projections, they can be integrated into a modular invariant theory.

Invariance under $\tau \to -1/\tau$ imposes further restrictions on the spectrum. Consider the partition function with $\tau = i\ell$, and let $\ell \to 0$; the convergence factor $q = \exp(-2\pi\ell)$ approaches 1. Thus, the partition function is determined by the density of the highest weight states. Using modular invariance, this equals the partition function with $\tau = i/\ell$, where the latter's $q = \exp(-2\pi/\ell)$ approaches 0. Thus, the lowest weight state dominates the sum, which is the identity state with $L_0 = \tilde{L}_0 = 0$, yielding

$$Z(i\ell) \overset{l\to 0}{\approx} \exp\left[\frac{\pi(c + \tilde{c})}{12\ell}\right] . \tag{7.2.30}$$

The density of highest weight states is thus determined by the central charge. This produces the result for free bosons, where $c = \tilde{c} = d$ counts the number of free bosons.

Theta Functions

The fundamental theta function is

$$\vartheta(\nu, \tau) = \sum_{n=-\infty}^{\infty} \exp\big(\pi i n^2 \tau + 2\pi i n \nu\big) . \tag{7.2.31}$$

It has periodicities

$$\vartheta(\nu+1,\tau)=\vartheta(\nu,\tau)\,, \tag{7.2.32a}$$

$$\vartheta(\nu+\tau,\tau)=\exp(-\pi i\tau-2\pi i\nu)\vartheta(\nu,\tau)\,, \tag{7.2.32b}$$

and under modular transformations

$$\vartheta(\nu,\tau+1)=\vartheta(\nu+1/2,\tau)\,, \tag{7.2.33a}$$

$$\vartheta(\nu/\tau,-1/\tau)=(-i\tau)^{1/2}\exp\big(\pi i\nu^2/\tau\big)\vartheta(\nu,\tau)\,. \tag{7.2.33b}$$

Aside from the periodicity (7.2.32), the theta function has only one zero at $v=\frac{1}{2}(1+\tau)$. It can also be written as an infinite product

$$\vartheta(\nu,\tau)=\prod_{m=1}^{\infty}(1-q^m)(1+zq^{m-1/2})(1+z^{-1}q^{m-1/2})\,, \tag{7.2.34}$$

where

$$q=\exp(2\pi i\tau),\quad z=\exp(2\pi i\nu)\,. \tag{7.2.35}$$

We often need the asymptotic behavior of theta functions as $q\to 0$ or $q\to 1$. $q\to 0$ can be read immediately from the product or sum. The behavior at $q\to 1$ is less obvious but can be obtained from the modular transformation $\tau\to -1/\tau$, which relates the two limits via $q\to\exp\left(4\pi^2/\ln q\right)$. It is often useful to define theta functions with characteristics

$$\begin{aligned}\vartheta\left[\begin{smallmatrix}a\\ b\end{smallmatrix}\right](\nu,\tau)&=\exp\Big[\pi i a^2\tau+2\pi i a(\nu+b)\Big]\vartheta(\nu+a\tau+b,\tau)\\ &=\sum_{n=-\infty}^{\infty}\exp\Big[\pi i(n+a)^2\tau+2\pi i(n+a)(\nu+b)\Big]\,.\end{aligned} \tag{7.2.36}$$

Other common notations are

$$\vartheta_{00}(\nu,\tau)=\vartheta_3(\nu|\tau)=\vartheta\left[\begin{smallmatrix}0\\ 0\end{smallmatrix}\right](\nu,\tau)=\sum_{n=-\infty}^{\infty}q^{n^2/2}z^n\,, \tag{7.2.37a}$$

$$\vartheta_{01}(\nu,\tau)=\vartheta_4(\nu|\tau)=\vartheta\left[\begin{smallmatrix}0\\ 1/2\end{smallmatrix}\right](\nu,\tau)=\sum_{n=-\infty}^{\infty}(-1)^n q^{n^2/2}z^n\,, \tag{7.2.37b}$$

$$\vartheta_{10}(\nu,\tau)=\vartheta_2(\nu|\tau)=\vartheta\left[\begin{smallmatrix}1/2\\ 0\end{smallmatrix}\right](\nu,\tau)=\sum_{n=-\infty}^{\infty}q^{(n-1/2)^2/2}z^{n-1/2}\,, \tag{7.2.37c}$$

$$\vartheta_{11}(\nu,\tau)=-\vartheta_1(\nu|\tau)=\vartheta\left[\begin{smallmatrix}1/2\\ 1/2\end{smallmatrix}\right](\nu,\tau)=-i\sum_{n=-\infty}^{\infty}(-1)^n q^{(n-1/2)^2/2}z^{n-1/2}\,. \tag{7.2.37d}$$

Product representations:

$$\vartheta_{00}(\nu,\tau)=\prod_{m=1}^{\infty}(1-q^m)(1+zq^{m-1/2})(1+z^{-1}q^{m-1/2})\,, \tag{7.2.38a}$$

$$\vartheta_{01}(\nu,\tau)=\prod_{m=1}^{\infty}(1-q^m)(1-zq^{m-1/2})(1-z^{-1}q^{m-1/2})\,, \tag{7.2.38b}$$

$$\vartheta_{10}(\nu,\tau)=2\exp(\pi i\tau/4)\cos\pi\nu\prod_{m=1}^{\infty}(1-q^m)(1+zq^m)(1+z^{-1}q^m)\,, \tag{7.2.38c}$$

$$\vartheta_{11}(\nu,\tau)=-2\exp(\pi i\tau/4)\sin\pi\nu\prod_{m=1}^{\infty}(1-q^m)(1-zq^m)(1-z^{-1}q^m)\,. \tag{7.2.38d}$$

Modular transformations

$$\vartheta_{00}(\nu,\tau+1) = \vartheta_{01}(\nu,\tau)\,, \tag{7.2.39a}$$
$$\vartheta_{01}(\nu,\tau+1) = \vartheta_{00}(\nu,\tau)\,, \tag{7.2.39b}$$
$$\vartheta_{10}(\nu,\tau+1) = \exp(\pi i/4)\vartheta_{10}(\nu,\tau)\,, \tag{7.2.39c}$$
$$\vartheta_{11}(\nu,\tau+1) = \exp(\pi i/4)\vartheta_{11}(\nu,\tau)\,, \tag{7.2.39d}$$

and

$$\vartheta_{00}(\nu/\tau,-1/\tau) = (-i\tau)^{1/2}\exp\bigl(\pi i\nu^2/\tau\bigr)\vartheta_{00}(\nu,\tau)\,, \tag{7.2.40a}$$
$$\vartheta_{01}(\nu/\tau,-1/\tau) = (-i\tau)^{1/2}\exp\bigl(\pi i\nu^2/\tau\bigr)\vartheta_{10}(\nu,\tau)\,, \tag{7.2.40b}$$
$$\vartheta_{10}(\nu/\tau,-1/\tau) = (-i\tau)^{1/2}\exp\bigl(\pi i\nu^2/\tau\bigr)\vartheta_{01}(\nu,\tau)\,, \tag{7.2.40c}$$
$$\vartheta_{11}(\nu/\tau,-1/\tau) = -i(-i\tau)^{1/2}\exp\bigl(\pi i\nu^2/\tau\bigr)\vartheta_{11}(\nu,\tau)\,. \tag{7.2.40d}$$

We will encounter these functions more frequently in superstrings.

Theta functions satisfy Jacobi's "abstruse identity", which is a special case of the Riemann quartic identity

$$\vartheta_{00}^4(0,\tau) - \vartheta_{01}^4(0,\tau) - \vartheta_{10}^4(0,\tau) = 0\,. \tag{7.2.41}$$

Also note

$$\vartheta_{11}(0,\tau) = 0\,. \tag{7.2.42}$$

Finally, the Dedekind eta function is

$$\eta(\tau) = q^{1/24}\prod_{m=1}^{\infty}(1-q^m) = \left[\frac{\partial_\nu\vartheta_{11}(0,\tau)}{-2\pi}\right]^{1/3}\,. \tag{7.2.43}$$

Modular transformations

$$\eta(\tau+1) = \exp(i\pi/12)\eta(\tau)\,, \tag{7.2.44a}$$
$$\eta(-1/\tau) = (-i\tau)^{1/2}\eta(\tau)\,. \tag{7.2.44b}$$

7.3 Torus Amplitudes

We now apply the general result (**??**) to the torus: expressing string scattering amplitudes as integrals with ghost and vertex operator insertions. Two CKVs require one vertex operator to be fixed, so

$$\begin{aligned} &S_{T^2}(1;2;\ldots;n) \\ &\quad= \frac{1}{2}\int_{F_0} d\tau d\bar\tau \left\langle B\tilde B\tilde c c\mathscr{V}_1(w_1,\bar w_1)\prod_{i=2}^{n}\int dw_i d\bar w_i\ \mathscr{V}_i(w_i,\bar w_i)\right\rangle_{T^2}\,. \end{aligned} \tag{7.3.1}$$

The fundamental region F_0 is the same as discussed in section **??**, and the $\frac{1}{2}$ comes from $w \to -w$. As with complex field variables, $d\tau d\bar\tau = 2d\tau_1 d\tau_2$, where $\tau = \tau_1 + i\tau_2$. The ghost insertion for $d\tau$ is

$$\begin{aligned} B = \frac{1}{4\pi}(b,\partial_\tau) &= \frac{1}{2\pi}\int d^2w\, b_{ww}(w)\partial_\tau g_{\bar w\bar w} \\ &= \frac{i}{4\pi\tau_2}\int d^2w\, b_{ww}(w) \\ &\to 2\pi i b_{ww}(0)\,. \end{aligned} \tag{7.3.2}$$

We used (7.2.13). In the last line, we used the fact that the ghost path integral (7.2.27) is position-independent.

The CKG is composed of translations of the torus. Since this group has a finite volume, we do not need to fix it; we can rewrite the amplitude to integrate over all vertex operators and then divide by the volume of the CKG. CKVs are constant, so the expectation value in (7.3.1) is independent of the position of c: we can move c away from w_1 and place them at some fixed location. For vertex operators, translation invariance implies the amplitude is invariant under $w_i \to w_i + w$. Average over this translation:

$$\int \frac{\mathrm{d}w\mathrm{d}\bar{w}}{2(2\pi)^2\tau_2}\,, \tag{7.3.3}$$

where the denominator is the area of the torus, which is the volume of the CKG. Using (7.3.2) and (7.3.3), the amplitude becomes

$$\begin{aligned} &S_{T^2}(1;2;\ldots;n) \\ &= \int_{F_0} \frac{\mathrm{d}\tau\mathrm{d}\bar{\tau}}{4\tau_2} \left\langle b(0)\tilde{b}(0)\tilde{c}(0)c(0) \prod_{i=1}^{n} \int \mathrm{d}w_i\mathrm{d}\bar{w}_i\, \mathscr{V}_i(w_i,\bar{w}_i) \right\rangle_{T^2}. \end{aligned} \tag{7.3.4}$$

Now all vertex operators have equal status.

> **Remark.** *The factor* $(4\tau_2)^{-1}$ *in* (7.3.4) *comes from* $1/2$ *in* (7.3.1), $(2\pi)^{-2}$ *in* (7.3.2), *and* $(2(2\pi)^2\tau_2)^{-1}$ *in* (7.3.3).

Without the intermediate step of fixing vertex operators, (7.3.4) can also be derived directly; even without vertex operators, it holds:

$$Z_{T^2} = \int_{F_0} \frac{\mathrm{d}\tau\mathrm{d}\bar{\tau}}{4\tau_2} \left\langle b(0)\tilde{b}(0)\tilde{c}(0)c(0) \right\rangle_{T^2}. \tag{7.3.5}$$

This vacuum amplitude is quite interesting and will be our main goal of study. It not only reveals that the UV behavior of string theory differs from field theory, but also has significant physical meaning itself.

For 26 flat spacetime dimensions, the path integrals of matter and ghost fields in (7.3.5) were calculated in the previous section, yielding

$$Z_{T^2} = iV_{26} \int_{F_0} \frac{\mathrm{d}\tau\mathrm{d}\bar{\tau}}{4\tau_2} (4\pi^2\alpha'\tau_2)^{-13}|\eta(\tau)|^{-48}\,. \tag{7.3.6}$$

This amplitude has an important property: modular invariance. According to (7.2.44), $\tau_2|\eta(\tau)|^4$ is modular invariant, and it is easy to verify that

$$\frac{\mathrm{d}\tau\mathrm{d}\bar{\tau}}{\tau_2^2} \tag{7.3.7}$$

is also modular invariant. Note that the exponent -48 comes from 24 sets of left-moving and 24 sets of right-moving oscillators. Ghost contributions cancel two sets of oscillators, leaving only the transverse mode contributions. By introducing vertex operators into the expectation value (7.2.4) and integrating over position, we get the n-tachyon amplitude.

> - According to (7.2.9), $[Z_X(\tau)]^{26} = (4\pi^2\alpha'\tau_2)^{-13}|\eta(\tau)|^{-52}$, and according to (7.2.27), $\langle cb\tilde{c}\tilde{b}\rangle = |\eta(\tau)|^4$. Multiplying them gives (7.3.6).
> - Regarding modular invariance, the invariance under $\tau \to \tau+1$ is trivial. Under $\tau \to -1/\tau$, we have $\tau_2 \to \tau_2/|\tau|^2$ and $\mathrm{d}\tau \to (-\tau^{-2})\mathrm{d}\tau$, which gives the modular invariance of (7.3.7).

For a general CFT, the path integral on the torus is expressed by the trace in (7.2.28). As long as there are $d \geq 2$ non-compact flat dimensions, the ghost fields still cancel two sets of bosonic operators, and the vacuum amplitude is

$$Z_{T^2} = V_d \int_{F_0} \frac{\mathrm{d}\tau \mathrm{d}\bar{\tau}}{4\tau_2} \int \frac{\mathrm{d}^d k}{(2\pi)^d} \exp\left(-\pi\tau_2\alpha' k^2\right) \sum_{i\in\mathscr{H}^\perp} q^{h_i-1}\bar{q}^{\tilde{h}_i-1} \tag{7.3.8a}$$

$$= iV_d \int_{F_0} \frac{\mathrm{d}\tau \mathrm{d}\bar{\tau}}{4\tau_2} (4\pi^2\alpha'\tau_2)^{-d/2} \sum_{i\in\mathscr{H}^\perp} q^{h_i-1}\bar{q}^{\tilde{h}_i-1} \tag{7.3.8b}$$

Here $\mathscr{H}^\perp$ is the closed string Hilbert space excluding ghosts, $\mu = 0, 1$ oscillators, and non-compact momentum. To understand the physics of these amplitudes, it is useful to compare them with the corresponding quantities in field theory, summing over all paths with circle topology. This is given by

$$\begin{aligned} Z_{S_1}(m^2) &= V_d \int \frac{\mathrm{d}^d k}{(2\pi)^d} \int_0^\infty \frac{\mathrm{d}l}{2l} \exp\left[-(k^2+m^2)l/2\right] \\ &= iV_d \int_0^\infty \frac{\mathrm{d}l}{2l} (2\pi l)^{-d/2} \exp\left(-m^2 l/2\right) \end{aligned} \tag{7.3.9}$$

This can be derived through gauge-fixed point-particle path integrals. The result is intuitive: l is the modulus of the circle, $\frac{1}{2}(k^2+m^2)$ is the Hamiltonian of the worldline, and $2l$ removes redundant counting from translations and reversing worldline coordinates.

We now adopt the point-particle result for particles of mass m^2 and sum over the string spectrum. As stated in Chapter 4, the physical spectrum corresponds one-to-one with $\mathscr{H}^\perp$, where the relationship between mass and transverse weight is

$$m^2 = \frac{2}{\alpha'}(h + \tilde{h} - 2) \tag{7.3.10}$$

subject to the constraint $h = \tilde{h}$. This constraint can be written in integral form

$$\delta_{h,\tilde{h}} = \int_{-\pi}^{\pi} \frac{\mathrm{d}\theta}{2\pi} \exp\left[i(h-\tilde{h})\theta\right]. \tag{7.3.11}$$

In the above, we assume $h - \tilde{h}$ is an integer. This is true in 26 flat dimensions; as discussed in the previous section, modular invariance is generally important. Then

$$\begin{aligned} \sum_{i\in\mathscr{H}\perp} Z_{S_1}(m_i^2) &= iV_d \int_0^\infty \frac{\mathrm{d}l}{2l} \int_{-\pi}^{\pi} \frac{\mathrm{d}\theta}{2\pi} (2\pi l)^{-d/2} \\ &\qquad \times \sum_{i\in\mathscr{H}^\perp} \exp\left[-(h_i+\tilde{h}_i-2)l/\alpha' + i(h_i-\tilde{h}_i)\theta\right] \\ &= iV_d \int_R \frac{\mathrm{d}\tau \mathrm{d}\bar{\tau}}{4\tau_2} (4\pi^2\alpha'\tau_2)^{-d/2} \sum_{i\in\mathscr{H}\perp} q^{h_i-1}\bar{q}^{\tilde{h}_i-1}, \end{aligned} \tag{7.3.12}$$

where $\theta + il/\alpha' = 2\pi\tau$. Integration region R:

$$R: \quad \tau_2 > 0, \quad |\tau_1| < \frac{1}{2}. \tag{7.3.13}$$

We now interpret this result. A single point-particle amplitude (7.3.9) diverges as $l \to 0$. This is the standard UV divergence of quantum field theory. Summing over the string spectrum as in (7.3.12) only worsens the situation, as all string contributions have the same sign. However, comparing (7.3.12) with the actual string amplitude (7.3.8), they are very similar with one significant difference. The integration regions differ: in string theory, it is the fundamental domain

$$F_0: \quad |\tau| > 1, \quad |\tau_1| < \frac{1}{2}, \quad \tau_2 > 0 \tag{7.3.14}$$

The UV divergence region is bypassed. We can also see this from the momentum integral (7.3.8a).

Another possible divergence comes from the limit $\tau_2 \to \infty$, when the torus becomes very long. In this region, the string amplitude in 26 dimensions has the expansion

$$iV_{26} \int^{\infty} \frac{\mathrm{d}\tau_2}{2\tau_2} \left(4\pi^2\alpha'\tau_2\right)^{-13} \Big[\exp(4\pi\tau_2) + 24^2 + \cdots\Big] , \tag{7.3.15}$$

The asymptotic behavior is controlled by the lightest string states. The first term in the series diverges, and from the field theory perspective, the explanation is obvious: the series is arranged in order of increasing mass squared. The first term comes from the tachyon, and the divergence arises from the positive exponent in the path sum (7.3.9). This is a product of a theory containing tachyons and does not affect more realistic theories. Incidentally, this path sum can be defined by analytic continuation of positive mass squared, but problems remain—the continued energy density is complex. This signifies instability, where the tachyon field rolls down its inverted potential. For general CFT, (7.3.12) becomes

$$iV_d \int^{\infty} \frac{\mathrm{d}\tau_2}{2\tau_2} \left(4\pi^2\alpha'\tau_2\right)^{-d/2} \sum_i \exp\left(-\pi\alpha' m_i^2 \tau_2\right) . \tag{7.3.16}$$

Since tachyons always appear in bosonic string theory, this is divergent; without tachyons, it would be convergent.

The torus exemplifies a general principle that holds for all string amplitudes: there is no UV region in moduli space giving rise to high-energy divergences. All limits are controlled by the lightest string states, i.e., long-range physics. For the torus with vertex operators, the τ integral is still truncated as above. But more limits exist when vertex operators approach each other. The same general principle applies to them, which we will discuss further in Chapter **??**.

By truncating the l integral, we can attempt to remove UV divergences in field theory. Similarly, one can try to make similar corrections to string theory. For example, replacing the usual fundamental domain F_0 with $|\tau_1| \leq \frac{1}{2}$, $\tau_2 > 1$. However, in either case, this would destroy the consistency of the theory: non-physical negative-norm states will no longer decouple. We saw in the general discussion of section **??** that the coupling of BRST-null states is proportional to the total derivative on moduli space. For the fundamental domain F_0, the boundaries I and I$'$, and II and II$'$ are equivalent, so boundary terms cancel. Modifying the integration region would introduce real boundaries, and surface terms on moduli space would no longer cancel. Thus, there would be non-zero amplitudes for null states and an inconsistent quantum theory. If we make a brute-force truncation to make a gravity theory finite, this is a direct analogy of the consequence: it is extremely difficult to do so without destroying local spacetime symmetry and making the theory inconsistent. String theory provides a subtle way that softens short-distance behavior and eliminates divergences without losing spacetime gauge invariance.

Physics of the Vacuum Amplitude

Besides exemplifying the behavior of string amplitudes, the one-loop vacuum amplitude has an interesting physical interpretation. In point-particle theory, "vacuum" paths consist of any n non-intersecting circles. Permutation symmetry introduces a factor $1/n!$, and summing over n gives

$$\boldsymbol{Z}_{\text{vac}}(m^2) = \exp\Big[Z_{S_1}(m^2)\Big] . \tag{7.3.17}$$

Switching to canonical field theory,

$$\begin{aligned} Z_{\text{vac}}(m^2) &= \langle 0| \exp(-iHT) |0\rangle \\ &= \exp(-i\rho_0 V_d)\,, \end{aligned} \tag{7.3.18}$$

where ρ_0 is the vacuum energy density:

$$\rho_0 = \frac{i}{V_d} Z_{S_1}(m^2)\,. \tag{7.3.19}$$

The l integral in $Z_{S_1}(m^2)$ diverges as $l \to 0$. In renormalizable field theory, this is cancelled by counterterms or supersymmetry. We truncate the integral $l \geq \epsilon$ and then take $\epsilon \to 0$, but discard the divergent terms. Under this treatment,

$$\int_0^\infty \frac{\mathrm{d}l}{2l} \exp\big[-(k^2+m^2)l/2\big] \to -\frac{1}{2}\ln\big(k^2+m^2\big) \tag{7.3.20}$$

and

$$i\int_0^\infty \frac{\mathrm{d}l}{2l} \int_{-\infty}^{\infty} \frac{\mathrm{d}k^0}{2\pi} \exp\big[-(k^2+m^2)l/2\big] \to \frac{\omega_{\mathbf{k}}}{2}\,, \tag{7.3.21}$$

where $\omega_{\mathbf{k}}^2 = \mathbf{k}\cdot\mathbf{k} + m^2$. Using (7.3.21), the vacuum energy density becomes

$$\rho_0 = \int \frac{\mathrm{d}^{d-1}\mathbf{k}}{(2\pi)^{d-1}} \frac{\omega_{\mathbf{k}}}{2}\,, \tag{7.3.22}$$

which is exactly the sum of the zero-point energies of each mode of the field. We encountered a similar sum in section 1.3, though that was on the worldsheet.

