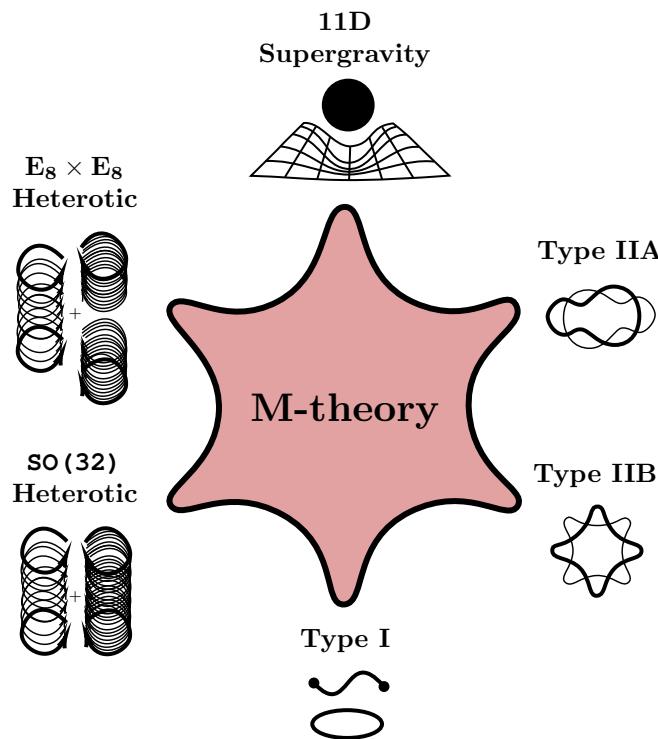


Introduction to string theory

Satoshi Nawata, Runkai Tao, Daisuke Yokoyama

Abstract

These are lecture notes of the course on string theory at Fudan University.



Contents

1 Motivation and Overview	4
1.1 Why do we study string theory?	4
1.2 Very very brief history	5
1.3 Very short highlights	6
1.4 Convention	8
2 Bosonic string theory	10
2.1 String sigma-model action	10
2.2 Quantizations	14
2.3 Path-integral formulations	17
3 Two-dimensional conformal field theory	18
3.1 Conformal transformations	19
3.2 State-operator correspondence	20
3.3 Operator product expansions in free scalar theory	21
3.4 Conformal Ward-Takahashi identity	24
3.5 Primary fields	26
3.6 Virasoro algebra	29
3.7 Vertex operators	31
4 Weyl anomaly	32
4.1 Weyl anomaly from curved world-sheet	32
4.2 Non-linear sigma model	34
5 BRST quantization	38
5.1 Quantization via path integral	38
5.2 BRST quantization	43
6 Bosonic string amplitudes	46
6.1 Teichmüller space and moduli space of Riemann surfaces	46
6.2 Gauge fixing and string amplitudes	52
6.3 Tree amplitude	55
6.4 One-loop amplitude	58
7 Superstring theories	60
7.1 RNS formulation	61
7.2 Physical spectrum and the GSO Projection	65
7.3 Torus partition functions for superstring theory	68
7.4 Modular functions	69
8 Type II superstring theories	71
8.1 Type II superstrings	71
8.2 Introduction to D-branes	72
9 T-duality	76
9.1 S^1 compactification in closed bosonic string	77
9.2 T-duality of open strings	80
9.3 T-duality of Type II superstrings	82

10 Unoriented strings and Type I theory	83
10.1 Unoriented string theory	84
10.2 Amplitude in unoriented string theory	86
10.3 T-duality of Type I theory	92
10.4 $O\!p$ -plane	93
11 Heterotic string theories	94
11.1 Bosonic construction	95
11.2 Fermionic construction	99
12 Supergravity	102
12.1 Local supersymmetry	102
12.2 $D = 11$ supergravity	103
12.3 $D = 10$ IIA supergravity	104
12.4 $D = 10$ IIB supergravity	105
13 String dualities	106
13.1 S-duality of Type IIB supergravity	107
13.2 S-duality between Type I and Heterotic $SO(32)$	108
13.3 Heterotic T-duality	110
13.4 M-theory	111
13.5 Heterotic $E_8 \times E_8$ string from M-theory	113
13.6 Duality between Heterotic on T^4 and Type IIA on K3	113
14 Green-Schwarz anomaly cancellation	117
14.1 Anomalies and the Wess-Zumino condition	117
14.2 Anomaly polynomials and decent equations	120
14.3 Green-Schwarz mechanism	125
15 D-brane dynamics	127
15.1 D-brane action	127
15.2 Branes and strings ending on branes	132
15.3 Brane creations/annihilations	136
16 Black holes in string theory	137
16.1 Black holes	138
16.2 Black hole thermodynamics and Bekenstein-Hawking entropy	140
16.3 Black holes in string theory	143
17 Introduction to AdS/CFT correspondence	146
17.1 Conformal field theory	147
17.2 Anti-de Sitter space	149
17.3 Introduction to AdS_5/CFT_4 correspondence	150
17.4 GKPW relation	155

1 Motivation and Overview

1.1 Why do we study string theory?

- **For quantum gravity**

One of the major problems in theoretical physics is to provide a unified description of all the forces in Nature. The standard model has unified electromagnetic, weak and strong force based on quantum field theories whereas general relativity for gravity is formulated within classical physics. A quantum theory of gravity is needed to reconcile general relativity with the principles of quantum mechanics. However, it is known that the renormalization procedure does not cure ultraviolet divergence for gravity. Therefore, naive quantization of general relativity does not give a consistent theory.

String theory not only yields the first quantization of gravity but also resolves the renormalization problem of gravity by replacing point particles by vibrating strings. To our knowledge, string theory is currently the only theory that has this property. Once we discover a candidate to unify gravity and quantum mechanics of all the forces, it is inevitable to try to understand it as well as we can although there is never any guarantee we can achieve.

Indeed, string theory has given deep insights into quantum gravity. The Beckenstein-Hawking formula [Bek73, BCH73, Haw75] was microscopically derived by counting D-brane states with fixed mass and charge for certain (called extremal) black holes in string theory [SV96]. The AdS/CFT correspondence [Mal99] conjectures that non-perturbative definition of string theory on AdS background is described by conformal field theory, and it partially resolves Hawking information paradox of black hole [Haw76]. Interestingly, a generalization of the Beckenstein-Hawking formula called the Ryu-Takayanagi entanglement entropy formula [RT06], connects quantum gravity and quantum information theory, which is actively studied in recent years. These developments certainly pose questions to basic concepts of spacetime.

- **Rich arena for physics theories**

String theory has had a significant impact on our understanding of established physical theories, ranging from proof of positive energy to quark confinement to quantum black holes to quantum information and more. There are several reasons for this.

First, string theory is capable of generating a vast number of quantum field theories (QFTs) in various dimensions. There are five types of string theories - Type I, IIA, IIB, Heterotic SO(32), and $E_8 \times E_8$ - and M-theory is the limit of large dilaton expectation value in Type IIA string theory. These theories can contain various types of branes, such as D_p -branes in Type II theories and M2 and M5-branes in M-theory. While there are only finitely many string theories, they can be used to "engineer" an infinite number of QFTs depending on the manifolds on which they are defined and the configurations of branes present.

Second, string theory is connected by a web of dualities, which are equivalences between different descriptions of a theory at the quantum level. The AdS/CFT correspondence is a particularly well-known example of this. These dualities can provide alternative perspectives on established physical theories and have played a crucial role in our understanding of them.

However, recent developments in the study of M5-branes have revealed that there are vast families of QFTs that do not admit a Lagrangian description and are inherently strongly coupled, making it difficult to use traditional QFT techniques to study

them. Currently, we are attempting to understand these QFTs on a case-by-case basis using techniques such as dualities, dimensional reduction on certain manifolds, RG flows, and perturbation theory. However, there is no unified framework for describing these QFTs. This motivates the search for a new approach to understanding QFTs without a Lagrangian description.