Describing quantum field theory as a sum of particle paths may be less familiar than describing it as a sum over field histories, but they are equivalent. Specifically, the free path integral of field theory can be found in modern field theory textbooks:

$$\begin{aligned} \ln \boldsymbol{Z}_{\text{vac}}(m^2) &= -\frac{1}{2}\operatorname{Tr}\ln\big(-\partial^2 + m^2\big) \\ &= -\frac{V_d}{2}\int \frac{\mathrm{d}^d k}{(2\pi)^d} \ln\big(k^2+m^2\big)\,. \end{aligned} \tag{7.3.23}$$

Using (7.3.20), this is identical to the path sum result (7.3.9).

In older quantum field theory books, vacuum amplitudes like (7.3.18) were usually considered irrelevant. If the scattering amplitudes are considered in a fixed background and gravity is ignored, this is correct. But vacuum energy density is significant in at least two cases. First, to compare the energy densities of different states to determine which is the vacuum ground state. For example, electroweak $SU(2)\times U(1)$ symmetry breaking is likely determined by quantum corrections to vacuum energy density. This was the original motivation for the Coleman-Weinberg formula (7.3.23). Extending to arbitrary spin,

$$\rho_0 = \frac{i}{V_d} \sum_i (-1)^{\mathbf{F}_i} Z_{S_1}(m_i^2)\,. \tag{7.3.24}$$

The sum is taken over all physical particle states. Each polarization is counted separately, giving a factor $2s_i+1$ for spin s_i in 4 dimensions. Here $\mathbf{F}_i$ is the spacetime fermion number, so fermions contribute with the opposite sign.

The second case is consideration of coupling with gravity. Vacuum energy provides the source term in the Einstein equations—the cosmological constant—and thus has observable effects. In fact, this cosmological constant is a major challenge. Because real spacetime is approximately flat and static, its value must be very small:

$$|\rho_0| \lesssim 10^{-44}\mathrm{GeV}^4 \,. \tag{7.3.25}$$

If only the vacuum energy contribution of known particles is considered (roughly below the electroweak scale), the zero-point energy is already about

$$m_{\mathrm{ew}}^4 \approx 10^8\mathrm{GeV}^4 \,, \tag{7.3.26}$$

which is 52 orders of magnitude higher than reality. The potential energy of the Higgs field and QCD vacuum energy are also too large. Finding a mechanism that cancels the net cosmological constant with great precision has proven difficult. For example, in supersymmetric theories, the contributions of degenerate bosons and fermions in the sum (7.3.24) cancel. But supersymmetry has not yet been seen in nature, so it must be a broken symmetry. Then the cancellation is imperfect, again leaving at least m_{ew}^4 as a remainder. The cosmological constant problem is one of the most persistent difficulties. Thus, one of the best clues likely lies in finding a unified theory containing gravity.

What about the cosmological constant problem in string theory? At the string tree level, we have a consistent theory compatible with flat metrics, so the cosmological constant is zero. In fact, when we chose 26 dimensions, we adjusted it manually. From the spacetime action (**??**), it is evident that otherwise, there would be a tree potential proportional to $D-26$. The one-loop vacuum energy density in bosonic string amplitudes is non-zero and must be on the scale of the string (ignoring tachyon divergence). In 4 dimensions, this corresponds to $10^{72}\mathrm{GeV}^4$, which is still too large. In supersymmetric string theory, there will be some amount of cancellation, but in a realistic theory with broken supersymmetry, a remainder of at least m_{ew}^4 can still be expected.

The cosmological constant problem tells us that there is still something about the vacuum we do not understand, both in field theory and in string theory.

7.4 Open and Unoriented One-loop Amplitudes

Cylinder

The results for the torus can be immediately extended to other surfaces with zero Euler number. For example, the vacuum amplitude from the cylinder in oriented theory is

$$\begin{aligned} Z_{C_2} &= \int_0^\infty \frac{\mathrm{d}t}{2t}\,\mathrm{Tr}_{\mathrm{o}}'[\exp(-2\pi t L_0)] = iV_d \int_0^\infty \frac{\mathrm{d}t}{2t}\,(8\pi^2\alpha' t)^{-d/2} \sum_{i\in\mathscr{H}_0^\perp} \exp[-2\pi t(h_i-1)] \\ &\to iV_{26}n^2 \int_0^\infty \frac{\mathrm{d}t}{2t}\,(8\pi^2\alpha' t)^{-13}\eta(it)^{-24} \,. \end{aligned} \tag{7.4.1}$$

This can be obtained either by writing the path integral in terms of ghost zero modes, as we did for the torus, or by guessing it, summing the point-particle result over the open string spectrum. In the first line, the trace is over the entire open string CFT, and the prime indicates ghost zero modes are dropped. In the last line, the trace is over 26 flat dimensions and n Chan-Paton degrees of freedom. Introducing tachyon vertex operators is straightforward.

The $t \to \infty$ limit of the cylinder is very similar to the $\tau_2 \to \infty$ limit of the torus. The cylinder looks like a long strip, and the leading order approximation is given by the lightest open string states. As with closed strings, there are divergences, but only from the open string tachyon.

The $t \to 0$ limit is quite interesting. Unlike the torus, there is no modular group, and the integration limits cannot be truncated, so UV divergences of field theory still manifest. However,

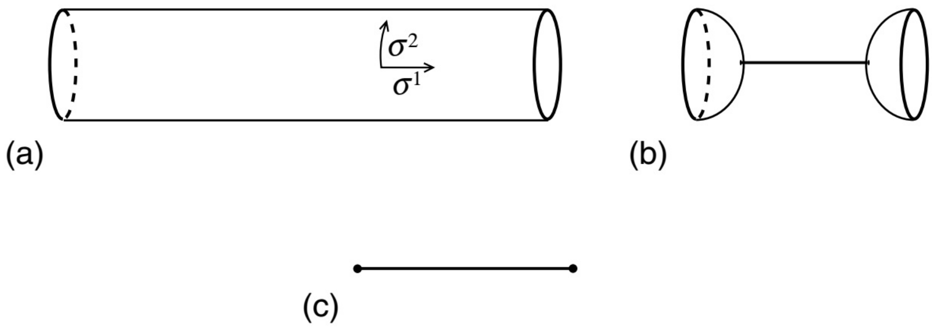


Figure 7.4.1: (a) Cylinder in the small t limit. (b) Amplitude decomposed into disk tadpole amplitudes and a closed string propagator. (c) Analogous field theory Feynman diagram. Heavy particles are analogous to tadpoles.

we will see that, like all divergences in string theory, this should actually be interpreted as a long-range effect. In the $t \to 0$ limit, this cylinder is very long, as shown in Figure 7.4.1a. It looks like a closed string emerging from vacuum, propagating a certain distance, and then disappearing back into the vacuum. To make it more obvious, using the modular transformation (7.2.44)

$$\eta(it) = t^{-1/2}\eta(i/t) \tag{7.4.2}$$

and changing the variable to $s = \pi/t$, the result is

$$Z_{C_2} = i\frac{V_{26}n^2}{2\pi(8\pi^2\alpha')^{13}}\int_0^\infty \mathrm{d}s\,\eta(is/\pi)^{-24}\,. \tag{7.4.3}$$

Rescaling the metric by $1/t$, the cylinder has a closed string circumference 2π, and the cylinder length is s. Expanding

$$\begin{aligned}\eta(is/\pi)^{-24} &= \exp(2s)\prod_{n=1}^{\infty}[1-\exp(-2ns)]^{-24}\\ &= \exp(2s) + 24 + O(\exp(-2s))\,,\end{aligned} \tag{7.4.4}$$

this is exactly the expected asymptotic form for an expansion in the complete closed string basis. In other words, if we consider σ^2 as worldsheet time and σ^1 as worldsheet distance, Figure 7.4.1a is a very short open string loop. If we swap their roles, it is a very long closed string worldline, with both the beginning and end on the boundary circles. In Euclidean path integrals, both descriptions are usable and useful in different limits of moduli space.

The leading order contribution to the vacuum amplitude comes from the closed string tachyon. This is uninteresting and can be defined through analytic continuation,

$$\int_0^\infty \mathrm{d}s\ \exp(\beta s) \equiv -\frac{1}{\beta}\,. \tag{7.4.5}$$

The second term comes from massless closed string states, which even with continuation give a divergence of the form $1/0$. To see the origin of this divergence, dissociate the process as in Figure 7.4.1b. As we saw in section 6.6, there is a non-zero amplitude for a closed string vertex operator on a disk. The tadpole diagram corresponds to a closed string disappearing or appearing in the vacuum. In momentum space, the massless propagator is proportional to $1/k^2$. Here, momentum conservation requires the momentum of the closed string emerging from vacuum to be zero, so the propagator diverges.

The same type of divergence exists in quantum field theory. Consider a massless scalar field ϕ with a linear term for ϕ in the Lagrangian. Then there will be a vertex connected to only one propagator, and Figure 7.4.1c exists. This divergence is due to the intermediate propagator being

$$\left.\frac{1}{k^2}\right|_{k^\mu=0} . \tag{7.4.6}$$

Since propagator poles correspond to long propagation distances in spacetime, this is a long-range (infrared) divergence. UV and IR divergences in quantum field theory have very different origins. UV divergences usually signal theory failure, requiring new physics at some short scale. IR divergences usually mean the question asked is wrong, or the method of expansion is wrong. This is the case here too. In general perturbation theory, we expand around $\phi(x) = 0$ or some other field configuration. For the action

$$-\frac{1}{g^2}\int \mathrm{d}^d x\, \left(\frac{1}{2}\partial_\mu\phi\partial^\mu\phi + g\Lambda\phi\right) , \tag{7.4.7}$$

the equation of motion

$$\partial^2\phi = g\Lambda \tag{7.4.8}$$

does not allow $\phi(x) = 0$ as a solution. We must expand around a solution of (7.4.8); any such solution must be position-dependent. The corresponding amplitude has no divergence, even if the right side is a perturbation (we introduced a suitable factor g to the disk), because it is a singular perturbation. Specifically, the correct background breaks some Poincaré symmetries of the zero-order solution.

The situation in string theory is the same. The disk tadpole is the source

$$-\Lambda\int \mathrm{d}^{26}x\,(-G)^{1/2}\mathrm{e}^{-\tilde{\Phi}} , \tag{7.4.9}$$

which is a source for the dilaton and metric. Expanding around the solution of the corresponding field equations (no longer constant) gives effective amplitudes. The details are somewhat complex and will be further discussed in Chapter **??**. Incidentally, in superstring theory, if the tree-level background is invariant under supersymmetry, it usually has no loop corrections.

The pole (7.4.6) is the same type of divergence we encountered in tree amplitudes, corresponding to resonances propagating over long spacetime distances. If we add open string vertex operators at each end of the cylinder so that it represents an open string one-loop amplitude, the momentum k^μ flowing from one boundary to another is generally non-zero. Then the large s limit (7.4.4) introduces the factor

$$\exp\bigl(-\alpha' k^2 s/2\bigr) , \tag{7.4.10}$$

and the divergence becomes a momentum pole, representing an open string scattering into a closed string intermediate state. Thus, as claimed in Chapter 3, open string theory must simultaneously incorporate closed strings. The mechanism for removing UV divergences is different in open and closed strings. In closed strings, it is an effective truncation on modular integrals; in open strings, in the form of long spacetime distances, it is reinterpreted as a dangerous limit of moduli space.

In (7.4.1), we associated the path integral on the cylinder with the open string spectrum by cutting the path integral with σ^2 as time. It can also be obtained through closed strings, where σ^1 is time. Let σ^2 have period 2π , with boundaries at $\sigma^1 = 0, s$. The closed string emerges in some state $|B\rangle$ at $\sigma^1 = 0$, and then disappears in the same way at $\sigma^1 = s$. Introducing a measure insertion, the path integral is proportional to

$$\langle B|c_0 b_0 \exp\Bigl[-s(L_0+\tilde{L}_0)\Bigr]|B\rangle . \tag{7.4.11}$$

The fact that $\partial_1 X^\mu, c^1$ and b_{12} are zero on the boundary determines the state $|B\rangle$. In Hamiltonian form, they must annihilate $|B\rangle$; in terms of Laurent coefficients,

$$(\alpha_n^\mu + \tilde{\alpha}_{-n}^\mu)|B\rangle = (c_n + \tilde{c}_{-n})|B\rangle = (b_n - \tilde{b}_{-n})|B\rangle = 0, \qquad \text{all } n\,. \tag{7.4.12}$$

This gives

$$|B\rangle \propto (c_0 + \tilde{c}_0)\exp\left[-\sum_{n=1}^{\infty}(n^{-1}\alpha_{-n}\cdot\tilde{\alpha}_{-n} + b_{-n}\tilde{c}_{-n} + \tilde{b}_{-n}c_{-n})\right]|0;0\rangle\,. \tag{7.4.13}$$

Using this in (7.4.11) yields (7.4.3), with the normalization of $|B\rangle$ yet to be determined. This representation is very useful for analyzing the $t \to 0$ limit and closed string poles.

By comparing string calculations with field theory, we can determine the disk tadpole Λ, but it will be more convenient to treat it as a special case of a more general result in the next chapter.

Klein Bottle

The vacuum amplitude from the Klein bottle is

$$\begin{aligned} Z_{K_2} &= \int_0^\infty \frac{\mathrm{d}t}{4t}\, \mathrm{Tr}_c'\Big\{\Omega \exp\Big[-2\pi t(L_0 + \tilde{L}_0)\Big]\Big\} \\ &= iV_d \int_0^\infty \frac{\mathrm{d}t}{4t}\,(4\pi^2\alpha' t)^{-d/2} \sum_{i\in\mathscr{H}_c^\perp} \Omega_i \exp\Big[-2\pi t(h_i + \tilde{h}_i - 2)\Big]\,, \end{aligned} \tag{7.4.14}$$

where the notation of the cylinder (7.4.1) is adopted. Compared with the torus of the oriented theory, this has an extra factor of $\frac{1}{2}$, coming from the projection operator $\frac{1}{2}(1+\Omega)$. For the same reason, the torus and cylinder amplitudes in unoriented theory both have an extra factor of $\frac{1}{2}$. This can also be seen as arising from the extra gauge invariance $w \to \bar{w}$. To calculate the trace in 26 flat dimensions, note that the only diagonal elements of Ω are those states where left-movers and right-movers are the same, and these states contribute $\Omega = +1$. Thus, the trace is actually taken only on one side, which is the same as the open string amplitude, except that the weight is doubled because left and right contributions are identical. The result is

$$Z_{K_2} \to iV_{26}\int_0^\infty \frac{\mathrm{d}t}{4t}\,(4\pi^2\alpha' t)^{-13}\eta(2it)^{-24}\,. \tag{7.4.15}$$

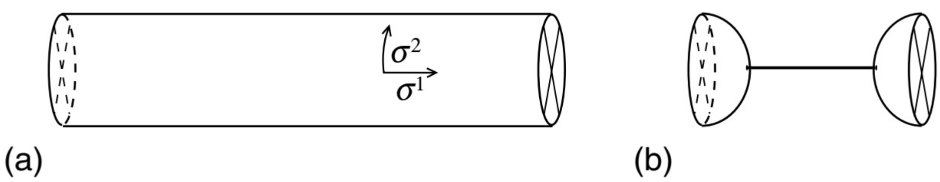


Figure 7.4.2: (a) In the small t limit, the Klein bottle becomes a cylinder with crosscaps at both ends. (b) Amplitude decomposed into RP_2 and D_2 tadpole amplitudes and a closed string propagator.

The modulus t has the same range $0 < t < \infty$ as for the cylinder, so the $t \to 0$ divergence is the same. Similarly, the $1/0$ pole can have a long-range explanation in the form of a closed string pole. To see this, consider the regions

$$0 \le \sigma^1 \le 2\pi, \qquad 0 \le \sigma^2 \le 2\pi t\,, \tag{7.4.16a}$$

$$0 \leq \sigma^1 \leq \pi, \qquad 0 \leq \sigma^2 \leq 4\pi t\,. \tag{7.4.16b}$$

Both are fundamental regions for the equivalence relation

$$w \cong w + 2\pi \cong -\bar{w} + 2\pi i t \tag{7.4.17}$$

In (7.4.16a), the left and right edges are periodically equivalent, while the top and bottom edges are equivalent after parity inversion, which explains the closed string loop with weight Ω. For (7.4.16b), the equivalence relation (7.4.17) simultaneously implies

$$w \cong w + 4\pi i t, \qquad w + \pi \cong -(\bar{w} + \pi) + 2\pi i t\,. \tag{7.4.18}$$

It follows that the upper and lower edges of the region (7.4.16b) are periodically equivalent. After translating half the length of the left edge, the inversion through that edge is equivalent to itself. Similarly for the right edge: this is the definition of a crosscap. Thus we have the interpretation of Figure 7.4.2a. Rescaling by $1/2t$, the cylinder base length is 2π, height $s = \pi/2t$, and both ends are connected to crosscaps. After modular transformation, the amplitude becomes

$$Z_{K_2} = i\frac{2^{26} V_{26}}{4\pi(8\pi^2\alpha')^{13}} \int_0^\infty \mathrm{d}s\, \eta(is/\pi)^{-24}\,. \tag{7.4.19}$$

All discussions regarding divergences of the cylinder can be applied to the Klein bottle. The difference is that the tadpole diagrams come from the projective plane rather than the disk.

Möbius Strip

For the Möbius strip,

$$Z_{M_2} = iV_d \int_0^\infty \frac{\mathrm{d}t}{4t}(8\pi^2\alpha' t)^{-d/2} \sum_{i \in \mathscr{H}_0^\perp} \Omega_i \exp[-2\pi t(h_i - 1)] \tag{7.4.20}$$

It differs from the cylinder result only by Ω and the $\frac{1}{2}$ of the projection operator. In 26 flat dimensions, the effect of the operator Ω in the trace is: there is an extra -1 on even mass levels, plus a proper count of Chan-Paton factors. Thus, the trace of oscillators is

$$\exp(2\pi t) \prod_{n=1}^{\infty} [1 - (-1)^n \exp(-2\pi n t)]^{-24} = \vartheta_{00}(0, 2it)^{-12} \eta(2it)^{-12}\,. \tag{7.4.21}$$

For $SO(n)$ theory, $\frac{1}{2}n(n+1)$ symmetric states have $\Omega = +1$, while $\frac{1}{2}n(n-1)$ antisymmetric states have $\Omega = -1$, giving a net contribution of n. For $Sp(k)$ theory, it is the opposite, giving $-n$ (in our notation $n = 2k$ Chan-Paton states). Then the amplitude is

$$Z_{M_2} = \pm i n V_{26} \int_0^\infty \frac{\mathrm{d}t}{4t}\,(8\pi^2\alpha' t)^{-13} \vartheta_{00}(0, 2it)^{-12} \eta(2it)^{-12}\,. \tag{7.4.22}$$

By the same construction as for the Klein bottle, the Möbius strip can also be represented as a cylinder with a base, where only one end is a crosscap, as shown in Figure 7.4.3a. The cylinder length is now $s = \pi/4t$. Through modular transformation, the amplitude is

$$Z_{M_2} = \pm 2in \frac{2^{13} V_{26}}{4\pi(8\pi^2\alpha')^{13}} \int_0^\infty \mathrm{d}s\, \vartheta_{00}(0, 2is/\pi)^{-12} \eta(2is/\pi)^{-12}\,. \tag{7.4.23}$$

As with the cylinder, this can also be written as an operator expression, with one boundary state in (7.4.11) replaced by a similar crosscap state $|C\rangle$. This is also a $t \to 0$ divergence; it

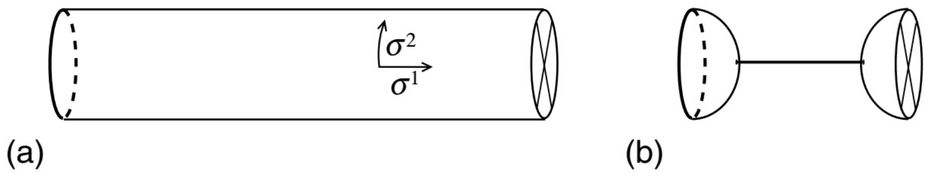


Figure 7.4.3: (a) In the small t limit, the Möbius strip becomes a cylinder with a crosscap at one end. (b) Amplitude decomposed into D_2 and RP_2 tadpole amplitudes and a closed string propagator.

corresponds to the process in Figure 7.4.3a, with one end from the disk and the other from the projective plane. In unoriented theory, the divergences of these three surfaces combine into

$$i\frac{24V_{26}}{4\pi(8\pi^2\alpha')^{13}}(2^{13}\mp n)^2\int_0^\infty \mathrm{d}s\,. \tag{7.4.24}$$

That is, the total tadpole is proportional to $2^{13}\mp n$. For the gauge group $SO(2^{13})=SO(8192)$, this is zero. The tadpole from the disk and the tadpole from the projective plane cancel each other. In bosonic string theory, this might not have special meaning, but in superstrings, there is indeed a similar cancellation for $SO(32)$.

Chapter 8

Toroidal Compactification and T-duality

The realistic compactification of string theory will be the subject of the final chapters of Volume II, but let us first examine the simplest compactifications of string theory, where one or more dimensions are periodically identified. We first examine the same compactifications in field theory, which are encountered in the Kaluza-Klein unification of gauge interactions and gravity. Then we extend this to string theory, where several new and intrinsic string phenomena arise: winding states, enhanced gauge symmetry, T-duality, and D-branes. We will also examine slightly more complex compactifications along this line, namely orbifolds and orientifolds.

8.1 Toroidal Compactification in Field Theory

In general relativity, the geometry of spacetime is dynamical. The three spatial dimensions we see are extended, but they might have once been highly curled up. There is the logical possibility that there are still extra dimensions that remain very small. In fact, this was proposed as early as 1914 as a way to unify the electromagnetic and gravitational fields into components of a single higher-dimensional field. Consider the 5-dimensional case, where x^4 is periodically identified:

$$x^4 \cong x^4 + 2\pi R\,, \tag{8.1.1}$$

while the other x^μ for $\mu = 0, \ldots, 3$ are non-compact. This is toroidal compactification, where the 5-dimensional metric splits into $G_{\mu\nu}, G_{\mu 4}$, and G_{44}. From a 4-dimensional perspective, these are a metric, a vector, and a scalar.