- **For mathematical structure**

String theory has garnered significant interest due to its ability to provide new insights and perspectives on mathematics, particularly geometry. This is because string theory has repeatedly given rise to new ways of understanding geometry as a natural consequence of seeking to understand physical theory. Some notable examples include mirror symmetry [GP90, CDLOGP91], Seiberg-Witten invariants [Wit94], and the AGT relation [AGT10]. In addition, string dualities have led to highly non-trivial conjectures that connect seemingly unrelated areas of mathematics.

One reason why string theory has had such a profound impact on mathematics is that it is not yet fully understood. While Einstein's theory of gravity was constructed based on Riemannian geometry, the geometric foundations of string theory are still not fully understood. However, physicists can develop new insights into geometry because physicists stumble upon a profound theory we do not understand well. There are certainly many more mathematical secrets to be uncovered within string theory.

- **Because we do not know what it is**

As previously mentioned, one of the reasons that string theory is a fascinating topic for students is that it is based on a concept that is not yet fully understood. This can sometimes lead to criticism that string theorists do not understand the theory. However, it is also true that if we did fully understand string theory, it would likely lead to major breakthroughs in both physics and mathematics. The fact that so much is still not understood and that even relatively small pieces of progress can be significant discoveries in their own right makes working on string theory an exciting prospect. Of course, there is still a lot of work to be done in this field.

1.2 Very very brief history

String theory was initially developed by accident in trying to solve a different problem, subsequently developed by a long and fortunate process of tinkering. Therefore, the history of string theory itself is fascinating. For detailed history of string theory, we refer to [Gre99, Sch00, Sch11, CCCDV12, Oog15, Pol17].

String theory was initially developed in an attempt to describe hadron physics, with the idea that mesons could be modeled as strings with charges at their ends and meson resonances as vibrational states of these strings. Although this physical picture, which first emerged from the Veneziano amplitude [Ven68], is now believed to be qualitatively correct at a description of strongly interacting particles, it was eventually surpassed by quantum chromodynamics (QCD) in explaining the details of strong interactions.

However, a small group of physicists continued to work on string theory in the 1970s and discovered that it contained massless spin-two particles, suggesting that it could potentially be used as a theory of gravity. Additionally, the incorporation of supersymmetry led to the removal of tachyons in bosonic string theory and the development of Type I, IIA, and IIB string theories.

In 1984, Green and Schwarz demonstrated [GS84, GS85] that the anomaly pointed

out by Alvarez-Gaumé and Witten could be canceled by setting the gauge group of Type I string theory to be $\text{SO}(32)$, marking the beginning of the first string theory revolution. This resulted in a surge of interest in the field and the development of Heterotic string theory.

In 1995, Witten proposed [Wit95b] that the five known types of string theory were all different limits of a single theory known as M-theory, leading to the second string theory revolution. During this time, D-branes were introduced and non-perturbative aspects of string theory were extensively studied, resulting in the discovery of various dualities, a better understanding of the quantum nature of black holes, and the development of the AdS/CFT correspondence.

The second string theory revolution has led to increased research on the world-volume theories of D-branes and their connections to quantum gravity. However, the dynamics of M2-branes and M5-branes, which are fundamental degrees of freedom in M-theory, have remained a mystery. Triggered by the work of Bagger-Lambert [BL07], the world-volume theory for M2-branes was found to be a highly supersymmetric exquisite mixture of Chern-Simons gauge theory with some scalars and spinors [ABJM08]. In contrast, the world-volume theory for M5-branes is challenging to analyze due to its lack of a Lagrangian description. However, the work of Gaiotto [Gai12] has sparked ongoing research on the properties of M5-branes by studying them when wrapped on certain manifolds.

The following picture shows the timeline of string theory and when the standard textbooks were written. Of course, there are many other books [Zwi06, Joh06, Din07, IU12, BLT13, Kir19, Tom22, Cec23] and notes [Ura05, Ton09, Wra11, Hos15, Wei15]. Obviously, the textbooks do not cover developments after their publications. We will mainly follow Polchinski's textbook [Pol98], and the reader can find more elaborate explanations in the textbooks listed above. However, I do not recommend you stick to only one book when you study a subject. Some parts are explained very well in a book, but some parts are sometimes hard to grasp for some people. Therefore, it is always a good idea to look for books, notes, and papers that suit you best.