Let us examine this carefully: take the more general case of $D = d+1$ spacetime dimensions, where x^d is periodic. Parameterize the metric as:

$$\mathrm{d}s^2 = G^D_{MN}\mathrm{d}x^M\mathrm{d}x^N = G_{\mu\nu}\mathrm{d}x^\mu\mathrm{d}x^\nu + G_{dd}(\mathrm{d}x^d + A_\mu\mathrm{d}x^\mu)^2\,. \tag{8.1.2}$$

Denote the metric of the entire D-dimensional spacetime as G^D_{MN}, and note that $G_{\mu\nu} \neq G^D_{\mu\nu}$; we use $G_{\mu\nu}$ to raise and lower indices in the d-dimensional action. For now, only allow $G_{\mu\nu}, G_{dd}$ and A_μ to depend on x^μ. (8.1.2) is the most general form of the metric that is invariant under translations of x^d. This form still allows for d-dimensional reparameterizations $x'^\mu(x^\nu)$ and:

$$x'^d = x^d + \lambda(x^\mu)\,. \tag{8.1.3}$$

Under this reparameterization:

$$A'_\mu = A_\mu - \partial_\mu\lambda\,, \tag{8.1.4}$$

so gauge transformations arise as part of the higher-dimensional coordinate group; this is the Kaluza-Klein mechanism. To see the effect of the x^d dependence, consider a massless scalar ϕ

in D dimensions. For simplicity, take the metric in the compact dimension to be $G_{dd} = 1$. The momentum in the periodic dimension is quantized, $p_d = n/R$. Expand the dependence of ϕ on x^d in a complete basis:

$$\phi(x^M) = \sum_{n=-\infty}^{\infty} \phi_n(x^\mu) \exp\Big(inx^d/R\Big) . \tag{8.1.5}$$

The D-dimensional wave equation $\partial_M \partial^M \phi = 0$ becomes:

$$\partial_\mu \partial^\mu \phi_n(x^\mu) = \frac{n^2}{R^2} \phi_n(x^\mu) . \tag{8.1.6}$$

The modes ϕ_n of the D-dimensional field thus become an infinite tower of d-dimensional fields. For fields where p_d is non-zero, the d-dimensional mass squared:

$$-p^\mu p_\mu = \frac{n^2}{R^2} \tag{8.1.7}$$

is non-zero. When the energy is much smaller than R^{-1}, only the x^d independence remains, and the physics is d-dimensional. When the energy exceeds R^{-1}, we see the Kaluza-Klein tower of states. The charge corresponding to the Kaluza-Klein gauge invariance (8.1.3) is the p_d momentum. In this simple example, all fields carrying Kaluza-Klein charge are massive. More generally, if there are also higher-spin fields and a curved background, there are also massless charged fields.

The effective action of massless fields is always an important topic. Define $G_{dd} = \mathrm{e}^{2\sigma}$. The Ricci scalar of the metric (8.1.2) is:

$$\boldsymbol{R} = \boldsymbol{R}_d - 2\mathrm{e}^{-\sigma}\nabla^2 \mathrm{e}^{\sigma} - \frac{1}{4}\mathrm{e}^{2\sigma} F_{\mu\nu}F^{\mu\nu} , \tag{8.1.8}$$

where $\boldsymbol{R}$ is constructed with G^D_{MN}, and $\boldsymbol{R}_d$ is constructed with $G_{\mu\nu}$. The graviton-dilaton action (**??**) becomes:

$$\begin{aligned}
\boldsymbol{S}_1 &= \frac{1}{2\kappa_0^2}\int \mathrm{d}^D x\, (-G_D)^{1/2} \mathrm{e}^{-2\Phi} (\boldsymbol{R} + 4\nabla_\mu \Phi \nabla^\mu \Phi) \\
&= \frac{\pi R}{\kappa_0^2} \int \mathrm{d}^d x\, (-G_d)^{1/2} \mathrm{e}^{-2\Phi+\sigma} \\
&\qquad \times \left(\boldsymbol{R}_d - 4\partial_\mu \Phi \partial^\mu \sigma + 4\partial_\mu \Phi \partial^\mu \Phi - \frac{1}{4}\mathrm{e}^{2\sigma} F_{\mu\nu}F^{\mu\nu} \right) \\
&= \frac{\pi R}{\kappa_0^2} \int \mathrm{d}^d x\, (-G_d)^{1/2} \mathrm{e}^{-2\Phi_d} \\
&\qquad \times \left(\boldsymbol{R}_d - \partial_\mu \sigma \partial^\mu \sigma + 4\partial_\mu \Phi_d \partial^\mu \Phi_d - \frac{1}{4}\mathrm{e}^{2\sigma} F_{\mu\nu}F^{\mu\nu} \right) ,
\end{aligned} \tag{8.1.9}$$

which gives the kinetic terms for all massless fields. Here G_d is the determinant of $G_{\mu\nu}$, and $\Phi_d = \Phi - \sigma/2$ is the effective d-dimensional dilaton. The sign error of the dilaton kinetic term is fictitious, because the mixing of the graviton with the metric trace must also be considered. The simplest way to achieve this is through the Weyl transformation in (**??**).

The field equations do not determine the radius of the compact dimension. For any values of Φ and σ, the flat metric and constant dilaton are solutions. In other words, Φ and σ have no potential, so these fields must be massless, much like Goldstone bosons. Different values of Φ and σ mark degenerate configurations (or states of the quantum theory), and if the field in the state varies slowly, its energy comes only from the gradient. It differs from the Goldstone phenomenon in that the degenerate states are not related to each other by any symmetry. In this bosonic theory, the degeneracy is accidental, and the one-loop energy discussed in the previous section breaks this degeneracy. In supersymmetric theories, the existence of physically inequivalent but degenerate vacua is quite common and plays a significant role in understanding dynamics. The fields used to mark inequivalent vacua are called moduli. In nature, supersymmetry breaking almost necessarily endows all moduli with mass; otherwise, they would transmit infinite long-range forces of gravitational magnitude.

Defining $A_\mu = R\tilde{A}_\mu$, the covariant derivative is:

$$\partial_\mu + ip_d A_\mu = \partial_\mu + in\tilde{A}_\mu \, , \tag{8.1.10}$$

which makes the charge an integer. The definitions of d-dimensional gauge coupling and gravitational coupling are as follows. The coefficient of $\tilde{F}_{\mu\nu}\tilde{F}^{\mu\nu}$ in the Lagrangian density is defined as $-1/4g_d^2$, and the coefficient of $\boldsymbol{R}_d$ is defined as $1/2\kappa_d^2$. In terms of gravitational coupling, the gauge coupling is written as:

$$g_d^2 = \frac{\kappa_0^2 \mathrm{e}^{2\Phi_d}}{\pi R^3 \mathrm{e}^{2\sigma}} = \frac{2\kappa_d^2}{\rho^2} \, . \tag{8.1.11}$$

The relationship between d-dimensional and D-dimensional gravitational couplings is:

$$\frac{1}{\kappa_d^2} = \frac{2\pi\rho}{\kappa^2} \, , \tag{8.1.12}$$

where $2\pi\rho$ is the volume of the compact dimension.

Remark.

- *$1/g_d^2$ has a factor of R^2 contributed by the definition $A_\mu = R\tilde{A}_\mu$. Other contributions come from (8.1.9), and combined with $\rho = R\mathrm{e}^\sigma$, this gives (8.1.11).*
- *According to (8.1.9) and the definition, $1/2\kappa_d^2 = \frac{\pi R \mathrm{e}^{-2\Phi_d}}{\kappa_0^2}$. The D-dimensional gravitational coupling is $\frac{1}{2\kappa^2} = \frac{\mathrm{e}^{-2\Phi}}{2\kappa_0^2}$, which yields (8.1.12).*

By extending the Kaluza-Klein mechanism, antisymmetric tensors can also provide gauge symmetry. Divide B_{MN} into $B_{\mu\nu}$ and $A'_\mu = B_{d\mu}$. The gauge parameter ζ_M in (**??**) is divided into the d-dimensional antisymmetric tensor transformation ζ_μ and the ordinary gauge invariant ζ_d. The gauge field is $B_{d\mu}$, and the field strength is $H_{d\mu\nu}$. The antisymmetric tensor action becomes:

$$\begin{aligned}
\boldsymbol{S}_2 &= -\frac{1}{24\kappa_0^2}\int \mathrm{d}^D x \, (-G_D)^{1/2} \mathrm{e}^{-2\Phi} H_{MNL}H^{MNL} \\
&= -\frac{\pi R}{12\kappa_0^2}\int \mathrm{d}^d x \, (-G_d)^{1/2} \mathrm{e}^{-2\Phi_d}\left(\tilde{H}_{\mu\nu\lambda}\tilde{H}^{\mu\nu\lambda} + 3\mathrm{e}^{-2\sigma} H_{d\mu\nu}H_d^{\mu\nu}\right) .
\end{aligned} \tag{8.1.13}$$

We defined:

$$\tilde{H}_{\mu\nu\lambda} = (\partial_\mu B_{\nu\lambda} - A_\mu H_{d\nu\lambda}) + \text{cyclic permutations} \, . \tag{8.1.14}$$

The term proportional to the vector potential originates from the inverse metric G^{MN}. It is called a Chern-Simons term, which represents a gauge potential coupled to any number of field strengths. Such terms frequently appear in supersymmetric theories and are connected with interesting physical phenomena. Note that $\tilde{H}_{\mu\nu\lambda}$ is gauge invariant because the variation of A_μ in (8.1.4) is cancelled by the following variation:

$$B'_{\nu\lambda} = B_{\nu\lambda} - \lambda H_{d\nu\lambda} \, . \tag{8.1.15}$$

Minimal coupling between B_{MN} and other fields does not exist, so it is unlike the Kaluza-Klein case. No field carries charge under antisymmetric tensor gauge symmetry; this is different in string theory.

Remark. *The inverse metric is:*

$$g^{MN} = \begin{bmatrix} g^{\mu\nu} & -A^{\mu} \\ -A^{\nu} & g_{\mu\nu}A^{\mu}A^{\nu} + \mathrm{e}^{-2\sigma} \end{bmatrix} .$$

8.2 Toroidal Compactification in CFT

Now let us examine the conformal field theory of a single periodic scalar field:

$$X \cong X + 2\pi R . \tag{8.2.1}$$

To keep the equations clean, we drop the superscript d of X^d and set $G_{dd} = 1$. The worldsheet action is the same as in the non-compact theory, $\int \mathrm{d}^2 z\, \partial X \bar{\partial} X / 2\pi\alpha'$, so the equations of motion, operator products, and energy-momentum tensor remain unchanged. Periodicity has two effects. First, string states must be single-valued under the equivalence relation (8.2.1). That is, the operator translating the string $\exp(2\pi i p)$ around the compact dimension once must leave the state invariant, so the center-of-mass momentum is quantized:

$$k = \frac{n}{R}, \quad n \in \mathbb{Z} . \tag{8.2.2}$$

This is the same as in field theory.

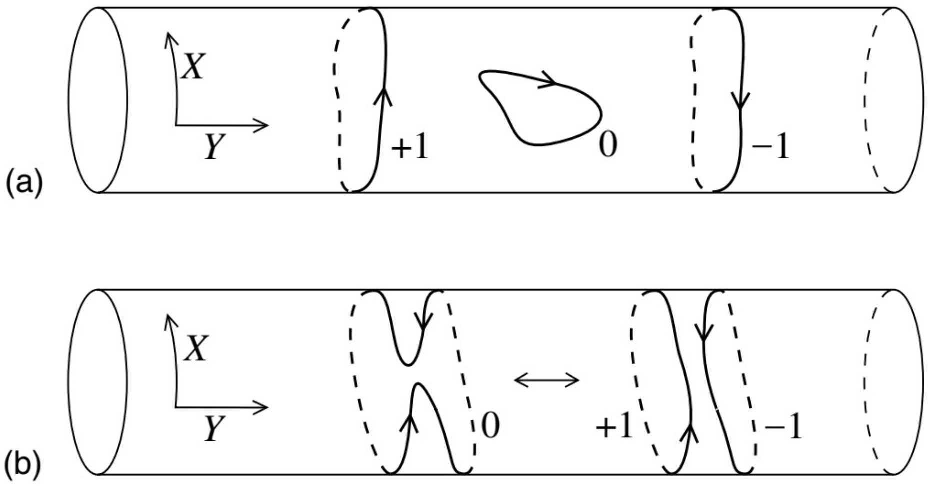


Figure 8.2.1: (a) Oriented closed strings with winding numbers $w = +1, 0, -1$. (b) Transition of a $w = 0$ string into $w = +1$ and $w = -1$ strings.

The second effect is unique to string theory. Closed strings can now wind around the compact dimension:

$$X(\sigma + 2\pi) = X(\sigma) + 2\pi R w, \quad w \in \mathbb{Z} . \tag{8.2.3}$$

The integer w is the winding number. States with winding numbers $+1, 0, -1$ are shown in Figure 8.2.1(a). From the viewpoint of worldsheet field theory, strings with non-zero winding number are topological solitons, i.e., states with topologically non-trivial field configurations. Consistent string theory must incorporate winding number states: through split-merge processes, $w = 0$ strings can become $w = +1$ and $w = -1$ strings, as shown in Figure 8.2.1(b). It is easy to see that winding number is always conserved in this example. To determine the states of the closed string CFT, examine the Laurent expansion:

$$\partial X(z) = -i\left(\frac{\alpha'}{2}\right)^{1/2}\sum_{m=-\infty}^{\infty}\frac{\alpha_m}{z^{m+1}}\,, \qquad \bar\partial X(\bar z) = -i\left(\frac{\alpha'}{2}\right)^{1/2}\sum_{m=-\infty}^{\infty}\frac{\tilde\alpha_m}{\bar z^{m+1}}\,. \tag{8.2.4}$$

After winding once, the change in X is:

$$2\pi R w = \oint (\mathrm{d}z\,\partial X + \mathrm{d}\bar z\,\bar\partial X) = 2\pi(\alpha'/2)^{1/2}(\alpha_0 - \tilde\alpha_0)\,. \tag{8.2.5}$$

The total Noether momentum is:

$$p = \frac{1}{2\pi\alpha'}\oint (\mathrm{d}z\,\partial X - \mathrm{d}\bar z\,\bar\partial X) = (2\alpha')^{-1/2}(\alpha_0 + \tilde\alpha_0)\,. \tag{8.2.6}$$

For non-compact dimensions, this gives $\alpha_0 = \tilde\alpha_0 = p(\alpha'/2)^{1/2}$ as usual, but for periodic dimensions we have:

$$p_L \equiv \left(2/\alpha'\right)^{1/2}\alpha_0 = \frac{n}{R} + \frac{wR}{\alpha'}\,, \tag{8.2.7a}$$

$$p_R \equiv \left(2/\alpha'\right)^{1/2}\tilde\alpha_0 = \frac{n}{R} - \frac{wR}{\alpha'}\,. \tag{8.2.7b}$$

The Virasoro generators are:

$$L_0 = \frac{\alpha' p_L^2}{4} + \sum_{n=1}^{\infty}\alpha_{-n}\alpha_n\,, \tag{8.2.8a}$$

$$\tilde L_0 = \frac{\alpha' p_R^2}{4} + \sum_{n=1}^{\infty}\tilde\alpha_{-n}\tilde\alpha_n\,. \tag{8.2.8b}$$

Partition Function

The partition function for X is:

$$\begin{aligned}
&(q\bar q)^{-1/24}\,\mathrm{Tr}\Big(q^{L_0}\bar q^{\tilde L_0}\Big)\\
&\quad= |\eta(\tau)|^{-2}\sum_{n,w=-\infty}^{\infty} q^{\alpha' p_L^2/4}\bar q^{\alpha' p_R^2/4}\\
&\quad= |\eta(\tau)|^{-2}\sum_{n,w=-\infty}^{\infty}\exp\left[-\pi\tau_2\left(\frac{\alpha' n^2}{R^2} + \frac{w^2R^2}{\alpha'}\right) + 2\pi i\tau_1 n w\right]\,.
\end{aligned} \tag{8.2.9}$$

The oscillators are the same as in the non-compact case, while the momentum integral is replaced by a sum over n and w. Modular invariance is not obvious, but using the Poisson resummation formula:

$$\sum_{n=-\infty}^{\infty}\exp\big(-\pi a n^2 + 2\pi i b n\big) = a^{-1/2}\sum_{m=-\infty}^{\infty}\exp\Big[-\frac{\pi(m-b)^2}{a}\Big]\,. \tag{8.2.10}$$

The partition function becomes:

$$2\pi R Z_X(\tau)\sum_{m,w=-\infty}^{\infty}\exp\Big(-\frac{\pi R^2|m - w\tau|^2}{\alpha'\tau_2}\Big)\,. \tag{8.2.11}$$

$Z_X(\tau)$ is the modular invariant expression (7.2.9) for the non-compact theory; this sum is clearly invariant under $\tau\to\tau+1$. Under $\tau\to -1/\tau$, it is invariant under $m\to -w$ and $w\to m$.

Remark. *Under $\tau \to -1/\tau$:*

$$\frac{|m - w\tau|^2}{\tau_2} \to \frac{|m + w/\tau|^2}{\tau_2/|\tau|^2} = \frac{|m\tau + w|^2}{\tau_2}$$

(8.2.11) has a simple path integral interpretation. Summing over all genus 1 worldsheets in periodic spacetime, each non-trivial closed curve on the worldsheet winds around the compact direction:

$$X(\sigma^1 + 2\pi, \sigma^2) = X(\sigma^1, \sigma^2) + 2\pi w R\,, \tag{8.2.12a}$$
$$X(\sigma^1 + 2\pi\tau_1, \sigma^2 + 2\pi\tau_2) = X(\sigma^1, \sigma^2) + 2\pi m R\,. \tag{8.2.12b}$$

That is, the path integral splits into topologically inequivalent sectors labeled by w and m. By writing X as a classical solution satisfying periodicity:

$$X_{\text{cl}} = \sigma^1 w R + \sigma^2 (m - w\tau_1) R/\tau_2\,, \tag{8.2.13}$$

plus a quantum part satisfying periodic boundary conditions. This Gaussian path integral can be integrated out. The path integral for the quantum part is the same as in the non-compact case, while the classical part appears as the exponential in (8.2.11). The effect of modular transformations on periodicity is simplified to swapping the summation variables m and w.

Vertex Operators

To construct winding states with $\alpha_0 \neq \tilde{\alpha}_0$, we need independent variables x_L and x_R satisfying:

$$[x_L, p_L] = [x_R, p_R] = i\,. \tag{8.2.14}$$

The field X splits into holomorphic and anti-holomorphic parts:

$$X(z, \bar{z}) = X_L(z) + X_R(\bar{z})\,, \tag{8.2.15}$$

where:

$$X_L(z) = x_L - i\frac{\alpha'}{2} p_L \ln z + i\left(\frac{\alpha'}{2}\right)^{1/2} \sum_{\substack{m=-\infty \\ m\neq 0}}^{\infty} \frac{\alpha_m}{m z^m}\,, \tag{8.2.16a}$$
$$X_R(\bar{z}) = x_R - i\frac{\alpha'}{2} p_R \ln \bar{z} + i\left(\frac{\alpha'}{2}\right)^{1/2} \sum_{\substack{m=-\infty \\ m\neq 0}}^{\infty} \frac{\tilde{\alpha}_m}{m \bar{z}^m}\,. \tag{8.2.16b}$$

If we restrict to states with $k_L = k_R$ and only use operators constructed from $X_L(z) + X_R(\bar{z})$, this reduces to the CFT of non-compact dimensions.

We now discuss OPEs and vertex operators. The form of the vertex operator is easily guessed, since the usual XX operator product $-(\alpha'/2)\ln(z_{12}\bar{z}_{12})$ can be split into holomorphic and anti-holomorphic functions, thus:

$$X_L(z_1) X_L(z_2) \sim -\frac{\alpha'}{2} \ln z_{12}, \quad X_R(\bar{z}_1) X_R(\bar{z}_2) \sim -\frac{\alpha'}{2} \ln \bar{z}_{12}\,, \tag{8.2.17a}$$
$$X_L(z_1) X_R(\bar{z}_2) \sim 0\,. \tag{8.2.17b}$$

The operator corresponding to the state $|0; k_L, k_R\rangle$ is:

$$\mathscr{V}_{k_L k_R}(z, \bar{z}) = :e^{ik_L X_L(z) + ik_R X_R(\bar{z})}: \tag{8.2.18}$$

The OPE is:

$$\mathscr{V}_{k_L k_R}(z_1, \bar{z}_1)\mathscr{V}_{k'_L k'_R}(z_2, \bar{z}_2) \sim z_{12}^{\alpha' k_L k'_L/2} \bar{z}_{12}^{\alpha' k_R k'_R/2} \mathscr{V}_{(k+k')_L (k+k')_R}(z_2, \bar{z}_2)\,. \tag{8.2.19}$$

We will encounter branch points and cuts in various expressions; this is not surprising, as the field X is no longer single-valued. However, it is important that the OPE of the entire vertex operator must be single-valued: when z_1 winds around z_2 once, the net phase is 1.

$$\exp\big[\pi i\alpha'(k_L k'_L - k_R k'_R)\big] = \exp\big[2\pi i(nw' + wn')\big] = 1\,. \tag{8.2.20}$$

This is exactly what is necessary for the string amplitude to be well-defined.

A Trick

The previous discussion is clearly correct, but we have been too casual about the handling of branch points and cuts in the logarithms. If we swap z_1 and z_2, and the momenta k and k', a problem arises in OPE (8.2.19). The left side is symmetric, but the right side will differ by $\exp[\pi i(nw' + wn')]$. Therefore, if $nw' + wn'$ is odd, there will be a negative sign difference. We can also see this problem in the following way. From the mode expansion, we can derive the equal-time ($|z_1| = |z_2|$) commutator:

$$[X_L(z_1), X_L(z_2)] = \frac{\pi i\alpha'}{2}\,\mathrm{sign}(\sigma_1^1 - \sigma_2^1)\,. \tag{8.2.21}$$

The CBH formula then tells us that if we define the operator through a creation-annihilation order, then when $nw' + wn'$ is odd, the operator $\mathscr{V}_{k_L k_R}$ anticommutes with $\mathscr{V}_{k'_k k'_R}$. The invisible branch point/cut (8.2.20) becomes two visible ones. The correct oscillator expression for the vertex operator is (with some ambiguity in the phase):

$$\mathscr{V}_{k_L k_R}(z, \bar{z}) = \exp\big[\pi i(k_L - k_R)(p_L + p_R)\alpha'/4\big] : \mathrm{e}^{ik_L X_L(z) + ik_R X_R(\bar{z})} :\,, \tag{8.2.22}$$

where p_L, p_R are operators as usual, and k is a number (the momentum carried by the given vertex operator). When $\mathscr{V}_{k_L k_R}$ and $\mathscr{V}_{k'_L k'_R}$ commute with each other, the extra factor is called cocycles, which gives an extra phase:

$$\begin{aligned}\exp\Big\{\pi i\Big[(k_L - k_R)(k'_L + k'_R) - (k'_L - k'_R)(k_L + k_R)\Big]\alpha'/4\Big\}&\\ = \exp\big[\pi i(nw' - wn')\big]\,,&\end{aligned} \tag{8.2.23}$$

which removes the branch points and cuts. In most cases, this complexity can be ignored, and the simpler expression from the previous paragraph can be used; cocycles only affect the relative signs of specific amplitudes.