1.3 Very short highlights

- The basic idea in string theory is that different elementary particles are all vibrational modes of a single type of string. There are two types of strings, open and closed, and a trajectory of a string is called the **string world-sheet** which is parametrized by (σ, τ) . If the typical size ℓ_s of a string is smaller than the resolution that an accelerator can provide, we cannot see this in the experiment involving elementary particles.

$$(\text{Planck scale}) \quad 10^{-33} \text{ cm} \leq \ell_s \leq 10^{-17} \text{ cm} \quad (\text{TeV scale})$$

We often use the parameter

$$\alpha' = \ell_s^2$$

which is the only free parameter in string theory. As we will see, the couplings in string theory are expectation values of dynamical fields (so-called moduli) which take their value dynamically.

- Application of the quantization rules provides us with the Fock space of string excitations. In bosonic string, the massless modes include among others

open string :	A_μ	spin 1	W-boson
closed string :	$g_{\mu\nu}$	spin 2	graviton

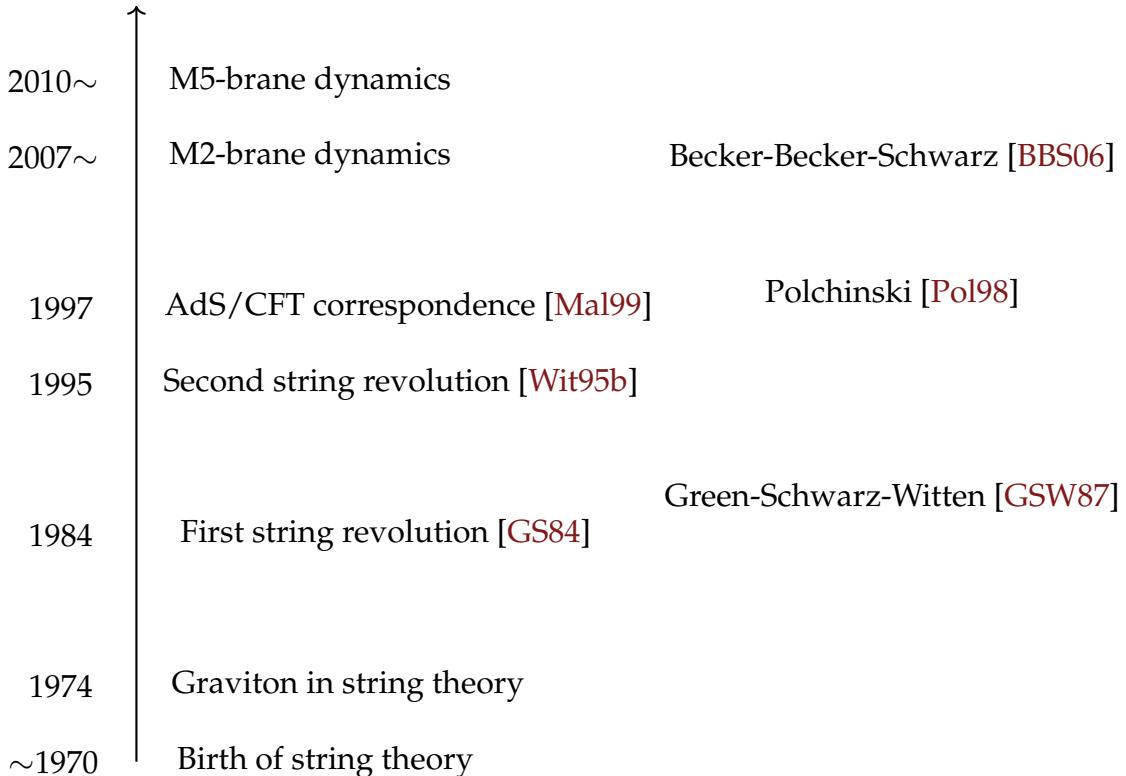


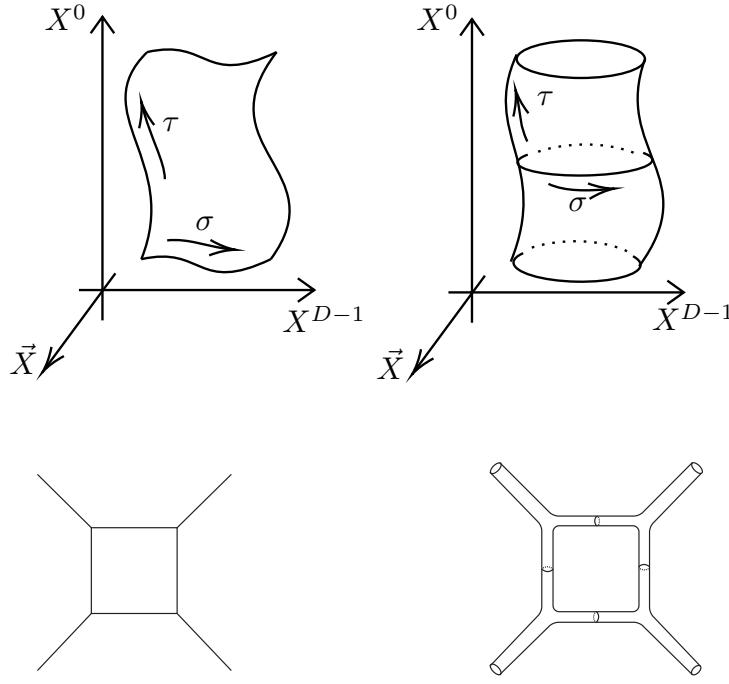
Figure 1: History of string theory and textbooks

In addition, one finds a tower of massive string excitations of mass

$$\begin{aligned} \text{open bosonic : } M^2 &= \frac{1}{\alpha'}(N - 1) \\ \text{closed bosonic : } M^2 &= \frac{4}{\alpha'}(N - 1) \quad N = 0, 1, 2, \dots \end{aligned}$$

Note that the lowest-lying state has a dimension of negative [mass]² ($N = 0$), which is called **tachyon**. The bosonic theory is consistent only when the number of spacetime dimensions is 26. We can also refer to [Wit18].

- In the point-particle picture, the divergence appears when all four vertices come close to each other. On the other hand, the string world-sheet has no vertices. Thus, when we sum over all surfaces, we do not encounter configuration analogous to collapsed vertices. String theory amplitudes have no ultraviolet (short distance) divergence. As a result, string theory provides a finite quantum theory of gravity. Moreover, this is even better than renormalizable quantum field theories since there is no divergence in the first place.