The general X^μ path integral (6.2.18) factors into holomorphic and anti-holomorphic parts in an obvious way, allowing for the substitution:

$$\prod_{i<j}^{n} |z_{ij}|^{\alpha' k_i k_j} \to \prod_{i<j}^{n} z_{ij}^{\alpha' k_{Li} k_{Lj}/2} \bar{z}_{ij}^{\alpha' k_{Ri} k_{Rj}/2}\,, \tag{8.2.24}$$

and we can also replace the $2\pi\delta(\sum k)$ from the non-compact integral over x_0 with:

$$2\pi R\delta_{\Sigma_i n_i, 0}\delta_{\Sigma_i w_i, 0}\,. \tag{8.2.25}$$

The exact expression (8.2.22) will yield some additional signs.

DDF Operators

We have seen that exponential operators can be split into holomorphic and anti-holomorphic parts, which has many applications. We now use this to discuss the DDF operators mentioned at the end of Chapter 4.

We will work with light-cone coordinates $X^{\pm} = 2^{-1/2}(X^0 \pm X^1)$, whose holomorphic part has the OPE (dropping the L subscript):

$$X^+(z)X^-(0) \sim \frac{\alpha'}{2}\ln z, \qquad X^+(z)X^+(0) \sim X^-(z)X^-(0) \sim 0\,. \tag{8.2.26}$$

Remark.

$$X^+(z)X^-(0) = 2^{-1}\left(X^0(z) + X'(z)\right)\left(X^0(0) - X'(0)\right) = 2^{-1}\left(\frac{\alpha'}{2}\ln z_{12} + \frac{\alpha'}{2}\ln z_{12}\right) = \frac{\alpha'}{2}\ln z_{12}.$$

From this, it follows that the operator:

$$V^i(nk_0, z) = \partial X^i(z)\mathrm{e}^{ink_0X^+(z)}(2/\alpha')^{1/2} \tag{8.2.27}$$

is a $(1,0)$ primary field, where i is a transverse index. The OPE is:

$$\begin{aligned} V^i(nk_0, z)V^j(mk_0, 0) \sim & -\frac{\delta^{ij}}{z^2}\mathrm{e}^{i(n+m)k_0X^+(0)} \\ & - \frac{ink_0\delta^{ij}}{z}\partial X^+(0)\mathrm{e}^{i(n+m)k_0X^+(0)}\,. \end{aligned} \tag{8.2.28}$$

Define the DDF operator as:

$$A_n^i = \oint \frac{\mathrm{d}z}{2\pi}\, V^i(nk_0, z)\,. \tag{8.2.29}$$

Unless $n + m = 0$, the z^{-1} term in OPE (8.2.27) is a total derivative, so the DDF operators satisfy the oscillator algebra:

$$[A_m^i, A_n^j] = m\delta^{ij}\delta_{m,-n}\frac{\alpha' k_0 p^+}{2}\,. \tag{8.2.30}$$

Examine the action of this operator on a state with given momentum q. The vertex operator for $V^i(nk_0, z)$ and this state will contain $z^{-\alpha' nk_0 q^-/2}$. Thus, if we fix $k_0 = 2/\alpha' q^-$, it will be single-valued. Then the contour integral (8.2.28) defines a reasonable operator in this sector. The DDF operator is the integral of a $(1,0)$ tensor and thus commutes with the Virasoro generators, transforming physical states into physical states. From this, we can build physical states by acting on the oscillator ground state with DDF operators. The states obtained in this way correspond one-to-one with light-cone states. DDF operators do not contain X^-, so these states have no α^-_{-m} excitations. Thus, compared to the more general but less obvious construction in (4.4.19), they actually yield the same states.

8.3 Closed Strings and T-duality

Since we are studying string theory, we take $D = 26$ and only assume X^{25} is periodic. The mass shell (L_0 and $\tilde{L}_0$) is:

$$\begin{aligned} m^2 = -k^\mu k_\mu &= (k_L^{25})^2 + \frac{4}{\alpha'}(N-1) \\ &= (k_R^{25})^2 + \frac{4}{\alpha'}(\tilde{N}-1)\,, \end{aligned} \tag{8.3.1}$$

or:

$$m^2 = \frac{n^2}{R^2} + \frac{w^2R^2}{\alpha'^2} + \frac{2}{\alpha'}(N + \tilde{N} - 2)\,, \tag{8.3.2a}$$

$$0 = nw + N - \tilde{N}\,. \tag{8.3.2b}$$

There are four contributions to mass squared: compact momentum, winding string potential energy, oscillators, and zero-point energy. The discussion of the physical spectrum in Chapter 4 applies to any compactification, so we obtain the correct counting of states just by examining the transverse oscillators $M = 2, \dots, 25$.

First, we reproduce the field theory result for the massless spectrum. At a general value of R, states can only be massless when $n = w = 0$ and $N = \tilde{N} = 1$. There are 24^2 identical states, but it is useful to classify them based on whether the oscillators are in the spacetime direction μ or the internal direction 25:

$$\begin{aligned} &\alpha^{\mu}_{-1}\tilde{\alpha}^{\nu}_{-1}|0;k\rangle\,, \qquad (\alpha^{\mu}_{-1}\tilde{\alpha}^{25}_{-1} + \alpha^{25}_{-1}\tilde{\alpha}^{\mu}_{-1})|0;k\rangle\,, \\ &(\alpha^{\mu}_{-1}\tilde{\alpha}^{25}_{-1} - \alpha^{25}_{-1}\tilde{\alpha}^{\mu}_{-1})|0;k\rangle\,, \qquad \alpha^{25}_{-1}\tilde{\alpha}^{25}_{-1}|0;k\rangle\,. \end{aligned} \tag{8.3.3}$$

The first set further splits into the 25-dimensional graviton plus dilaton plus antisymmetric tensor. The second is the graviton on spacetime and internal indices, which is the Kaluza-Klein vector. The third is the vector from the antisymmetric tensor. The last state is a scalar, which is the modulus of the radius in the compact direction; its vertex operator $:\partial X^{25}\bar{\partial}X^{25}\mathrm{e}^{ik\cdot X}:$ is a perturbation of the metric $G_{25,25}$. This is the same spectrum found when examining low-energy field theory.

Examining the transformation of massive states under $U(1)\times U(1)$ symmetry is very beneficial. Once again, Kaluza-Klein gauge symmetry comes from translations in the X^{25} direction, so the corresponding charge is the compact momentum p_{25}. For antisymmetric tensor gauge symmetry, examine the zero-momentum vertex operator, which measures the transformation of any state coupled to it. It is proportional to:

$$\partial X^{\mu}\bar{\partial}X^{25} - \partial X^{25}\bar{\partial}X^{\mu} = \bar{\partial}(X^{25}\partial X^{\mu}) - \partial(X^{25}\bar{\partial}X^{\mu})\,. \tag{8.3.4}$$

This is a total derivative, but on winding states, its integral is non-zero since X^{25} is not single-valued. So the $B_{\mu,25}$ charge is the winding number. This is the first "stringy" physics we encounter in this compactification: no state carries this charge in field theory.

We verify the gauge coupling from the string three-point coupling. The gauge boson vertex operator is:

$$\frac{2^{1/2}g_{\mathrm{c},25}}{\alpha'}\,:(\partial X^{\mu}\bar{\partial}X^{25} \pm \partial X^{25}\bar{\partial}X^{\mu})\mathrm{e}^{ik\cdot X}:\,. \tag{8.3.5}$$

We set the 25-dimensional string coupling $g_{\mathrm{c},25} = g_{\mathrm{c}}(2\pi R)^{-1/2}$. The factor $(2\pi R)^{-1/2}$ comes from the normalization of the zero-mode wave function. For general compact momentum and winding number, the tachyon vertex operator is:

$$g_{\mathrm{c},25}\,:\mathrm{e}^{ik_L\cdot X_L(z)+ik_R\cdot X_R(\bar{z})}:\,. \tag{8.3.6}$$

The three-point amplitude for one gauge boson and two tachyons, if the compact momentum and winding number of the tachyons are non-zero, is similar to (6.6.14):

$$\begin{aligned} &-2^{-1/2}\pi i g_{\mathrm{c},25}(2\pi)^{25}\delta^{25}(\textstyle\sum_i k_i)k^{\mu}_{23}(k^{25}_{L23} \pm k^{25}_{R23}) \\ \to&-2^{3/2}\pi i g_{\mathrm{c},25}(2\pi)^{25}\delta^{25}(\textstyle\sum_i k_i)k^{\mu}_{2}(k^{25}_{L2} \pm k^{25}_{R2})\,. \end{aligned} \tag{8.3.7}$$

In the second line, we take the gauge boson momentum $k_1 \to 0$, which defines the gauge coupling. Two gauge bosons couple to $k^{25}_{L2} \pm k^{25}_{R2}$ respectively, i.e., compact momentum and winding number. This is exactly what we expect. Replacing the gauge boson with a graviton and continuing to examine this amplitude, we reproduce the relationships between couplings in (8.1.11) and (8.1.12), which are the same as those found in the effective action.

Enhanced Gauge Symmetry

Thus far, everything is identical to field theory with one compact dimension. The 26-dimensional graviton yields the 25-dimensional graviton plus vector plus modulus. The only string effect is the winding state carrying the $B_{\mu,25}$ charge.

Further and more magical effects come from special compact radii. We omitted the discussion of the massless spectrum previously; states are only massless at special values of the radius R. The most enriched case is $R = \alpha^{1/2}$, so $k^{25}_{L,R} = (n \pm w)\alpha'^{-1/2}$, and the condition for massless states is:

$$(n+w)^2 + 4N = (n-w)^2 + 4\tilde{N} = 4\,. \tag{8.3.8}$$

Besides the general solution $n = w = 0, N = \tilde{N} = 1$, we now have:

$$n = w = \pm 1,\ N = 0,\ \tilde{N} = 1\,, \quad n = -w = \pm 1,\ N = 1,\ \tilde{N} = 0\,, \tag{8.3.9}$$

as well as:

$$n = \pm 2,\ w = N = \tilde{N} = 0\,, \quad w = \pm 2,\ n = N = \tilde{N} = 0\,. \tag{8.3.10}$$

States in (8.3.9) introduce four gauge bosons, with vertex operators:

$$:\bar{\partial} X^\mu e^{ik\cdot X} \exp\left[\pm 2i\alpha'^{-1/2} X^{25}_L\right]:\,, \quad :\partial X^\mu e^{ik\cdot X} \exp\left[\pm 2i\alpha'^{-1/2} X^{25}_R\right]:\ . \tag{8.3.11}$$

The exact definition of the exponential operator is the same as in the previous section. These states possess internal momentum and winding number, so they carry Kaluza-Klein charge and antisymmetric tensor gauge charge. The only consistent theory for charged massless vectors is a non-Abelian gauge theory, so the new gauge bosons must combine with the old gauge bosons to form a non-Abelian theory. It is now convenient to treat the previous vectors using $\partial X^{25} \bar{\partial} X^\mu$ and $\partial X^\mu \bar{\partial} X^{25}$. The first one couples to k^{25}_L; under this coupling, the first pair of states in (8.3.11) carries charge ± 1, while the second pair is neutral. The second couples to k^{25}_R, and so on. This indicates that the gauge group is $SU(2) \times SU(2)$; the three vectors containing $\bar{\partial} X^\mu$ constitute an $SU(2)$. The other three constitute another $SU(2)$. To present this $SU(2) \times SU(2)$, define three $(1, 0)$ currents:

$$j^1(z) =:\cos[2\alpha^{-1/2} X^{25}_L(z)]:\,, \tag{8.3.12a}$$

$$j^2(z) =:\sin[2\alpha'^{-1/2} X^{25}_L(z)]:\,, \tag{8.3.12b}$$

$$j^3(z) = i\partial X^{25}_L(z)/\alpha'^{1/2}\,. \tag{8.3.12c}$$

They can be normalized such that the OPE is:

$$j^i(z) j^j(0) \sim \frac{\delta^{ij}}{2z^2} + i\frac{\epsilon^{ijk}}{z} j^k(0)\,. \tag{8.3.13}$$

For $(0, 1)$ currents $\tilde{j}^i$, there are similar OPEs. The single-pole term implies the corresponding charges form an $SU(2)$ algebra. In fact, since these currents are holomorphic, there is an infinite-dimensional algebra formed by Laurent coefficients:

$$j^i(z) = \sum_{m=-\infty}^{\infty} \frac{j^i_m}{z^{m+1}}\,, \tag{8.3.14a}$$

$$[j^i_m, j^j_n] = \frac{m}{2}\delta_{m,-n}\delta^{ij} + i\epsilon^{ijk} j^k_{m+n}\,, \tag{8.3.14b}$$

which is called a current algebra, affine Lie algebra, or Kac-Moody algebra. We will encounter such algebras frequently at the beginning of Chapter 11. It will be proven that the coefficient of z^{-2} is quantized, and the value in (8.3.13) is the minimum. So this is called a level 1 $SU(2)$ current algebra.

This $SU(2)\times SU(2)$ symmetry shows for the first time that string theory perceives spacetime geometry in a way that is completely different from what we use in field theory, where only $U(1)\times U(1)$ symmetry is obvious at any radius. The origins of j^3 and $j^{1,2}$ are completely different—they are oscillator excitations and winding number-momentum excitations, respectively, and their representations in terms of X^{25} are also completely different, but in their action on the string spectrum, they are related to each other by symmetry. Since the energy of internal momentum and winding number is cancelled by negative zero-point energy, the extra vectors are massless. This situation occurs not only in tachyon theory, but also in other cases in Volume II.

Scale and Coupling

At the $SU(2)\times SU(2)$ radius, the relation (8.1.11) becomes: [1]

$$g_{25}^2 = 2\kappa_{25}^2/\alpha' \,. \tag{8.3.15}$$

If we compactify multiple dimensions, each dimension scales the same way. So, specifically, in 4 dimensions:

$$g_4^2 = 2\kappa_4^2/\alpha' \,. \tag{8.3.16}$$

In nature, non-Abelian gauge couplings are of order 1, which implies that the string length $\alpha^{1/2}$ is not far from the gravitational length κ_4.

We can also think of this in the following way. The 4-dimensional gauge coupling is dimensionless, but the 4-dimensional gravitational coupling has dimensions. Define an effective dimensionless gravitational coupling, which depends on the energy scale E:

$$g_{\mathrm{G},4}^2(E) = \kappa_4^2 E^2 \,, \tag{8.3.17}$$

which grows with the square of energy. Obviously, the dimensionless gauge coupling runs with energy, but this is slower than the dimensional scaling of gravitational coupling. Then by (8.3.16), the string mass scale is where the gauge and gravitational couplings are roughly equal, $g_{\mathrm{G},4}^2(E) \approx g_4^2$.

It is also beneficial to examine the possibility of compactification to 4 dimensions, where some compactification dimensions might be much larger than others. This is best analyzed by starting from low energy and working up, where a dimensionless gauge coupling g_4^2 exists at low energy. At energy ρ_5^{-1}, one compact dimension becomes a visible dimension, and physics becomes 5-dimensional (for simplicity, we only examine this one dimension at this scale). The effective 5-dimensional coupling is:

$$g_5^2 = 2\pi\rho_5 g_4^2 \,. \tag{8.3.18}$$

Similar to (8.1.12), this originates from the effective action. The coupling g_5^2 has dimensions of length. That is, 5-dimensional Yang-Mills theory, like 4-dimensional gravity, is non-renormalizable. Correspondingly, the effective dimensionless coupling is:

$$\hat{g}_5^2 = g_5^2 E = 2\pi\rho_5 E g_4^2 \,, \tag{8.3.19}$$

which grows linearly with energy. However, g_4^2 is not much less than 1 in nature, so the coupling becomes strong quickly at high energy; it is presumed that string theory is then required to make the short-distance behavior reasonable. So this compactification scale must be close to the gravitational scale. This analysis is modified in open string theory, where modern understandings of strong-coupling string theory are taken into account. There is a very low probability that string or Kaluza-Klein excitations are far below the Planck scale.

[1] More accurately, this is a coupling for an $(n,w)=(1,0)$ state. Non-Abelian gauge bosons have $(|n|,|w|) = (1,1)$, so their coupling is $g_{SU(2),25}^2 = 4\kappa_{25}^2/\alpha'$, which is the traditional quantum physics definition of $SU(2)$ coupling. We will see in Chapter 18 that this holds for all level 1 current algebras.

Higgs Mechanism

Let R move away from the $SU(2) \times SU(2)$ radius. It is most instructive to examine what happens. The extra gauge bosons now acquire mass:

$$m = \frac{|R^2 - \alpha'|}{R\alpha'} \approx \frac{2}{\alpha'}\left|R - \alpha'^{1/2}\right| , \tag{8.3.20}$$

where the approximation holds when the radius R is very close to the $SU(2) \times SU(2)$ radius $\alpha'^{1/2}$. Near this radius, the mass is much smaller than the string scale, and we should use low-energy field theory to understand the physics.

There is only one way to endow gauge bosons with such mass: spontaneous symmetry breaking, which is exactly what occurs. When $R = \alpha^{1/2}$, there are 10 massless scalars: the dilaton, the modulus $G_{25,25}$, the four states in (8.3.9), and the four states in (8.3.10), the latter nine generated by the vertex operators:

$$:j^i(z)\tilde{\jmath}^j(\bar{z})\mathrm{e}^{ik\cdot X(z,\bar{z})}: \ . \tag{8.3.21}$$

Index i is a vector under the left-moving $SU(2)$, and index j is a vector under the right-moving $SU(2)$. That is, they transform according to the $(\mathbf{3}, \mathbf{3})$ representation under $SU(2) \times SU(2)$. In particular, the modulus of the radius is $j^3\tilde{\jmath}^3$. Moving away from the $SU(2) \times SU(2)$ radius gives this field an expectation value, so the gauge symmetry is broken down to the $U(1) \times U(1)$ symmetry that preserves the z-axis. Indeed, this is exactly the gauge group at a general radius. Near the $SU(2)$ radius, the mass (8.3.20) is linear with respect to the modulus dimension $|R - \alpha'^{1/2}|$, just like standard spontaneous breaking.

We denote the spacetime fields corresponding to these nine scalar fields as M_{ij}. One variation in the modulus is $M_{33} \neq 0$, but near the $SU(2) \times SU(2)$ point, it is very instructive to investigate more general backgrounds. The interactions of these massless fields will be described by some spacetime action, and this action will contain a potential $U(M)$. These fields have no mass at the symmetry point, so there is no M^2 term in U, but there may be $(SU(2) \times SU(2))$-invariant cubic terms:

$$U(M) \propto \epsilon^{ijk}\epsilon^{i'j'k'} M_{ii'} M_{jj'} M_{kk'} = \det M \, . \tag{8.3.22}$$

It is not difficult to prove from the three-point string amplitude that this term indeed exists. It is a cubic term for bosons and is unbounded from below, but like the tachyon, this is an artifact of bosonic string theory.

Ignoring the fact that this potential is unstable for small variations, it is meaningful to find its static classical solutions. To have a static background solution, the field equations for the graviton and dilaton require the potential to be zero, and the field equations for M_{ij} require the potential to be stable:

$$U(M) = \frac{\partial U(M)}{\partial M_{ij}} = 0 \, . \tag{8.3.23}$$

Diagonalizing M using $SU(2) \times SU(2)$ rotations, the condition (8.3.23) becomes:

$$M_{11}M_{22}M_{33} = M_{11}M_{22} = M_{11}M_{33} = M_{22}M_{33} = 0 \, . \tag{8.3.24}$$

Therefore, only one diagonal component can be non-zero. By $SU(2) \times SU(2)$ rotations, we can take this non-zero diagonal component to be M_{33}. So we have not constructed any physically new string background.

A continuous family of static background solutions is called a flat direction. There are nine massless scalars here, which can be reduced to three diagonal fields by $SU(2) \times SU(2)$ rotations, but there is only one flat direction (ignoring gauge equivalent directions). Of course, we only calculated the cubic terms in the fields, but in general, the cubic terms can be zero, while higher-order terms remain non-zero. To construct an exact flat direction, we need a more specialized discussion. That is, for any value of R, we can construct a free CFT. Note that if we consider $j^1\tilde{\jmath}^1$ or $j^2\tilde{\jmath}^2$ as moduli, since they are sines and cosines of X^{25}, the worldsheet action would look very complex, but since it is equivalent to the free theory obtained by varying R, the worldsheet theory is solvable.

T-duality

From the mass formula:

$$m^2 = \frac{n^2}{R^2} + \frac{w^2R^2}{\alpha'^2} + \frac{2}{\alpha'}(N + \tilde{N} - 2)\,, \tag{8.3.25}$$

we see that as $R \to \infty$, the mass of the winding states becomes infinite, while the compact momentum approaches a continuous spectrum, which is exactly what is expected for non-compact dimensions. Examine the opposite limit $R \to 0$. States with compact momentum acquire infinite mass, but the spectrum of winding states now approaches a continuum—winding strings around small circles does not cost too much energy. Thus, as the radius approaches zero, the spectrum approaches a continuum once again, just as in non-compact dimensions. This is completely different from field theory behavior, where there is only compact momentum n and no winding number w, and at $R \to 0$, no state becomes light.

In fact, the limits $R \to 0$ and $R \to \infty$ are physically equivalent. The spectrum (8.3.25) is invariant under the following transformation:

$$R \to R' = \frac{\alpha'}{R}, \quad n \leftrightarrow w\,. \tag{8.3.26}$$

This equivalence simultaneously extends to interactions. Note that swapping n and w is equivalent to:

$$p_L^{25} \to p_L^{25}\,, \quad p_R^{25} \to -p_R^{25}\,. \tag{8.3.27}$$

Examine this theory at radius R. Recall the decomposition $X^{25}(z,\bar{z}) = X_L^{25}(z) + X_R^{25}(\bar{z})$, and define:

$$X'^{25}(z,\bar{z}) = X_L^{25}(z) - X_R^{25}(\bar{z})\,. \tag{8.3.28}$$

The OPE and energy-momentum tensor of field X'^{25} are the same as those of X^{25}; negative signs will always appear in pairs. The only change to the CFT caused by replacing X^{25} with X'^{25} is that it causes the sign change (8.3.27), which will transform the spectrum of the theory at radius R into the spectrum of the theory at radius R'. That is, they are the same theory, just one written as X^{25} and the other as X'^{25}. This equivalence is called T-duality. That is, the $R \to 0$ limit and $R \to \infty$ are physically completely equivalent, which is completely different from the behavior of point particles, and this is another proof that the geometry of strings at short distances is completely different. The space of inequivalent theories is the ray $R \geq \alpha'^{1/2}$. We could also take the range to be $0 \leq R \leq \alpha^{1/2}$, but taking the larger one is more natural: to us, the momentum continuum is more familiar than the winding continuum. And in the picture of larger R, the issue of locality will be clearer. Therefore, the smallest radius is the self-dual radius:

$$R_{\text{self-dual}} = R_{SU(2)\times SU(2)} = \alpha'^{1/2}\,. \tag{8.3.29}$$

The smallest distance scale is of the order of the string length. This phenomenon will repeatedly appear in string perturbation theory, but non-perturbatively, we will see structures at even smaller scales.

Many applications of T-duality are reflected in its non-trivial effect on the string dilaton Φ. Examine the scattering amplitude of gravitons with neither winding nor compact momentum. These states are invariant under T-duality, so the amplitude must be as well. The latter can be read from the low-energy action (8.1.9), so the 25-dimensional coupling κ_{25} must be invariant under duality. And for the 26-dimensional coupling $\kappa = (2\pi\rho)^{1/2}\kappa_{25}$, the effect of duality is:

$$\rho' = \frac{\alpha'}{\rho}\,, \qquad \kappa' = \frac{\alpha'^{1/2}}{\rho}\kappa\,. \tag{8.3.30}$$

Since $\kappa \propto \mathrm{e}^{\Phi}$, this implies:

$$\mathrm{e}^{\Phi'} = \frac{\alpha^{1/2}}{\rho}\mathrm{e}^{\Phi}\,. \tag{8.3.31}$$

T-duality is a symmetry that relates different states (backgrounds) in a single theory; in fact, it is a gauge symmetry. We saw that the modulus δR is the 33-component of the $(\mathbf{3},\mathbf{3})$ field. Rotating around the 1-axis of one of the $SU(2)$s by an angle π reflects the sign of the modulus, so decreasing R from the $SU(2)\times SU(2)$ radius is gauge equivalent to increasing R. Therefore, the $\mathbb{Z}_2$ symmetry of duality is a small part of the $SU(2)\times SU(2)$ gauge symmetry. This further implies that duality is not only a symmetry of string perturbation theory, but also a symmetry of the exact theory. If we have massless gauge bosons in the leading approximation, even a small explicit breaking of symmetry would lead to inconsistency (spontaneous symmetry breaking is not a problem; at this point T-duality is already spontaneously broken, moving away from the self-dual radius).

The final discussion exemplifies an important idea: the mutual mapping of discussions on strings and spacetime. The appearance of enhanced symmetry is a pure string theory phenomenon. We do not have a non-perturbative understanding of string theory. But we know a lot about low-energy field theory, even non-perturbatively, and we can utilize this.

8.4 Compactification of Several Dimensions

Now extend the analysis to k periodic dimensions:

$$X^m \cong X^m + 2\pi R\,, \quad 26-k \le m \le 25\,. \tag{8.4.1}$$

Let $d = 26-k$ be the number of non-compact dimensions. Now spacetime is $M^d \times T^k$. Assume for now that the coordinate periodicity (8.4.1) remains unchanged, but the actual geometry of the k-torus actually depends on the internal metric G_{mn}. When the compact dimension is greater than 1, the antisymmetric tensor also has scalar components B_{mn}. This altogether gives k^2 scalars. There are also Kaluza-Klein gauge bosons A^m_μ and antisymmetric tensor gauge bosons $B_{m\mu}$. The low-energy effective action is:

$$\begin{aligned}\boldsymbol{S} = \frac{(2\pi R)^k}{2\kappa_0^2}\int \mathrm{d}^d x\,(-G_d)^{1/2}\mathrm{e}^{-2\Phi_d}\Big[&\boldsymbol{R}_d + 4\partial_\mu\Phi_d\partial^\mu\Phi_d \\ &- \frac{1}{4}G^{mn}G^{pq}(\partial_\mu G_{mp}\partial^\mu G_{nq} + \partial_\mu B_{mp}\partial^\mu B_{nq}) \\ &- \frac{1}{4}G_{mn}F^m_{\mu\nu}F^{n\mu\nu} - \frac{1}{4}G^{mn}H_{m\mu\nu}H^{\mu\nu}_n - \frac{1}{12}H_{\mu\nu\lambda}H^{\mu\nu\lambda}\Big]\,,\end{aligned} \tag{8.4.2}$$

where $\Phi_d = \Phi - \frac{1}{4}\ln\det G_{mn}$.

String Spectrum

The main new issue is the antisymmetric tensor background B_{mn}. The contribution of this term to the worldsheet Lagrangian density is proportional to:

$$B_{mn}\partial_a(g^{1/2}\epsilon^{ab}X^m\partial_b X^n)\,, \tag{8.4.3}$$

which is a total derivative if B_{mn} is constant. It has no local effect, and the worldsheet theory remains a CFT. But it changes the string spectrum. We will examine this using canonical methods and path integrals.

In canonical methods, focus on the contribution of zero modes to the worldsheet action:

$$X^m(\sigma^1, \sigma^2) = x^m(\sigma^2) + w^m R\sigma^1 \,, \tag{8.4.4}$$

substituting it into the worldsheet action:

$$L = \frac{1}{2\alpha'} G_{mn}(\dot{x}^m \dot{x}^n + w^m w^n R^2) - \frac{i}{\alpha'} B_{mn} \dot{x}^m w^n R \,. \tag{8.4.5}$$

The dot represents differentiation with respect to worldsheet time σ^2. The canonical momentum is:

$$p_m = -\frac{\partial L}{\partial v^m} = \frac{1}{\alpha'}(G_{mn} v^n + B_{mn} w^n R) \,. \tag{8.4.6}$$

where $v^m = i\dot{x}^m$. Some less familiar notation appears because we used Euclidean time. Extended to Minkowski time, v^m becomes the velocity $\partial_0 x^m$. The periodicity of the wave function implies the quantization of canonical momentum, $p_m = n_m/R$, thus:

$$v_m = \alpha' \frac{n_m}{R} - B_{mn} w^n R \,. \tag{8.4.7}$$

The zero-mode contribution to the worldsheet Hamiltonian is:

$$\frac{1}{2\alpha'} G_{mn}(v^m v^n + w^m w^n R^2) \,, \tag{8.4.8}$$

and the mass of the closed string is:

$$m^2 = \frac{1}{2\alpha'^2} G_{mn}(v_L^m v_L^n + v_R^m v_R^n) + \frac{2}{\alpha'}(N + \tilde{N} - 2) \,, \tag{8.4.9a}$$

$$v_{L,R}^m = v^m \pm w^m R \,. \tag{8.4.9b}$$

Therefore, since v^m depends on B_{mn}, the B_{mn} background shifts the mass of winding states. The $L_0 - \tilde{L}_0$ constraint is:

$$\begin{aligned} 0 &= G_{mn}(v_L^m v_L^n - v_R^m v_R^n) + 4\alpha'(N - \tilde{N}) \\ &= 4\alpha'(n_m w^m + N - \tilde{N}) \,. \end{aligned} \tag{8.4.10}$$

On the other hand, consider the torus path integral for the partition function. The antisymmetric tensor term is locally a total derivative, so it only depends on the topology of the configuration. Consider the configuration:

$$X^m = (w_1^m \sigma^1 + w_2^m \sigma^2)R \,. \tag{8.4.11}$$

The spatial direction on the torus winds w_1^m times around each periodic spacetime dimension, while the temporal direction winds w_2^m times. The B_{mn} worldsheet action is:

$$2\pi i b_{mn} w_1^m w_2^n \,, \tag{8.4.12}$$

where $b_{mn} = B_{mn} R^2/\alpha'$. It can be proven through the Poisson resummation formula that summing over path integral sectors with phase (8.4.12) is related to the partition function with the shifted spectrum (8.4.9). In the previous description, the metric G_{mn} appears in the worldsheet field action. For vertex operator calculations, it is more convenient to retain the usual action. Write the metric in terms of the frame ${e_m}^r$:

$$G_{mn} = {e_m}^r {e_n}^r \,, \tag{8.4.13}$$

where $r, s, \dots$ are tangent space coordinates. The OPE of the coordinates $X^r = e_m{}^r X^m$ is standard. The vertex operator momentum is:

$$k_{rL} = e_r{}^m \frac{v_{mL}}{\alpha'}, \qquad k_{rR} = e_r{}^m \frac{v_{mR}}{\alpha'}, \tag{8.4.14}$$

where $e_r{}^m$ is the inverse frame. The mass-shell conditions become:

$$m^2 = \frac{1}{2}(k_{rL}k_{rL} + k_{rR}k_{rR}) + \frac{2}{\alpha'}(N + \tilde{N} - 2), \tag{8.4.15a}$$

$$0 = \alpha'(k_{rL}k_{rL} - k_{rR}k_{rR}) + 4(N - \tilde{N}). \tag{8.4.15b}$$

In the following, we will use coordinates X^r in any discussion of vertex operators.

Narain Compactification

There is a very elegant description for general toroidal compactifications. Consider the winding state vertex operator $e^{ik_L \cdot X_L + ik_R \cdot X_R}$. For any given compactification, the momentum spectrum (k_{rL}, k_{rR}) forms a lattice in the $2k$-dimensional momentum space $\mathbb{R}^{2k}$. That is, the momentum spectrum consists of all linear combinations of $2k$ mutually independent basis vectors with integer coefficients. We now work with dimensionless momentum $l_{L,R} = k_{L,R}(\alpha'/2)^{1/2}$, and we denote the corresponding lattice as Γ. The OPE of two vertex operators is:

$$:e^{ik_L \cdot X_L(z) + ik_R \cdot X_R(\bar{z})}::e^{ik'_L \cdot X_L(0) + ik'_R \cdot X_R(0)}:$$
$$\sim z^{l_L \cdot l'_L} \bar{z}^{l_R \cdot l'_R} :e^{i(k_L + k'_L) \cdot X_L(0) + i(k_R + k'_R) \cdot X_R(0)}: . \tag{8.4.16}$$

When one vertex operator winds around another vertex operator once, this product yields the phase $\exp[2\pi i(l_L \cdot l'_L - l_R \cdot l'_R)]$. The single-valuedness of the operator product requires:

$$l_L \cdot l'_L - l_R \cdot l'_R \equiv l \circ l' \in \mathbb{Z}, \tag{8.4.17}$$

where l and l' belong to Γ. We defined the inner product $\circ$, which has signature (k, k) in $\mathbb{R}^{2k}$. The dual lattice Γ^* is defined as the set of points in $\mathbb{R}^{2k}$ with integer-valued $\circ$ inner products with points in Γ. Then the single-valuedness condition is equivalent to:

$$\Gamma \subset \Gamma^*. \tag{8.4.18}$$

Modular invariance also constrains Γ. It was proven in section 7.2 that invariance under $\tau \to \tau + 1$ requires $L_0 - \tilde{L}_0$ to be an integer. Then:

$$l \circ l \in 2\mathbb{Z} \quad \forall l \in \Gamma. \tag{8.4.19}$$

Remark. $L_0 - \tilde{L}_0 = \frac{\alpha'}{4}(k_{rL}k_{rL} - k_{rR}k_{r_R}) = \frac{1}{2}(l_{rL}l_{rR} - l_{rR}l_{rR}) = \frac{1}{2} l \circ l$.

(8.4.19) actually implies (8.4.17), which is given by the closure of the OPE:

$$2l \circ l' = (l + l') \circ (l + l') - l \circ l - l' \circ l' \in 2\mathbb{Z}. \tag{8.4.20}$$

Modular invariance under $\tau \to -1/\tau$ requires more careful calculation. The partition function of the compact dimension is:

$$Z_\Gamma(\tau) = |\eta(\tau)|^{-2k} \sum_{l \in \Gamma} \exp\left(\pi i \tau l_L^2 - \pi i \tau l_R^2\right). \tag{8.4.21}$$

To determine the transformation properties of Z_Γ, use Poisson resummation. Use:

$$\sum_{l' \in \Gamma} \delta(l - l') = V_\Gamma^{-1} \sum_{l'' \in \Gamma^*} \exp\left(2\pi i l'' \circ l\right). \tag{8.4.22}$$

Here, V_Γ is the volume of the unit cell of lattice Γ. (8.4.22) is similar to a Fourier series: only when $l \in \Gamma$ does the phase average not vanish, and normalization is given by integration over the unit cell. Using (8.4.22), the sum (8.4.21) can be written as:

$$\begin{aligned} Z_\Gamma(\tau) &= V_\Gamma^{-1}|\eta(\tau)|^{-2k} \sum_{l'' \in \Gamma^*} \int \mathrm{d}^{2k}l \, \exp\bigl(2\pi i l'' \circ l + \pi i \tau l_L^2 - \pi i \bar{\tau} l_R^2\bigr) \\ &= V_\Gamma^{-1}(\tau\bar{\tau})^{-k/2}|\eta(\tau)|^{-2k} \sum_{l'' \in \Gamma^*} \exp\bigl(-\pi i l_L''^2/\tau + \pi i l_R''^2/\bar{\tau}\bigr) \\ &= V_\Gamma^{-1} Z_{\Gamma^*}(-1/\tau) , \end{aligned} \tag{8.4.23}$$

where we used modular transformation (7.2.44) in the last line. Then the sufficient condition for modular invariance is:

$$\Gamma = \Gamma^* , \tag{8.4.24}$$

Since $V_\Gamma = V_{\Gamma^*}^{-1}$, if (8.4.24) holds, it equals 1. If modular invariance holds for all τ, it can further be proven that this is a sufficient condition.

Consistency conditions can be summarized in two requirements: Γ is an even self-dual lattice with signature (k, k); even refers to property (8.4.19), and self-dual refers to (8.4.24).

All such lattices have been classified. Note that consistency conditions (8.4.19) and (8.4.24) depend on momentum l through the $\circ$ product, which is invariant under Lorentz boosts in $2k$-dimensional space, $O(k, k, \mathbb{R})$ transformations. If Γ is an even self-dual lattice, then:

$$\Gamma' = \Lambda \tag{8.4.25}$$

is also an even self-dual lattice, where Λ is an $O(k, k, \mathbb{R})$ transformation. Note that $O(k, k, \mathbb{R})$ is not a symmetry of the theory. The mass-shell condition and operator product also contain individual dot products $l_L \cdot l'_L$ and $l_R \cdot l'_R$, thus they are only invariant under $O(k, \mathbb{R}) \times O(k, \mathbb{R})$. So, most $O(k, k, \mathbb{R})$ transformations yield inequivalent theories. Examine the $k = 1$ case, where:

$$l_{L,R} = \frac{n}{r} \pm \frac{mr}{2} , \tag{8.4.26}$$

$r = R(2/\alpha')^{1/2}$ is the dimensionless radius. This is indeed an even self-dual lattice. The transformation:

$$l'_L = l_L \cosh\lambda + l_R \sinh\lambda , \quad l'_R = l_L \sinh\lambda + l_R \cosh\lambda \tag{8.4.27}$$

transforms the lattice into one with $r' = r\mathrm{e}^{-\lambda}$.

Given the Lorentz signature, all even self-dual lattices can be obtained from one lattice through $O(k, k, \mathbb{R})$ transformations. Starting from a given solution, such as all compact dimensions being mutually orthogonal and at the $SU(2)$ radius with $B_{mn} = 0$. The corresponding momentum lattice Γ_0 is called kP_2 and yields the gauge group $SU(2)^{2k}$. As discussed earlier, if $\Lambda' \in O(k, \mathbb{R}) \times O(k, \mathbb{R})$, then two $O(k, k, \mathbb{R})$ transformations Λ and $\Lambda'\Lambda$ yield equivalent string theories; so, the space of inequivalent theories obtained in this way is:

$$\frac{O(k, k, \mathbb{R})}{O(k, \mathbb{R}) \times O(k, \mathbb{R})} . \tag{8.4.28}$$

There are also some discrete equivalence relations, which we will discuss briefly later. This is equivalent to the description using G_{mn} and B_{mn} made earlier. In particular, the number of parameters in the coset (8.4.28) is:

$$\frac{2k(2k-1)}{2} - k(k-1) = k^2 , \tag{8.4.29}$$

which is the same as the number of components in G_{mn} and B_{mn}. Compared with one-dimensional compactification, T-duality symmetry is greatly expanded. In this abstract description, $O(k, k, \mathbb{R})$ has some discrete subgroups that transform the lattice Γ_0 into itself. These subgroups are customarily denoted as $O(k,k,\mathbb{Z})$. If Λ'' belongs to this subgroup, then obviously $\Lambda\Gamma_0$ and $\Lambda\Lambda''\Gamma_0$ are the same lattice. In summary, we have the following equivalence relations:

$$\Lambda\Gamma_0 \cong \Lambda'\Lambda\Lambda''\Gamma_0\,, \tag{8.4.30a}$$
$$\Lambda' \in O(k,\mathbb{R})\times O(k,\mathbb{R}), \quad \Lambda'' \in O(k,k,\mathbb{Z})\,. \tag{8.4.30b}$$

Then the space of inequivalent lattices and inequivalent backgrounds is:

$$\frac{O(k,k,\mathbb{R})}{O(k,\mathbb{R})\times O(k,\mathbb{R})\times O(k,k,\mathbb{Z})}\,. \tag{8.4.31}$$

Note that the continuous group in the denominator is a left action, while the discrete group is a right action.

From the form of the background, the T-duality group $O(k,k,\mathbb{Z})$ contains several transformations. One is the $R\to\alpha'/R$ duality on a single axis. Another is the large coordinate transformation reflecting periodicity:

$$x''^m = L^m{}_n x^n\,, \tag{8.4.32}$$

where $L^m{}_n$ are integers and $\det L = 1$. This is the group $SL(k,\mathbb{Z})$. The last is the shift in the antisymmetric tensor background:

$$b_{mn}\to b_{mn}+N_{mn} \tag{8.4.33}$$

where N_{mn} are integers. From canonical results (8.4.7) or path integral phases (8.4.12), it can be seen that these leave the spectrum invariant. These together generate the entire T-duality group. For a general Λ, there are no $\Lambda_{1,2}$ in the denominator group such that $\Lambda_1\Lambda\Lambda_2=\Lambda$, so there is no T-duality that leaves points in the covering space fixed. At some special Λ, solutions to $\Lambda_1\Lambda\Lambda_2=\Lambda$ exist, and they are different points of the T-duality element Λ_2. In the one-dimensional case, we can find such points, which are also the points of enhanced gauge symmetry. In higher-dimensional cases, besides $SU(2)$, larger gauge groups appear at different points, which we will discuss further in Chapter 11.

When there is a moduli space parameterized by scalar fields ϕ^i, the kinetic term:

$$-\frac{1}{2}g_{ij}(\phi)\partial_\mu\phi^i\partial^\mu\phi^j \tag{8.4.34}$$

defines a natural metric on the moduli space. More precisely, a definite definition requires performing a Weyl transformation first, as in (**??**), to make the coefficient of the gravitational action modulus-independent and eliminate the mixing between the modulus and the spacetime metric. In low-energy action (8.4.2), the corresponding scalar kinetic term is proportional to:

$$\frac{16}{d-2}\partial_\mu\Phi\partial^\mu\Phi + G^{mn}G^{pq}(\partial_\mu G_{mp}\partial^\mu G_{nq}+\partial_\mu B_{mp}\partial^\mu B_{nq})\,. \tag{8.4.35}$$

The coset space (8.4.31) has a unique $O(k,k,\mathbb{R})$-invariant metric, which is exactly the second term in (8.4.35). This $O(k,k,\mathbb{R})$ is not a symmetry of the entire theory; only its discrete T-duality subgroup $O(k,k,\mathbb{Z})$ is. The difference between the two is invisible at low energy—it comes from the quantization of string zero modes, thus only affecting the massive spectrum. Therefore, $O(k,k,\mathbb{R})$ is only an accidental symmetry of the low-energy theory. In this example, the moduli space is the product of the dilaton moduli space and the compactification moduli space.

Example

Two compactification dimensions are a good example. The four moduli $G_{24,24}$, $G_{24,25}$, $G_{25,25}$, and $B_{24,25}$ are usually combined into two complex fields $\tau = \tau_1 + i\tau_2$ and $\rho = \rho_1 + i\rho_2$. Define:

$$\begin{aligned}\rho &= \frac{R^2}{\alpha'}\left(B_{24,25} + i\det{}^{1/2} G_{mn}\right)\\ &= b_{24,25} + i\frac{V}{4\pi^2\alpha'}\,,\end{aligned} \tag{8.4.36}$$

where V is the volume of the compact 2-torus. Parameterize the compact metric as:

$$\mathrm{d}s^2 = \frac{\alpha'\rho_2}{R^2\tau_2}\left|\mathrm{d}X^{24} + \tau\mathrm{d}X^{25}\right|^2 . \tag{8.4.37}$$

Coordinate transformation (8.4.32) acts on the spacetime 2-torus just as the modular group acts on the worldsheet 2-torus, generating $PSL(2,\mathbb{Z})$ on τ while leaving ρ unchanged. The antisymmetric tensor shift (8.4.33) is $\rho \to \rho + N_{24,25}$ while τ remains unchanged. Simultaneously using T-duality on $X^{24,25}$ causes $\rho \to -1/\rho$ while τ remains unchanged. The latter two transformations generate $PSL(2,\mathbb{Z})$ on ρ. Additionally, T-duality on X^{24} yields $(\tau,\rho) \to (\rho,\tau)$, and spacetime parity $X^{24} \to -X^{24}$ yields $(\tau,\rho) \to (-\bar{\tau},-\bar{\rho})$. The entire T-duality group is:

$$PSL(2,\mathbb{Z}) \times PSL(2,\mathbb{Z}) \rtimes \mathbb{Z}_2^2 . \tag{8.4.38}$$

Further equivalence relations come from worldsheet parity $(\tau,\rho) \to (\tau,-\bar{\rho})$, which flips the sign of the $\circ$ product.

The moduli space is thus the product of two torus moduli spaces, plus some discrete equivalence relations. The kinetic term is proportional to:

$$\frac{\partial_\mu\tau\partial^\mu\bar{\tau}}{\tau_2^2} + \frac{\partial_\mu\rho\partial^\mu\bar{\rho}}{\rho_2^2} . \tag{8.4.39}$$

$\mathrm{d}^2\tau/\tau_2^2$ is encountered in genus zero amplitudes. The continuous group $PSL(2,\mathbb{R})$ transforms τ from the upper half-plane to another point, while the $U(1)$ subgroup leaves each τ invariant. Therefore, the upper half-plane can be regarded as $PSL(2,\mathbb{R})/U(1)$, and the entire moduli space is:

$$\frac{PSL(2,\mathbb{R}) \times PSL(2,\mathbb{R})}{U(1) \times U(1) \times PSL(2,\mathbb{Z}) \times PSL(2,\mathbb{Z}) \times \mathbb{Z}_2^2} . \tag{8.4.40}$$

This is the same as the previous result (8.4.31). In particular, just as $SU(2) \times SU(2)$ is locally the same as $O(4,\mathbf{R})$, $PSL(2,\mathbb{R}) \times PSL(2,\mathbb{R})$ is locally the same as $O(2,2,\mathbb{R})$.