- Strings, besides vibrating in usual spacetime, can also have some internal degrees of freedom. These internal degrees of freedom are fermionic. This can be quantized in a similar manner to that for bosonic strings. Then, a superstring can be considered as a string with symmetry between bosonic states and fermionic states. In superstring theories, the good features (e.g. emergence of gravity and UV finiteness) are preserved, and one can incorporate fermions. Moreover, as mentioned, the tachyonic mode is absent.



1. the theory is consistent in 10 spacetime dimensions.
 2. there are five fully consistent string theories in $d = 10$.
Type IIA, Type IIB, Type I (open+closed string), $E_8 \times E_8$ Heterotic, $SO(32)$ Heterotic
 3. Type II theories can have D-branes that accommodate open strings. Type IIA (resp. IIB) has Dp -branes where p is even (resp. odd), and Type I does D9, D5, D1-branes.
 4. Furthermore, all the five apparently different string theories are different limits of the same underlying theory called M-theory
 5. In the low energy scale $\ll 1/\ell_s$, a theory is described by supergravity which combines supersymmetry and general relativity.
- How do we reduce the number of dimensions from 10 to 4? The answer is Kaluza-Klein mechanism/compactification. If we take the 10-dimensional space of the form $\mathbb{R}^{1,3} \times M$ where M is a 6-dimensional compact space and its size is much smaller than the resolution of the most powerful accelerator $\sim 10^{-17}$ cm, then a (9+1)-dimensional theory looks like (3+1)-dimensional. In fact, in order for a theory to be consistent, M has to be a Calabi-Yau manifold (which is endowed with a Ricci-flat metric). More importantly, the extra dimensions give “room” to derive the complexity of the real world from a simple setting.