8.5 Orbifolds

Now we examine an equivalence other than the periodic equivalence of X^{25}, namely reflection equivalence:

$$X^{25} \cong -X^{25} \tag{8.5.1}$$

The current fundamental region is the ray $X^{25} \geq 0$, and the hyperplane $X^{25} = 0$ becomes a boundary. This hyperplane is singled out because points on it are fixed under reflection (8.5.1). More generally, we can examine the simultaneous inversion of k coordinates:

$$X^m \to -X^m\,, \qquad 26-k \leq m \leq 25\,. \tag{8.5.2}$$

Similarly, there still exists a space of fixed points, $X^{26-k} = \cdots = X^{25} = 0$. When $k \geq 2$, this space is singular. In particular, when $k = 2$, this is a conical singularity with a deficit angle of π.

By combining reflection with periodic equivalence $X^{25} \cong X^{25} + 2\pi R$, we can construct a compact space; we call the former r and the latter t. We can think of first forming a circle S^1 through the t equivalence, and then identifying points on the circle through r. The corresponding compact space is the segment $0 \leq X^{25} \leq \pi R$. Both ends of the segment are fixed points; point πR is fixed by tr. Altogether, the equivalence relations on the real line form a group:

$$t^m : \quad X^{25} \cong X^{25} + 2\pi R m \,, \tag{8.5.3a}$$

$$t^m r : \quad X^{25} \cong 2\pi R m - X^{25} \,. \tag{8.5.3b}$$

where m is any integer. Similarly, reflection (8.5.2) can be imposed on the k-torus. In this case, there are 2^k different fixed points. Singular spaces obtained in this way are called orbifolds. The space obtained in a non-compact space through equivalence relation (8.5.2) is $\mathbb{R}^k/\mathbb{Z}_2$; if it is a compact space k-torus, what is obtained is $T^k/\mathbb{Z}_2$.

It is not obvious that string theory should be meaningful in such a singular space, but we will see that string theory is indeed meaningful in such spaces. It is not completely inconsistent with toroidal compactification, but it breaks more symmetries (including subsequent supersymmetry), thus it is very valuable in building string models.

Equivalence relation (8.5.1) or (8.5.2) has two effects. First, wave functions must be invariant under inversion and be the same at equivalent points. Second, there are new sectors in the closed string spectrum, where:

$$X^{25}(\sigma^1 + 2\pi) = -X^{25}(\sigma^1) \,, \tag{8.5.4}$$

because they are the same point in equivalent spacetime. Strings in this sector are called twisted states. Similar to periodic equivalence. In this case, the invariance of the wave function discretizes momentum, and closed strings that differ only by periodic equivalence (winding strings) appear. By analogy with Figure 8.1, it can be proven that twisted strings can be generated from untwisted strings. As we will see below, modular invariance requires these strings to appear in the spectrum.

We will focus on the compact one-dimensional orbifold $S^1/\mathbb{Z}_2$. First, verify the untwisted sector; the spectrum of the periodic theory must project onto invariant states, which will reduce it. The effect of r on a general state is:

$$|N, \tilde{N}; k^\mu, n, w\rangle \to (-1)^{\sum_{m=1}^{\infty}(N_m^{25} + \tilde{N}_m^{25})} |N, \tilde{N}; k^\mu, -n, -w\rangle \,, \tag{8.5.5}$$

In particular, the effect of r reverses compact winding and compact momentum. We must construct linear combinations that are invariant under this transformation. For a general R, massless states have $n = w = 0$, so the projection only requires the excitation number for the 25th dimension to be even. Thus, the spacetime graviton, antisymmetric tensor, dilaton, and tachyon survive this projection. The modulus $\alpha_{-1}^{25}\tilde{\alpha}_{-1}^{25}|0; k^\mu, 0, 0\rangle$ also still exists—since R can take any value, this is expected. However, Kaluza-Klein gauge bosons no longer exist. Other massless states at the self-dual point are quite interesting and will be discussed in detail below. In the sector twisted by r, X^{25} is anti-periodic, thus it has a half-integer mode expansion:

$$X^{25}(z, \bar{z}) = i\left(\frac{\alpha'}{2}\right)^{1/2} \sum_{m=-\infty}^{\infty} \frac{1}{m + 1/2}\left(\frac{\alpha_{m+1/2}^{25}}{z^{m+1/2}} + \frac{\tilde{\alpha}_{m+1/2}^{25}}{\bar{z}^{m+1/2}}\right) . \tag{8.5.6}$$

Anti-periodicity makes center-of-mass coordinates and momentum zero, so the string cannot leave the fixed point $X^{25} = 0$: if string vibrations are small, then $X^{25}(z, \bar{z}) \approx 0$. There are also twisted strings around other fixed points:

$$X^{25}(\sigma^1 + 2\pi) = 2\pi R - X^{25}(\sigma^1) \,. \tag{8.5.7}$$

The field is changed by tr after winding around the compact dimension once. The mode expansion differs from above only by an additional constant term πR. All other fixed points $n\pi R$ are images under t transformations.

According to the discussion of zero-point energy in section 2.9, the zero-point energy of X^{25} shifts from $-\frac{1}{24}$ for a periodic boson to $\frac{1}{48}$ for an anti-periodic boson, a net increase of $+\frac{1}{16}$. In twisted states, the mass-shell condition is:

$$m^2 = \frac{4}{\alpha'}\left(N - \frac{15}{16}\right), \quad N = \tilde{N}\,. \tag{8.5.8}$$

Additionally, the contribution of the 25th-dimension oscillators to level N is half-integer. The r-projection once again requires the total number of excitations in the 25th dimension to be even, which can also be derived from the level-matching condition $N = \tilde{N}$. In the twisted sector, the ground states are $|T_{1,2}\rangle$, where subscripts refer to the two fixed points. They are tachyons, and their first excited state $\alpha^{25}_{-1/2}\tilde{\alpha}^{25}_{-1/2}|T_{1,2}\rangle$ is also a tachyon. Massless twisted states do not exist. The extension to the more general case (8.5.2) is straightforward. Although there are singularities in spacetime, the quantization is completely consistent.

A difficulty brought by orbifolds is that twisted state vertex operators are not as simple as winding states; there is no explicit formula like the free field exponential. For two twisted strings and any number of untwisted strings, tree-level amplitudes can be written in the operator form $\langle T|\mathscr{V}_U\mathscr{V}_U\ldots|T\rangle$, where only untwisted vertex operators are used. For four or more twisted state vertex operators (for the path integral to be meaningful, there must be an even number of twisted state vertex operators), there are two methods available. One is the stress tensor method, which we used in section 7.2 to calculate the partition function. The other is to go to the covering space where X^{25} is single-valued (if the sphere has $2g+2$ twisted state vertex operators, the covering space is a surface of genus g), and then calculate the path integral.

Tree-level coupling of untwisted states yields a simple result. On the sphere, if all external states are untwisted, twisting and projection do not enter the calculation at all, and the amplitude is the same as in the untwisted theory. For example, the low-energy effective action of massless untwisted fields is the same as we obtained in toroidal compactification, differing only in the absence of vector fields. This property is usually very useful in analyzing the physics of orbifold theories; it is often called the inheritance principle. Now let us examine the partition function of the X^{25} CFT. In the untwisted sector, we project onto states with $r = +1$:

$$(q\bar{q})^{-1/24}\,\mathrm{Tr}_U\left(\frac{1+r}{2}q^{L_0}\bar{q}^{\tilde{L}_0}\right)\,. \tag{8.5.9}$$

The first term of the projection operator yields exactly half of the toroidal compactification partition function (8.2.9). For the term with r, the diagonal elements of r must have $n = w = 0$. And $(-1)^{N^{25}+\tilde{N}^{25}}$ changes the oscillator sum to $1 - q + q^2 - q^3 + \cdots = (1+q)^{-1}$. Thus, the partition function for untwisted states is:

$$\frac{1}{2}Z_{\text{tor}}(R,\tau) + \frac{1}{2}(q\bar{q})^{-1/24}\prod_{m=1}^{\infty}|1+q^m|^{-2}\,. \tag{8.5.10}$$

The contribution of the twisted sector comes from the sum of half-integer mode oscillators:

$$\begin{aligned}&(q\bar{q})^{1/48}\,\mathrm{Tr}_T\left(\frac{1+r}{2}q^{L_0}\bar{q}^{\tilde{L}_0}\right)\\ &\quad= (q\bar{q})^{1/48}\left[\prod_{m=1}^{\infty}\left|1-q^{m-1/2}\right|^{-2} + \prod_{m=1}^{\infty}\left|1+q^{m-1/2}\right|^{-2}\right]\,,\end{aligned} \tag{8.5.11}$$

the number of twisted sectors gives a factor of 2, which exactly cancels with the $\frac{1}{2}$ in the projection operator. Combined, the partition function can be written as:

$$Z_{\text{orb}}(R,\tau) = \frac{1}{2}Z_{\text{tor}}(R,\tau) + \left|\frac{\eta(\tau)}{\vartheta_{10}(0,\tau)}\right| + \left|\frac{\eta(\tau)}{\vartheta_{01}(0,\tau)}\right| + \left|\frac{\eta(\tau)}{\vartheta_{00}(0,\tau)}\right|\,. \tag{8.5.12}$$

The first term is known to be modular invariant, and the sum of the last three terms is modular invariant according to the modular transformation properties in section 7.2. In the form of path integrals on the torus, $Z_{\text{tor}}(R,\tau)$ is given by fields that are periodic under translation. ϑ_{ab} comes from the path integral of:

$$X^{25}(\sigma^1+2\pi,\sigma^2)=(-1)^{a+1}X^{25}(\sigma^1,\sigma^2)\,, \tag{8.5.13a}$$

$$X^{25}(\sigma^1+2\pi\tau_1,\sigma^2+2\pi\tau_2)=(-1)^{b+1}X^{25}(\sigma^1,\sigma^2) \tag{8.5.13b}$$

For example, the ϑ_{10} term comes from the untwisted sector with r in the trace. That is, anti-periodic in the temporal direction and periodic in the spatial direction. The ϑ_{01} term is the term with 1 in the trace within the twisted sector, which has opposite boundary conditions. They are swapped under $\tau\to-1/\tau$. Thus, the twisted sector must be modular invariant.

Twisting

Orbifold compactification and toroidal compactification are two examples of a construction called twisting, which can be used to construct new string theories from old ones. In a given CFT, if its symmetry contains a discrete group H, we can obtain a new CFT in two steps. First, add twisted sectors, in which closed string worldsheet fields are periodic up to some $h\in H$:

$$\phi(\sigma^1+2\pi)=h\cdot\phi(\sigma^1)\,, \tag{8.5.14}$$

where ϕ refers to general worldsheet fields. Second, restrict the spectrum to H-invariant states. Although (8.5.14) is no longer periodic, the vertex operator is periodic due to the projection. Since the product of H-invariant operators remains H-invariant, it is still closed.

In the partition function, we sum over the twisting h_1 and introduce a projection onto invariant states:

$$P_H=\frac{1}{\text{order}(H)}\sum_{h_2\in H}\hat{h}_2\,. \tag{8.5.15}$$

The operator $\hat{h}_2$ in the trace makes the periodicity of the field in the temporal direction differ by h_2, so in the end, we sum over twisting in the temporal and spatial directions on the torus:

$$Z=\frac{1}{\text{order}(H)}\sum_{h_1,h_2\in H}Z_{h_1,h_2}\,. \tag{8.5.16}$$

The sum over boundary conditions is modular invariant after shifting the summation variables. The modular transformation $\tau\to-1/\tau$ takes (h_1,h_2) to (h_2,h_1^{-1}), while $\tau\to\tau+1$ takes (h_1,h_2) to (h_1,h_1h_2). Therefore, if we project onto H-invariant states by summing over h_2, we also obtain the sum over twisted sectors. In more general cases, particularly in right-left asymmetric theories, phase issues may destroy the natural modular invariance of the sum over boundary conditions.

In the case of toroidal compactification, the original theory is a non-compact theory. H is formed by $t^m=\exp(2\pi iRmp)$, where m is an integer. This exactly has infinite order, but we can regularize it by putting the theory in a box. For orbifolds, we can think of starting from a toroidal theory and then twisting it by r.

If H is non-Abelian, for non-commuting h_1 and h_2, the path integral boundary conditions are inconsistent and the path integral is zero. This can also be derived from the exact expression (8.5.15). If ϕ has spatial twisting h_1, then:

$$\phi'(\sigma^1)=h_2\cdot\phi(\sigma^1) \tag{8.5.17}$$

has a different spatial twisting:

$$\phi'(\sigma^1+2\pi)=h_1'\cdot\phi'(\sigma^1)\,, \tag{8.5.18}$$

where:

$$h_1' = h_2 h_1 h_2^{-1} \tag{8.5.19}$$

So the diagonal elements of $\hat{h}_2$ are zero. We also see that sectors twisted by h_1 and h_1' are not independent—they are related through projection onto h_2-invariant states. Therefore, independent twisted sectors correspond one-to-one with the conjugacy classes of H.

Twisting can be regarded as the gauging of the discrete group H, which is not obvious. Due to discreteness, the gauge parameter must be constant. In particular, there is no corresponding gauge field. The key point here is that on a worldsheet with non-trivial closed paths, only gauge-invariant quantities need to be periodic. So the periodicity of the fields can differ by an H transformation. Gauging a discrete symmetry adds new sectors to the path integral and correspondingly projects onto invariant states. This is just the twisting construction. One can also gauge continuous worldsheet symmetries to produce new CFTs, as we will discuss in Chapter 15.

For general twisting theories, the inheritance principle still holds.

$c = 1$ CFT

We discovered two types of compact $c = 1$ CFTs. They are toroidal compactifications with the radius restricted to $R \geq \alpha'^{1/2}$ and orbifold compactifications with the same range for R. These two types of theories are related to each other. Starting from the $SU(2) \times SU(2)$ radius $R = \alpha'^{1/2}$, further twist by:

$$r' : \quad X^{25} \to X^{25} + \pi\alpha'^{1/2}\,. \tag{8.5.20}$$

This gives a toroidal theory with $R = \alpha'^{1/2}/2$, which is equivalent to $R = 2\alpha'^{1/2}$ via T-duality. In terms of $SU(2)$ currents (8.3.12), r' flips the signs of $j^{1,2}$ and $\tilde{\jmath}^{1,2}$, so it rotates by an angle of π around the 3-axis in each $SU(2)$. Orbifold r flips the signs of $j^{2,3}$ and $\tilde{\jmath}^{2,3}$, so it is a rotation by angle π around the 1-axis in $SU(2)$. However, clearly these rotations are equivalent under $SU(2) \times SU(2)$, so the corresponding theories are equivalent: the torus at $R = 2\alpha'^{1/2}$ and the orbifold at $R = \alpha'^{1/2}$. The corresponding partition functions are equal:

$$Z_{\text{orb}}(\alpha'^{1/2}, \tau) = Z_{\text{tor}}(2\alpha'^{1/2}, \tau) \tag{8.5.21}$$

This is also proven by θ-function identities.

This equivalence holds only at these radii. For example, for a general R, toroidal theory only has $U(1) \times U(1)$ gauge symmetry, while the orbifold has no gauge symmetry. Therefore, the relationship between the two moduli spaces is as shown in Figure 8.5.1. Moduli spaces with different branches meet at special points; in supersymmetric theories, this structure has very general and important features. Generally speaking, there will be different low-energy effective theories in each branch, manifested here as low-energy gauge symmetry. More light fields exist near special points. As we move away from special points along different branches, different subsets of them become massive.

The low-energy physics near the intersection of the toroidal and orbifold branches is very instructive. Describe it with the $SU(2)$ theory twisted by r' in (8.5.20). The massless scalars remaining under this twisting are $j^3\tilde{\jmath}^3$ and $j^i\tilde{\jmath}^j$ for $i, j \in \{1, 2\}$. Since they are untwisted, the potential is the same as before the twist, i.e., $\det M$. Scalar field M_{ij} can still be diagonalized through $U(1) \times U(1)$ rotations, and the potential is flat only when one diagonal element is non-zero. However, the modulus M_{33} cannot be rotated to M_{11} or M_{22} now, so these directions are physically different. Modulus M_{33} preserves $U(1) \times U(1)$ and corresponds to movement along the toroidal branch. Modulus M_{11} or M_{22} breaks the gauge symmetry and corresponds to movement along the orbifold branch. Since this modulus is charged, its sign reversal is a gauge transformation, so physically different orbifolds end at the intersection of branches. We cannot transform two moduli simultaneously because their linear combination is not a flat direction.

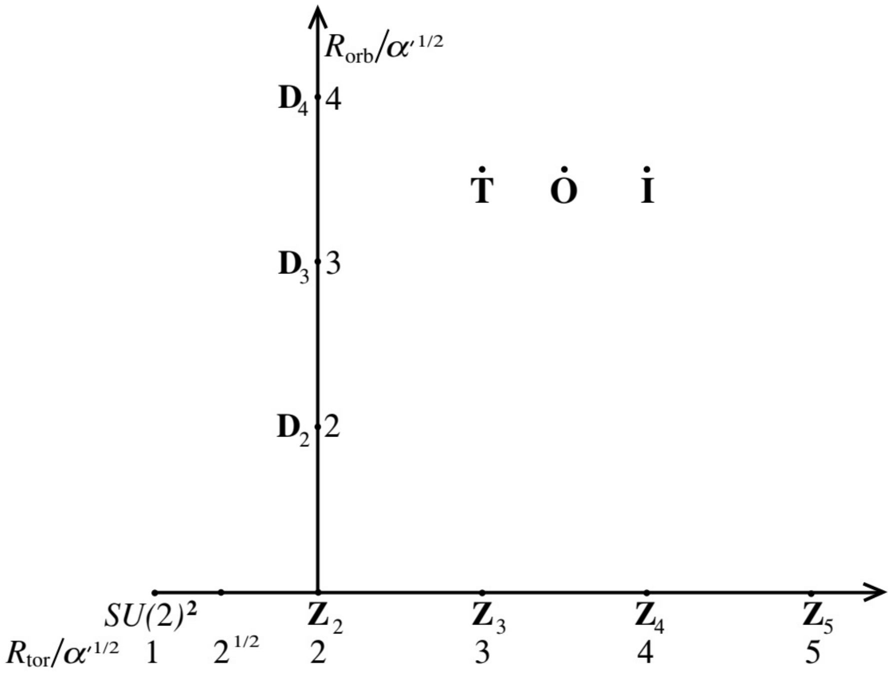


Figure 8.5.1: Moduli space of $c = 1$ CFT. The horizontal axis is the toroidal branch, while the vertical axis is the orbifold branch. The points marked in the figure are obtained by twisting with a discrete subgroup of the $SU(2)$ group, including three isolated points obtained from the tetrahedral group, the octahedral group, and the icosahedral group. Other special radii, such as the torus at $R = (2\alpha')^{1/2}$, will appear later.

String theory on an orbifold is equivalent to string theory on a circle, which is quite remarkable. This shows that strings perceive spacetime geometry in a way that is not completely identical to ours.

Moving along the toroidal branch, there will be extra massless states at integer multiples of the $SU(2)$ radius $R = k\alpha'^{1/2}$. These states have $(n, w) = (\pm 2k, 0)$. We can think of these points as being obtained this way: under T-duality, $R = \alpha'^{1/2}/k$, they are again obtained by twisting $SU(2) \times SU(2)$ theory using $\mathbb{Z}_k$ symmetry:

$$X^{25} \to X^{25} + \frac{2\pi\alpha'^{1/2}}{k} . \tag{8.5.22}$$

This multiplies $j^1 + ij^2$ and $\tilde{\jmath}^1 + i\tilde{\jmath}^2$ by $\exp(2\pi i/k)$, thus the surviving massless scalars are:

$$j^3\tilde{\jmath}^3, \quad j^1\tilde{\jmath}^1 + j^2\tilde{\jmath}^2, \quad j^1\tilde{\jmath}^2 - j^2\tilde{\jmath}^1 . \tag{8.5.23}$$

The first of these changes the radius of the torus, thus it is always a flat direction. However, because the flatness condition (8.3.23) for the potential is not met in direction $M_{11} = M_{22}$ or $M_{12} = -M_{21}$, no other branches emerge from these points.

Offset (8.5.22) generates the $\mathbb{Z}_k$ subgroup of $SU(2) \times SU(2)$. Let us examine twisting with other subgroups. Dihedral group $\mathbf{D}_k$ consists of offset (8.5.22) and reflection $X^{25} \to -X^{25}$, which gives the orbifold at radius $k\alpha'^{1/2}$. The other three discrete subgroups are the tetrahedral group, octahedral group, and icosahedral group. They remove all moduli from the CFT, so twisting CFT is an isolated point, not a member of a continuous family. Unlike $\mathbb{Z}_k$ and $\mathbf{D}_k$ twisting, which can be defined at all radii, these discrete subgroups containing $SU(2)$ elements exist only at discrete radii, thus the radius cannot vary after twisting. String theory with such a compactified space will only have one scalar—the dilaton.

These are all known $c = 1$ CFTs, and thus they give all known 25-dimensional bosonic string backgrounds.

8.6 Open Strings

In toroidal compactification of open strings, there is a new feature: the possible existence of non-trivial Wilson lines, flat backgrounds for gauge fields. First study some $U(1)$ gauge theories with charged fields, examining constant backgrounds:

$$A_{25}(x^M) = -\frac{\theta}{2\pi R} = -i\Lambda^{-1}\frac{\partial\Lambda}{\partial x^{25}}, \quad \Lambda(x^{25}) = \exp\left(-\frac{i\theta x^{25}}{2\pi R}\right), \tag{8.6.1}$$

where θ is a constant. Locally, this is pure gauge, the field strength is zero, and field equations are trivial. However, gauge parameter Λ does not satisfy spacetime periodicity, so the background has physical effects. The gauge invariant measuring this effect is the Wilson line:

$$W_q = \exp\left(iq\oint \mathrm{d}x^{25}\, A_{25}\right) = \exp(-iq\theta)\,. \tag{8.6.2}$$

First, consider a point particle with charge q; its gauge-fixed action is:

$$S = \int \mathrm{d}\tau \left(\frac{1}{2}\dot{X}^M\dot{X}_M + \frac{m^2}{2} - iqA_M\dot{X}^M\right). \tag{8.6.3}$$

The gauge action is $-iq\int \mathrm{d}x^M\, A_M$, so a path once around the compact dimension will pick out a phase equal to the Wilson line W_q. Canonical momentum is:

$$p_{25} = -\frac{\partial L}{\partial v^{25}} = v^{25} - \frac{q\theta}{2\pi R}\,, \tag{8.6.4}$$

where $v^{25} = i\dot{X}^{25}$ as in (8.4.6). The wave function must be periodic in the compact dimension, so $p_{25} = l/R$ for integer l, and:

$$v_{25} = \frac{2\pi l + q\theta}{2\pi R}\,. \tag{8.6.5}$$

The Hamiltonian that annihilates physical states is:

$$H = \frac{1}{2}(p_\mu p^\mu + v_{25}^2 + m^2)\,, \tag{8.6.6}$$

since v_{25} depends on θ, $-p_\mu p^\mu$ is shifted. The same spectrum can be obtained in the field description. Note that v_{25} is exactly the gauge-invariant momentum $-i\partial_{25} - qA_{25}$. We could also perform the gauge transformation Λ^{-1} to set A_{25} to zero; under this gauge, the charged field is no longer periodic, picking out the phase $\exp(iq\theta)$ under $x^{25} \to x^{25} + 2\pi R$. Physical quantities remain periodic. The canonical momentum is now shifted, so we have:

$$v_{25} = p_{25} = \frac{2\pi l + q\theta}{2\pi R}\,, \tag{8.6.7}$$

the same result is obtained for gauge-invariant momentum v_{25} as before.

Now we return to strings and introduce $U(n)$ Chan-Paton factors. A general constant A_{25} can be diagonalized by a gauge transformation:

$$A_{25} = -\frac{1}{2\pi R}\,\mathrm{diag}(\theta_1, \theta_2, \ldots, \theta_n)\,. \tag{8.6.8}$$

This gauge field lies in the diagonal group of $U(n)$, i.e., $U(1)^n$. The coupling of the gauge field with a general state's Chan-Paton factor is $[A_M, \lambda]$, so strings in the Chan-Paton state $|ij\rangle$ have charge $+1$ under $U(1)_i$, charge -1 under $U(1)_j$, and are neutral otherwise. Thus it has:

$$v_{25} = \frac{2\pi l - \theta_j + \theta_i}{2\pi R} \,. \tag{8.6.9}$$

Then the open string spectrum is:

$$m^2 = \frac{(2\pi l - \theta_j + \theta_i)^2}{4\pi^2 R^2} + \frac{1}{\alpha'}(N-1) \,. \tag{8.6.10}$$

Examine the gauge boson, which has $l = 0$ and $N = 1$, so:

$$m^2 = \frac{(\theta_j - \theta_i)^2}{4\pi^2 R^2} \,. \tag{8.6.11}$$

In a general background, all θ are different, and only diagonal vectors with $i = j$ are massless. In this case, the unbroken gauge group is $U(1)^n$. If r of the θ are equal, the corresponding vector $r \times r$ matrices are massless and carry $U(r)$ gauge symmetry. Assigning n of the θ values to sets r_i, the gauge symmetry is:

$$U(r_1) \times \ldots \times U(r_s), \quad \sum_{i=1}^{s} r_i = n \,. \tag{8.6.12}$$

As before, this has a low-energy interpretation. The gauge field A_{25} is a 25-dimensional scalar in the adjoint representation of $U(n)$, and its vacuum expectation value breaks symmetry down to $U(r_1) \times \ldots \times U(r_s)$.

Readers can fill in the details of the low-energy effective action. Note that if there are k compact dimensions, the potential will contain:

$$\mathrm{Tr}\Big([A_m, A_n]^2\Big) \,. \tag{8.6.13}$$

which comes from the field strength in the 26-dimensional Yang-Mills action. This forces gauge fields in different directions to commute in static solutions, allowing them to be simultaneously diagonalized. There are kn moduli from gauge fields, and k^2 moduli from the metric and antisymmetric tensor.

T-duality

Now examine the $R \to 0$ limit of the open string spectrum. If open string boundaries are Neumann boundaries, there is no quantum number analogous to w; they can always be untied from compact dimensions. So, as $R \to 0$, states with non-zero momentum acquire infinite mass, but there are no new continuous states. This is consistent with field theory behavior: corresponding states move in 25 spacetime dimensions.

We know that theories with open strings must have closed strings, which seems to cause a contradiction. It makes closed strings move in 26 spacetime dimensions while open strings move in 25 spacetime dimensions in the $R \to 0$ limit. From this it can be inferred that: the interior of an open string is identical to a closed string, composed of the same "matter", so it still vibrates in 26 dimensions. What distinguishes open strings is exactly their endpoints, so they must be restricted to a 25-dimensional hyperplane. Indeed, we can use new embedding coordinates to describe the T-duality theory:

$$X'^{25}(z, \bar{z}) = X_L^{25}(z) - X_R^{25}(\bar{z}) \,. \tag{8.6.14}$$

Then:

$$\partial_n X^{25} = -i\partial_t X'^{25} \,, \tag{8.6.15}$$

where n is the normal direction to the boundary, and t is the tangent direction to the boundary. Neumann boundaries in the original coordinates become Dirichlet boundaries in the dual coordinates: the X'^{25} coordinate of each string endpoint is fixed.

Let us first examine compactification without Wilson lines. Then all endpoints are constrained to the same hyperplane. To see this, integrate:

$$\begin{aligned} X'^{25}(\pi) - X'^{25}(0) &= \int_0^\pi \mathrm{d}\sigma^1\, \partial_1 X'^{25} = -i\int_0^\pi \mathrm{d}\sigma^1\, \partial_2 X^{25} \\ &= -2\pi\alpha' v^{25} = -\frac{2\pi\alpha' l}{R} = -2\pi l R' \,. \end{aligned} \tag{8.6.16}$$

The total change of X'^{25} between two endpoints is an integer multiple of the dual dimension period $2\pi R'$, so the endpoints lie on the same hyperplane in the periodic T-dual space. For two different open strings, exchanging a graviton between them yields a connected worldsheet. We can apply the same argument as in (8.6.16) to this, so all string endpoints lie on the same hyperplane, as seen in Figure 8.6.1. Endpoints move freely in the other 24 dimensions.

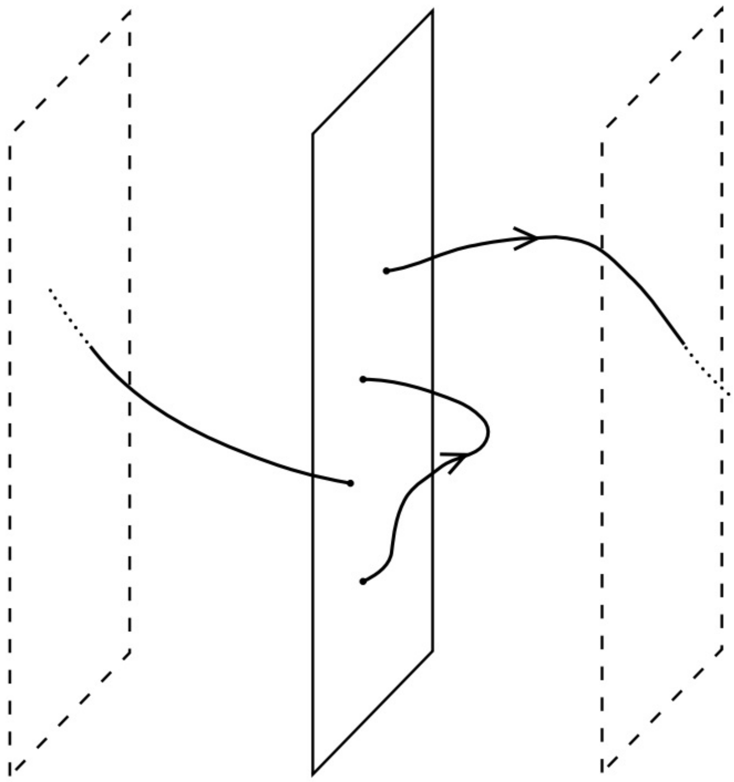


Figure 8.6.1: Open strings with endpoints attached to a hyperplane. The planes marked with dashed lines are periodically equivalent planes. Two strings are shown in the figure, with winding numbers 1 and 0 respectively.

Now let us look at the effect of Wilson lines. Due to the shift in v_{25} in (8.6.9), (8.6.16) is replaced by:

$$\Delta X'^{25} = X'^{25}(\pi) - X'^{25}(0) = -(2\pi l - \theta_j + \theta_i)R' \,. \tag{8.6.17}$$

In other words, discarding an extra normalization, endpoints on state i are at:

$$X'^{25} = \theta_i R' = -2\pi\alpha' A_{25,ii} \,. \tag{8.6.18}$$

Thus, there are generally n different hyperplanes, as seen in Figure 8.6.2.

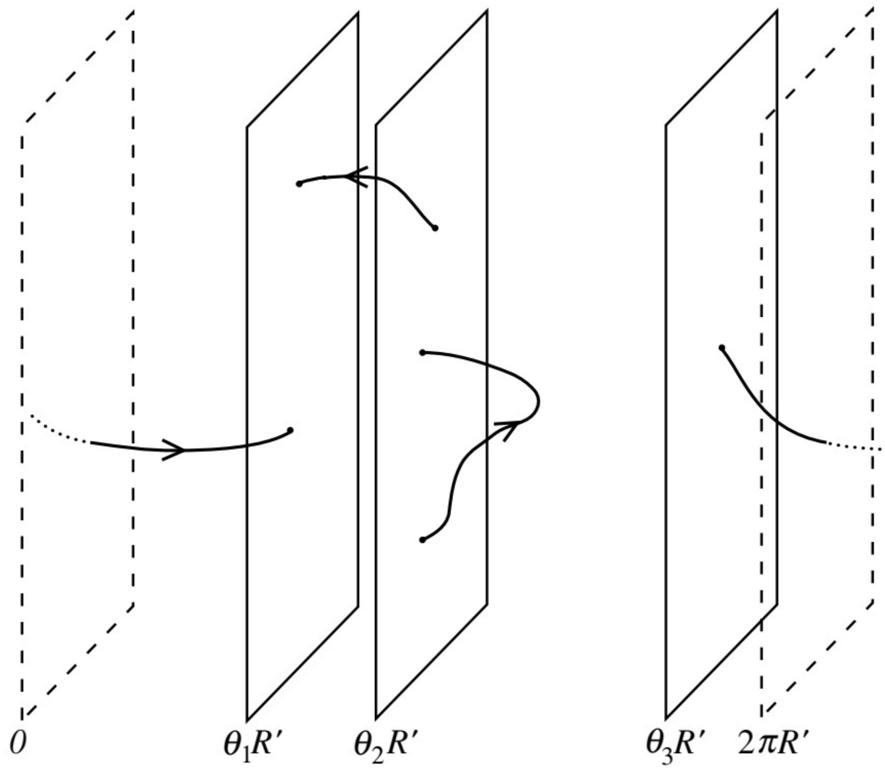


Figure 8.6.2: $n = 3$ hyperplanes at different positions, with different strings attached. Ignoring tachyons, the lightest strings with both ends on the same hyperplane are massless, while strings connecting two non-coincident hyperplanes are massive due to tension. When two hyperplanes coincide, the lightest strings connecting them become massless.

Let us examine the mode expansion for the Chan-Paton state $|ij\rangle$, assuming it winds l times around the compact dimension:

$$\begin{aligned} X'^{25}(z,\bar{z}) &= \theta_i R' - \frac{iR'}{2\pi}(2\pi l - \theta_j + \theta_i)\ln(z/\bar{z}) + i\left(\frac{\alpha'}{2}\right)^{1/2} \sum_{\substack{m=-\infty \\ m\neq 0}}^{\infty} \frac{\alpha_m^{25}}{m}(z^{-m} - \bar{z}^{-m}) \\ &= \theta_i R' + \frac{\sigma^1}{\pi}\Delta X'^{25} - (2\alpha')^{1/2} \sum_{\substack{m=-\infty \\ m\neq 0}}^{\infty} \frac{\alpha_m^{25}}{m} \exp\bigl(-m\sigma^2\bigr)\sin m\sigma^1 \,. \end{aligned} \tag{8.6.19}$$

The spectrum (8.6.10) becomes:

$$m^2 = \left(\frac{\Delta X'^{25}}{2\pi\alpha'}\right)^2 + \frac{1}{\alpha'}(N-1)\,, \tag{8.6.20}$$

where $\Delta X'^{25}$ is given by (8.6.17). It is the shortest length of a string in a given sector.

Now examine how this picture generalizes if there are multiple compact dimensions. Let X^m be periodic dimensions, where $p+1 \leq m \leq 25$. We write periodic dimensions in the form of dual coordinates. Neumann conditions on X^m again become Dirichlet conditions on dual coordinates X'^m, so open string endpoints are restricted to n parallel $(p+1)$-dimensional subspaces. For each subspace, Wilson lines in different directions become independent coordinates. Since translation invariance is broken by open string boundary conditions, the dual background is quite strange. This reflects the fact that: in the original open string theory, winding number is not conserved, and winding number T-duals to momentum. T-duality is just a different description of the same theory. Furthermore, when the compactification radius is very small, the T-duality picture should naturally be used.

8.7 D-branes

We now know that when an open string theory is compactified on a small torus, its physics can be described as compactification on a large torus, but with open string endpoints constrained to subspaces. In fact, these subspaces themselves are new dynamical objects. Consider the spectrum of massless open strings in a general configuration, where all θ_i are different, and for simplicity, only one coordinate is dualized. Ignoring tachyons, mass (8.6.20) is only zero when $N = 1$ and $\Delta X'^{25} = 0$. That is, both endpoints are on the same hyperplane and the winding number is zero. Thus we have massless states:

$$\alpha^{\mu}_{-1}|k;ii\rangle\,, \quad \mathscr{V} = i\partial_t X^{\mu}\,, \tag{8.7.1a}$$

$$\alpha^{25}_{-1}|k;ii\rangle\,, \quad \mathscr{V} = i\partial_t X^{25} = \partial_n X'^{25}\,. \tag{8.7.1b}$$

They are obviously massless states in the original theory; T-duality just provides a different view of the same spectrum. For general Wilson lines, the only massless open string states will be n massless $U(1)$ vectors. In (8.7.1), we classified them according to whether they are tangent or normal to the hyperplane. The 25 states with tangential polarizations constitute a gauge field on the hyperplane, which has a simple and important interpretation in the T-duality theory: it is a set of coordinates for the shape of the hyperplane. This can be seen from the fact that a constant gauge field (8.6.18) corresponds to a uniform translation of the hyperplane. A background depending on x^μ can shift to a translation depending on x^μ, a curved hyperplane, and the quantum of field A^{25}_{ii} corresponds to a small oscillation of the hyperplane.

This is identical to phenomena occurring in spacetime. We start with strings in a flat background and discover that massless closed string states correspond to fluctuations in geometry. Here, we first discover hyperplanes, and then discover that specific open string states correspond to fluctuations of the hyperplanes. We should not be surprised that hyperplanes have dynamics. String theory includes gravity. Gravitational waves passing through a hyperplane will perturb spacetime itself, so it is difficult for hyperplanes to remain rigid. Thus, hyperplanes are indeed dynamical objects; these are D-branes. Performing duality on $25 - p$ dimensions results in p-dimensional branes, called Dp-branes. Using this terminology, the original $U(n)$ open string theory contains n D25-branes. D25-branes fill the space, so there are no constraints on string endpoints: it just corresponds to general Chan-Paton factors. Since T-duality swaps Neumann and Dirichlet boundary conditions, T-duality in a direction tangent to a Dp-brane reduces it to a D(p−1)-brane, while T-duality in a normal direction reduces it to a D(p+1)-brane. Non-trivial angle cases appear immediately.

We could have started by studying Dirichlet boundaries themselves. Here the route we took was discovering them from T-duality, a method that better develops their properties. It can also be proven that through a continuous process, we can move from "ordinary" states to states containing D-branes. That is, taking the $R \to 0$ limit of the ordinary theory, which results in n D-branes in non-compact space. In superstrings, we will use this to argue that D-branes are an important part of the non-perturbative definition of the theory. For bosonic strings, since there is no strong evidence that it has a non-perturbative definition, we do not discuss this. Bosonic string theory possibly exists only as a part of superstrings.

Let us look at $U(n)$ symmetry breaking in the T-duality picture. When no D-branes coincide, there are massless vectors only on each D-brane, yielding the gauge group $U(1)^n$. If r D-branes coincide, there will be extra massless states, with the length of the strings connecting these branes being zero: when $n = 0$, if both i and j are in set r, $\Delta X'^{25}$ in mass formula (8.6.20) is zero. Thus, there are r^2 vectors, forming the adjoint representation of the $U(r)$ gauge group. This is the same as discovered in the dual Wilson line picture. However, surprisingly, there will be r^2 massless scalars in directions normal to the D-branes. We will discuss its meaning below.

D-brane Action

On the worldvolume of a Dp-brane, the massless fields are $U(1)$ vectors and $25-p$ worldvolume scalars describing fluctuations. The low-energy effective action for this system is always worth examining. We take the dual radius R' to be infinite and only consider a single D-brane in the corresponding 26-dimensional spacetime. Worldvolume fields interact with massless closed string fields, and its action was discussed in section 3.7. Introduce coordinates ξ^a on the brane, where $a = 0, \dots p$. Fields on the brane are the embedding $X^\mu(\xi)$ and the gauge field $A_a(\xi)$. We declare the action to be:

$$\boldsymbol{S}_p = -T_p \int \mathrm{d}^{p+1}\xi \, \mathrm{e}^{-\Phi}[-\det(G_{ab} + B_{ab} + 2\pi\alpha' F_{ab})]^{1/2} \,, \tag{8.7.2}$$

where T_p is a constant to be determined. Here:

$$G_{ab}(\xi) = \frac{\partial X^\mu}{\partial \xi^a}\frac{\partial X^\nu}{\partial \xi^b} G_{\mu\nu}(X(\xi)), \quad B_{ab}(\xi) = \frac{\partial X^\mu}{\partial \xi^a}\frac{\partial X^\nu}{\partial \xi^b} B_{\mu\nu}(X(\xi)) \tag{8.7.3}$$

are the induced metric and induced antisymmetric tensor on the brane. All features of (8.7.2) can be understood based on general reasoning. First, considering only the spacetime metric and embedding, the simplest coordinate-independent action with fewest derivatives is the integral of $(-\det G_{ab})^{1/2}$, i.e., the worldvolume. Note that this term has an implicit relationship to $X^\mu(\xi)$ via the induced fields (8.7.3). Expanding around a flat D-brane gives the action for fluctuations, just as the Nambu-Goto action describes string fluctuations.

The dependence on the dilaton $\mathrm{e}^{-\Phi} \propto g_\mathrm{c}^{-1}$ appears because this is an open string tree action. Self-interactions of open string fields, and their coupling with closed string fields, originally come from the disk.

The dependence on F_{ab} can be understood through T-duality. Consider a D-brane spread along X^1 and X^2 directions, with other directions unspecified, and let there be a constant gauge field F_{12} on it. Go to the gauge $A_2 = X^1 F_{12}$. Now take T-duality along the X^2 direction. The Neumann condition in this direction becomes a Dirichlet condition, so the D-brane loses one dimension. However, the relationship between the potential and coordinates (8.6.18) implies that the D-brane is tilted in the (1-2)-plane:

$$X'^2 = -2\pi\alpha' X^1 F_{12} \,. \tag{8.7.4}$$

This tilt gives the action a geometric factor:

$$\int \mathrm{d}X^1 \left[1 + (\partial_1 X'^2)^2\right]^{1/2} = \int \mathrm{d}X^1 \left[1 + (2\pi\alpha' F_{12})^2\right]^{1/2} . \tag{8.7.5}$$

For any D-brane, the action can be reduced to the product of factors (8.7.5) in each plane by pushing it to align with the coordinate axes and rotating, equivalent to the F_{ab} term in the determinant. The determinant composed of gauge fields is called the Born-Infeld action, which was originally intended to solve the short-range divergence problem of quantum electrodynamics.

Finally, the dependence on B_{ab} is given by the following discussion. In the worldsheet action of strings, closed string field $B_{\mu\nu}$ and open string field A^μ appear in the following form: [2]

$$\frac{i}{4\pi\alpha'} \int_M \mathrm{d}^2\sigma \, g^{1/2} \epsilon^{ab} \partial_a X^\mu \partial_b X^\nu B_{\mu\nu} + i \int_{\partial M} \mathrm{d}X^\mu \, A_\mu \,. \tag{8.7.6}$$

[2] For readers familiar with differential forms, this is:

$$\frac{i}{2\pi\alpha'} \int_M B + i \int_{\partial M} A \,.$$

The subsequent equation can be translated in a similar way.

Each of these fields is accompanied by a spacetime gauge invariance, which must be preserved to be consistent with spacetime theory. The usual gauge transformation:

$$\delta A_\mu = \partial_\mu \lambda \tag{8.7.7}$$

is an invariance of action (8.7.6), where the boundary term differs by the integral of a total derivative. The antisymmetric tensor variation:

$$\delta B_{\mu\nu} = \partial_\mu \zeta_\nu - \partial_\nu \zeta_\mu \tag{8.7.8}$$

makes the "bulk" action differ by a total derivative, but on a worldsheet with boundaries, this yields a surface term. To make it cancel, under tensor gauge symmetry, the open string field A_μ must transform as:

$$\delta A_\mu = -\zeta_\mu / 2\pi\alpha' \,. \tag{8.7.9}$$

Now, only the combination:

$$B_{\mu\nu} + 2\pi\alpha' F_{\mu\nu} \equiv 2\pi\alpha' \mathscr{F}_{\mu\nu} \tag{8.7.10}$$

is invariant under both symmetries, so it must be this combination that appears in the action. Thus, the form of the action (8.7.2) is completely determined. As a check, since the same combination appears in the worldsheet action under conformal gauge, it is natural that the combination $G_{\mu\nu} + B_{\mu\nu}$ should appear. The action (8.7.2) could also be determined by taking low-energy limits of various open string and open-closed string amplitudes, but this would be much more laborious.

For n discrete D-branes, the action is n copies of a single D-brane action. However, we have seen that when D-branes coincide, there will be n^2 massless vectors and scalars on the brane instead of n, and we will write the effective action for this case. Fields $X^\mu(\xi)$ and $A_a(\xi)$ will now be $n \times n$ matrices. For gauge fields, the meaning is obvious—it becomes a non-Abelian $U(n)$ gauge field. However, for coordinates X^μ, the meaning is unclear: the embedding coordinate set for n D-branes into spacetime extends into $n \times n$ matrices. Non-commutative geometry has been proven to play an important role in the dynamics of D-branes, and here is a conjecture: it is an important clue to the characteristics of spacetime.

We can obtain more enlightenment by examining the effective action. There are now non-derivative terms in the action, which is the potential for the coordinates, and this can be inferred from T-duality for constant A_m fields. For such vector potentials, the field strength is $[A_m, A_n]$, which becomes $(2\pi\alpha')^{-2}[X_m, X_n]$ in the T-duality picture. Expanding in the field strength, the leading order of the action is approximately:

$$V \propto \mathrm{Tr}([X_m, X_n][X^m, X^n]) \,. \tag{8.7.11}$$

The second derivative of this potential at $X^m = 0$ is zero, making all kn^2 scalars massless; as before $k = 25-p$ is the dual dimension. However, the space of flat dimensions is smaller. The potential is a sum of squares, so it is zero only when all $[X^m, X^n]$ are zero. We can go to the gauge where X^m are all diagonal using $U(n)$ symmetry. Thus there will be kn flat directions, exactly the number of diagonal elements. In this way, the kn diagonal elements can be interpreted as the coordinates for n Dp-branes. Thus (8.7.11) correctly intervenes in the physics of separating D-branes and coinciding D-branes.

When X^m commute, the action should reduce to n separate D-branes, thus:

$$\begin{aligned} \boldsymbol{S}_p = -T_p \int \mathrm{d}^{p+1}\xi \ \mathrm{Tr}\Big\{ &\mathrm{e}^{-\Phi}[-\det(G_{ab} + B_{ab} + 2\pi\alpha' F_{ab})]^{1/2} \\ &+ O([X^m, X^n]^2) \Big\} \,. \end{aligned} \tag{8.7.12}$$

The determinant is taken over worldvolume indices ab, and the trace is taken over n Chan-Paton indices. The trace is a suitable $U(n)$ invariant, which for diagonal matrices reduces to the sum over separate D-branes. The complete dependence on the commutator is much more complex than the simple potential (8.7.11), involving couplings with other fields and higher-order corrections to the commutator. However, the key property, the form of the flat directions, is unaffected. By starting from the Born-Infeld action in the complete Neumann case and performing T-duality, we can obtain the full dependence. Incidentally, higher-order derivative terms containing field strength commutators cannot be determined just by T-duality.

D-brane Tension

Calculating the constant T_p is very meaningful, and for superstrings, its exact value is very important. Before performing explicit calculation, note that the recurrence relation for T_p can be obtained from T-duality. Note that in a constant dilaton background, the tension of a Dp-brane is $T_p \mathrm{e}^{-\Phi}$: this is the negative of the action per unit volume for a static Dp-brane. Now examine such a Dp-brane where p directions tangent to the Dp-brane are periodically equivalent. That is, the Dp-brane is wrapped on a p-torus in spacetime. Its mass is its tension multiplied by the area of its torus:

$$T_p \mathrm{e}^{-\Phi} \prod_{i=1}^{p} (2\pi R_i)\ . \tag{8.7.13}$$

Now, take T-duality along one of the periodic directions X^p. This does not change the mass, just a new description of the same state. Using the dilaton of the T-dual theory (8.3.31), the mass (8.7.13) is:

$$2\pi\alpha'^{1/2} T_p \mathrm{e}^{-\Phi'} \prod_{i=1}^{p-1} (2\pi R_i)\ . \tag{8.7.14}$$

Remark. *Utilized* $\sqrt{\alpha'}\mathrm{e}^{-\Phi'}/R_p = \mathrm{e}^{-\Phi}$.

However, in the dual theory, this corresponds to a D(p−1)-brane wrapped on a (p−1)-torus, so its mass is:

$$T_{p-1} \mathrm{e}^{-\Phi'} \prod_{i=1}^{p-1} (2\pi R_i)\ . \tag{8.7.15}$$

Combining (8.7.14) and (8.7.15) yields:

$$T_p = T_{p-1}/2\pi\alpha'^{1/2}\ , \tag{8.7.16}$$

which completely determines T_p except for a total normalization.

To determine the actual value of D-brane tension, we need to calculate string amplitudes. For example, we could infer it from the gravitational coupling of D-branes, which is given by a disk with a graviton vertex operator. However, this involves an unknown ratio $g_\mathrm{c}/g_\mathrm{o}^2$, since closed string coupling comes from vertex operators and open string coupling from disks. We can obtain the absolute normalization through another method. Consider two parallel Dp-branes at $X_1^m = 0$ and $X_2^m = y^m$ respectively. By exchanging closed strings, these two can sense each other's existence, as shown in Figure 8.7.1. The string amplitude is an annulus without vertex operators, which can be calculated using the method from the previous chapter. Then the pole given by exchanging gravitons and dilatons gives the coupling T_p between closed string states and D-branes.

In section 7.4, we calculated the annulus vacuum amplitude using an open string loop, but discovered closed string poles from the $t \to 0$ limit of the modular integral. This is exactly the pole we are interested in, but the simplest way to calculate the amplitude is to treat it as an open string loop. In fact, with a slight change, the previous result (7.4.1) can be used. The number of momentum integrals decreases from 26 to p+1, and similarly, V_{26} becomes V_{p+1} ; weights

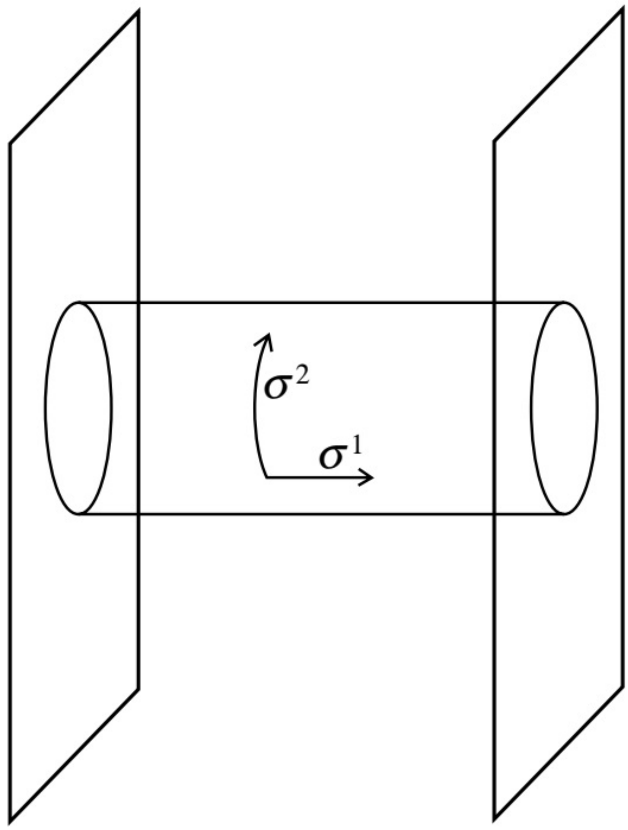


Figure 8.7.1: Exchanging a closed string between two D-branes. Equivalently, this is the vacuum loop for an open string with endpoints attached to two D-branes.

h_i acquire an extra term $y^2/4\pi^2\alpha'$ due to stretched open strings; the Chan-Paton weight n^2 becomes 2 (since there are two directions for connecting the strings). Therefore we have:

$$\begin{aligned}\mathscr{A} &= iV_{p+1}\int_0^\infty \frac{\mathrm{d}t}{t}(8\pi^2\alpha' t)^{-(p+1)/2}\exp\bigl(-ty^2/2\pi\alpha'\bigr)\eta(it)^{-24} \\ &= \frac{iV_{p+1}}{(8\pi^2\alpha')^{(p+1)/2}}\int_0^\infty \mathrm{d}t\, t^{(21-p)/2}\exp\bigl(-ty^2/2\pi\alpha'\bigr) \\ &\qquad\qquad \times\left[\exp(2\pi/t)+24+\cdots\right], \end{aligned} \tag{8.7.17}$$

where the asymptotic expansion is obtained by the method in section 7.4. The first term is tachyon exchange, thus it is a bosonic artifact. Integrating the second term gives:

$$\begin{aligned}\mathscr{A} &= iV_{p+1}\frac{24}{2^{12}}(4\pi^2\alpha')^{11-p}\pi^{(p-23)/2}\Gamma\left(\frac{23-p}{2}\right)|y|^{p-23} \\ &= iV_{p+1}\frac{24\pi}{2^{10}}(4\pi^2\alpha')^{11-p}G_{25-p}(y)\,, \end{aligned} \tag{8.7.18}$$

where $G_d(y)$ is the Green function for a massless scalar in d dimensions, the inverse of $-\nabla^2$. We now want to compare this with the field theory calculation of the same amplitude. Since antisymmetric tensors do not couple to D-branes, from the spacetime action (**??**), the relevant terms are:

$$\boldsymbol{S} = \frac{1}{2\kappa^2}\int \mathrm{d}^{26}X\,(-\tilde{G})^{1/2}\left(\tilde{\boldsymbol{R}} - \frac{1}{6}\nabla_\mu\tilde{\Phi}\tilde{\nabla}^\mu\tilde{\Phi}\right). \tag{8.7.19}$$

The tildes denote the Einstein metric. Because this form of action decouples from the dilaton, it is convenient. The dilaton with tildes has been shifted such that its vacuum expectation value is zero. In terms of the same variable, the relevant term in the D-brane action (8.7.2) is:

$$\boldsymbol{S}_p = -\tau_p\int \mathrm{d}^{p+1}\xi\ \exp\left(\frac{p-11}{12}\tilde{\Phi}\right)(-\det\tilde{G}_{ab})^{1/2}\,. \tag{8.7.20}$$

We defined $\tau_p = T_p\mathrm{e}^{-\Phi_0}$; when the background value of the dilaton is Φ_0, this is the true physical tension of the Dp-brane.

The field theory diagram analogous to Figure 8.5 is the exchange of dilatons or gravitons between D-branes. To obtain the propagator, we expand the spacetime action up to second order in $h_{\mu\nu} = \tilde{G}_{\mu\nu} - \eta_{\mu\nu}$ and $\tilde{\Phi}$. Additionally, for gravity calculation, we need to choose a gauge. The simplest gauge for perturbative calculation is:

$$F_\nu \equiv \partial^{\hat{\mu}}h_{\mu\nu} - \frac{1}{2}\partial_\nu h^{\hat{\mu}}{}_\mu = 0\,, \tag{8.7.21}$$

where the "hat" indicates using the flat metric $\eta^{\mu\nu}$ to raise and lower indices. Expanding the action to second order and adding the gauge fixing term $-F_\nu F^{\hat{\nu}}/4\kappa^2$, the spacetime action becomes:

$$\boldsymbol{S} = -\frac{1}{8\kappa^2}\int \mathrm{d}^{26}X\left(\partial_\mu h_{\nu\lambda}\partial^{\hat{\mu}}h^{\hat{\nu}\hat{\lambda}} - \frac{1}{2}\partial_\mu h^{\hat{\nu}}{}_\nu\partial^{\hat{\mu}}h^{\hat{\lambda}}{}_\lambda + \frac{2}{3}\partial_\mu\tilde{\Phi}\partial^{\hat{\mu}}\tilde{\Phi}\right). \tag{8.7.22}$$

Observe the kinetic terms, yielding the momentum space propagator:

$$\langle\tilde{\Phi}\tilde{\Phi}\rangle = -\frac{(D-2)i\kappa^2}{4k^2}\,, \tag{8.7.23a}$$

$$\langle h_{\mu\nu}h_{\sigma\rho}\rangle = -\frac{2i\kappa^2}{k^2}\left(\eta_{\mu\sigma}\eta_{\nu\rho} + \eta_{\mu\rho}\eta_{\nu\sigma} - \frac{2}{D-2}\eta_{\mu\nu}\eta_{\sigma\rho}\right). \tag{8.7.23b}$$

For later reference, we gave general D. Expanding around a general flat configuration, the D-brane action is:

$$\boldsymbol{S}_p = -\tau_p \int \mathrm{d}^{p+1}\xi \left(\frac{p-11}{12}\tilde{\Phi} - \frac{1}{2}h_{aa} \right) . \tag{8.7.24}$$

Note that $h_{\mu\nu}$ here is traced only over directions tangent to the D-brane; we have taken ξ as Cartesian coordinates with metric δ_{ab}. We can now read the Feynman diagram:

$$\begin{aligned} \mathscr{A} &= \frac{i\kappa^2\tau_p^2}{k_\perp^2} V_{p+1} \left\{ 6\left[\frac{p-11}{12}\right]^2 + \frac{1}{2}\left[2(p+1) - \frac{1}{12}(p+1)^2\right] \right\} \\ &= \frac{6i\kappa^2\tau_p^2}{k_\perp^2} V_{p+1} . \end{aligned} \tag{8.7.25}$$

Remark. *The above comes from* $\langle S_p S_p \rangle$.

Comparison with (8.7.18) gives:

$$\tau_p^2 = \frac{\pi}{256\kappa^2}(4\pi^2\alpha')^{11-p} . \tag{8.7.26}$$

This is consistent with the recurrence relation given by T-duality.

As an application, consider a state with n 25-branes, which is identical to the ordinary n-valued Chan-Paton factor. Expanding the 25-brane action (8.7.12) to second order in the gauge field yields:

$$\frac{\tau_{25}}{4}(2\pi\alpha')^2 \operatorname{Tr}(F_{\mu\nu}F^{\mu\nu}) . \tag{8.7.27}$$

This relates the open string gauge coupling and closed string gravitational coupling. Using (6.5.14), (6.5.16), (6.6.15), and (6.6.18), we can write this as a relationship between the vertex operator normalizations g_c and g_o:

$$\frac{g_\mathrm{o}^2}{g_\mathrm{c}} = \frac{4\pi\alpha' g_\mathrm{o}'^2}{\kappa} = 2^{18}\pi^{25/2}\alpha'^6 . \tag{8.7.28}$$

It has the correct form, with the square of the open string coupling being proportional to the closed string coupling, but now with a numerical coefficient. We decompose closed string poles from the torus via unitarity, or see Exercise 7.9.

8.8 T-duality of Unoriented Theories

In unoriented string theory, the $R \to 0$ limit also yields interesting new physics. First consider the closed string theory. To form an unoriented theory, we impose $\Omega = +1$ on the states. Following the discussion in section 8.5, we can also think of this as gauging Ω. In particular, the transition functions used to construct the worldsheet now reverse directions, and this produces an unoriented worldsheet.

The T-dual theory is obtained by replacing $X^m(z,\bar{z}) = X_L^m(z) + X_R^m(\bar{z})$ with coordinates $X'^m(z,\bar{z}) = X_L^m(z) - X_R^m(\bar{z})$. In the original description, we gauge worldsheet parity Ω, whose action is:

$$\Omega : X_L^M(z) \leftrightarrow X_R^M(z) . \tag{8.8.1}$$

In T-dual coordinates, this is:

$$\Omega : \quad X'^m(z,\bar{z}) \leftrightarrow -X'^m(\bar{z},z) , \tag{8.8.2a}$$

$$X^\mu(z,\bar{z}) \leftrightarrow X^\mu(\bar{z},z) , \tag{8.8.2b}$$

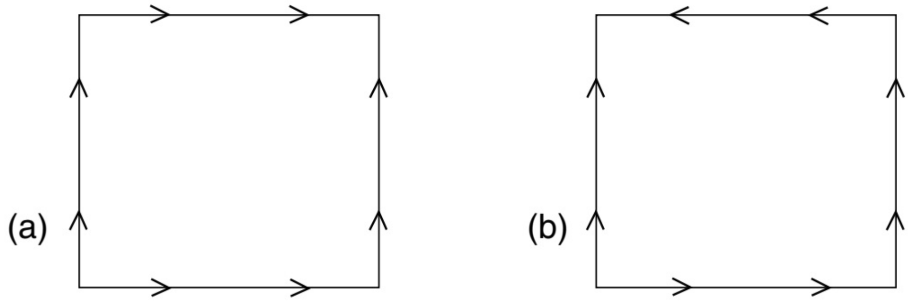


Figure 8.8.1: Gluing opposite sides gives (a) a torus and (b) a Klein bottle.

as before, m labels coordinates for which T-duality is taken, and μ labels coordinates for which T-duality is not taken. In the dual picture, symmetry (8.8.2) is thus the product of worldsheet parity transformation and spacetime inversion. We see that gauging worldsheet parity only produces an unoriented theory, while gauging inversion only produces an orbifold. The result of the combined projection is called an *orientifold*.

An orientifold is very similar to an orbifold. Divide the string wave function into its interior and its center-of-mass x^m dependent part, taking the interior wave function to be an eigenstate of Ω. Then the projection onto $\Omega = +1$ indicates that the string wave function at $-x^m$ is the same as at x^m, differing only by a sign. For example, the components of the metric and antisymmetric tensor satisfy:

$$G_{\mu\nu}(x') = G_{\mu\nu}(x)\,, \quad B_{\mu\nu}(x') = -B_{\mu\nu}(x)\,, \tag{8.8.3a}$$
$$G_{\mu n}(x') = -G_{\mu n}(x)\,, \quad B_{\mu n}(x') = B_{\mu n}(x)\,, \tag{8.8.3b}$$
$$G_{mn}(x') = G_{mn}(x)\,, \quad B_{mn}(x') = -B_{mn}(x)\,, \tag{8.8.3c}$$

where $(x^\mu, x^m)' = (x^\mu, -x^m)$. There is one negative sign for each of m, n and another negative sign for B_{MN}; this is identical to an orbifold. In other words, the T-dual spacetime is the torus T^k modulo the $\mathbb{Z}_2$ reflection in k compact dimensions, exactly the same as the construction of an orbifold. For example, if there is only one periodicity, the dual spacetime is the segment $0 \le x^{25} \le \pi R'$, with orientifold fixed planes at the ends.

Note that when away from the orientifold fixed plane, the local physics is that of an oriented string theory. Unlike the original unoriented theory, where projection locally removes half of the states, here it relates the string wave function at any point to its value at the mirror point, as in (8.8.3). One difference between orientifold and orbifold constructions is that the former has no direct analogue of twisted states because the Klein bottle has no modular transformation $\tau \to -1/\tau$. Examine Figure 8.6, which shows the torus and the Klein bottle. On the torus, the projection operator inserts a twist in the time-like direction in Figure 8.6a; rotating this figure by 90°, this becomes a twist in the spatial direction, implying the existence of twisted states in the spectrum. However, if the Klein bottle in Figure 8.6b is rotated by 90°, the time directions on relative sides cannot match. So there is no intermediate state interpretation in this channel. Thus it should be noted that the orientifold plane cannot be dynamical. Unlike the D-brane case, no string modes are tethered to the orientifold plane to represent its shape fluctuations. Our heuristic discussion—that gravitational waves force D-branes to oscillate here—does not apply to orientifold planes. Ultimately, the equivalence relation (8.8.3) becomes a boundary condition on fixed planes, causing incident waves to cancel with reflected waves. For D-branes, the reflected wave is a higher-order quantity of the string coupling.

In drawing oriented strings, we use arrows to indicate directions. In unoriented theory, we can either omit the arrows or take a linear combination of the two directions. The latter picture

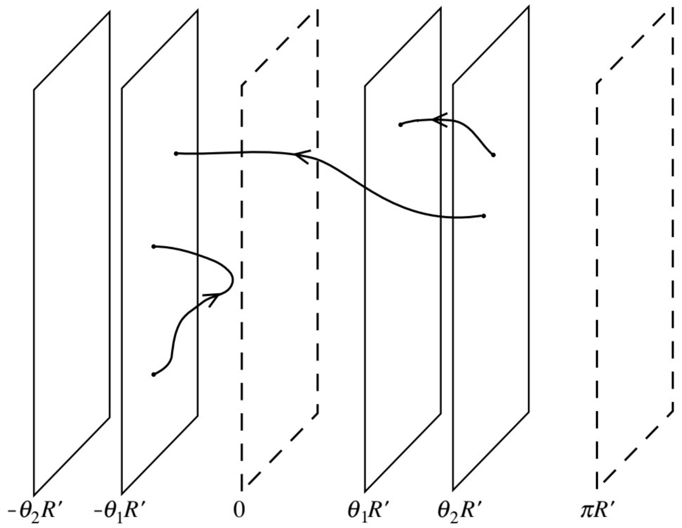


Figure 8.8.2: Orientifold planes are at 0 and $\pi R'$, D-branes are at $\theta_1 R'$ and $\theta_2 R'$, and D-brane images are at $-\theta_1 R'$ and $-\theta_2 R'$. The effect of the twisting operator Ω on any string is a combination of spacetime reversal and reversal of oriented arrows.

is more consistent with the idea of gauging discrete symmetries and is more widespread: in an orientifold, parity operators must be accompanied by a spacetime transformation, so we cannot simply forget the arrows.

Although we introduced orientifold construction through T-duality, one can also examine orientifolds that are not T-duals of toroidal compactifications, combining worldsheet parity with various spacetime symmetries to gauge a group.

Open Strings

In the case of open strings, the situation is similar. For simplicity, we only examine the case of one compact dimension. As before, there are orientifold fixed planes at 0 and $\pi R'$. Introduce $SO(n)$ Chan-Paton factors (symplectic symmetry is similar), and a general Wilson line can be transformed into a diagonal form; when n is even, the eigenvalues appear in pairs:

$$W = \operatorname{diag}\left(\mathrm{e}^{i\theta_1}, \mathrm{e}^{-i\theta_1}, \mathrm{e}^{i\theta_2}, \mathrm{e}^{-i\theta_2}, \cdots, \mathrm{e}^{i\theta_{n/2}}, \mathrm{e}^{-i\theta_{n/2}}\right) . \tag{8.8.4}$$

Thus, in the dual picture, there are $\frac{1}{2}n$ D-branes in the segment $0 \leq X'^{25} \leq \pi R'$, and another $\frac{1}{2}n$ at mirror points. Strings can open between D-branes and mirrors, as shown in Figure 8.8.2. The general gauge group is $U(1)^{n/2}$. Similar to oriented theory, if r D-branes coincide, then a $U(r)$ gauge group exists. If these r D-branes are also on one of the fixed planes, then strings opening between one of these branes and its mirror brane are also massless. We would then have the correct spectrum of extra states to fill $SO(2r)$. If all branes are on an orientifold plane, the maximum $SO(n)$ is restored.

If n is odd, the last eigenvalue of W is ± 1, causing it to be fixed on one of the two orientifold planes in the T-dual picture. Without mirrors, this is actually a half-D-brane. Additionally, if we examine the $n = 2$ Wilson line $\operatorname{diag}(1, -1)$ (which is in $O(2)$ instead of $SO(2)$), what appears is not one D-brane and its image, but two half-D-branes.

As with D-branes, orientifold planes couple with the dilaton and metric. The amplitude is the same as in the previous section, except for replacing the annulus with Klein bottles and Möbius strips. In fact, we already performed relevant calculations in Chapter 7. There we found the total dilaton coupling cancels for $SO\left(2^{13}\right)$. In the T-dual picture, the total dilaton coupling of 2^{12} D-branes cancels (images are not included in the count). If we compactify $k = 25 - p$ dimensions, there will be 2^k fixed planes, whose coordinates are all combinations of 0 and $\pi R'_m$. Thus the effective action of a single fixed plane is:

$$2^{12-k} T_p \int \mathrm{d}^{p+1}\xi \, \mathrm{e}^{-\Phi} (-\det G_{ab})^{1/2} \,, \tag{8.8.5}$$

where integration runs over fixed planes.

Although there is no direct analogue of twisted sectors in orientifolds, in some sense, twisted states are open strings. Since this analogy is not complete, we do not tend to emphasize it: gauging Ω itself does not introduce worldsheet boundaries. However, adding open strings to an orientifold theory to cancel the dilaton tadpole has certain physical significance. This cancellation will play an important role in superstrings.