1.4 Convention

The world-sheet coordinate indices are denoted by the alphabet letter a, b, \dots , and the target-space coordinate indices are denoted by the Greek letter μ, ν, \dots

The world-sheet metric is denoted by h_{ab} .

spacetime

In anticipation of string theory, we consider D -dimensional Minkowski space $\mathbb{R}^{1,D-1}$. Throughout these notes, we work with signature

$$\eta_{\mu\nu} = \text{diag}(-1, +1, +1, \dots, +1)$$

World-sheet

We use σ^a for a local coordinate of the world-sheet and we denote the metric by h_{ab} in this local coordinate so that the measure is $\sqrt{h}d^2\sigma$ for a curved world-sheet.

We often consider the case that the world-sheet is topologically a cylinder or a plane. The Minkowski cylindrical coordinates are denoted by (τ, σ) and the Euclidean cylindrical coordinates (t, σ) where the Wick rotation is performed as $\tau = -it$. In the Euclidean coordinate, we introduce the holomorphic complex coordinate $w = it + \sigma$ for the cylinder. The conformal map brings the cylindrical coordinate to the plane coordinate by $z = e^{-iw}$.

- right-moving or holomorphic coordinate: $\sigma^- = \tau - \sigma$ (Minkowski cylinder), $w = it + \sigma$ (Euclidean cylinder), z (Euclidean plane),
- left-moving or anti-holomorphic coordinate: $\sigma^+ = \tau + \sigma$ (Minkowski cylinder), $\bar{w} = -it + \sigma$ (Euclidean cylinder), \bar{z} (Euclidean plane),

In the plane coordinate, the holomorphic coordinates are given by

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

The metric is given by

$$d^2z = dz d\bar{z} = 2d\sigma^1 d\sigma^2,$$

where the factor 2 comes from the Jacobian. One can also write

$$h_{z\bar{z}} = h_{\bar{z}z} = \frac{1}{2}, \quad h_{zz} = h_{\bar{z}\bar{z}} = 0, \quad h^{z\bar{z}} = h^{\bar{z}z} = 2, \quad h^{zz} = h^{\bar{z}\bar{z}} = 0.$$

Parameters and constants

- String theory has only one parameter α' , and string length and tension are written as

$$\ell_s = \sqrt{\alpha'}, \quad T = \frac{1}{2\pi\alpha'} \tag{1.1}$$

- String coupling is given by an expectation value $g_s = e^\Phi$ of the Dilaton field
- Planck length

$$\ell_P = \sqrt{\frac{\hbar G}{c^3}}, \quad \ell_P^3 = g_s \ell_s^3 \tag{1.2}$$

- Coupling constants of $D = 11$ and $D = 10$ supergravities

$$\frac{1}{2\kappa_{11}^2} = \frac{2\pi}{(2\pi\ell_p)^9}, \quad \frac{1}{2\kappa_{10}^2} = \frac{2\pi}{(2\pi\ell_s)^8} \tag{1.3}$$

Acknowledgments and disclaimer

We are grateful to the students at Fudan University who provided valuable comments on this note during the lecture series in 2017, 2021 and 2024. Special thanks go to Jiaqi Guo and Jack Yang for their assistance with creating the figures. We also thank Y. Tachikawa for bringing the historical account in [CCCDV12] to our attention.

Given the vast literature on string theory, it is not possible to include all relevant citations in this note. Instead, we encourage readers to delve into the original sources for a more comprehensive understanding of the subject. As mentioned above, there are already many good books on string theory, and this note covers only a small portion of the vast field of string theory. Nevertheless, we would be delighted if it serves as a concise introduction for students.

2 Bosonic string theory

The first few sections are devoted to the study of bosonic string theory. Bosonic string theory is a framework in which the point-like particles of particle physics are replaced by one-dimensional objects called “strings”. These strings can vibrate at different frequencies, and each frequency corresponds to a different particle.

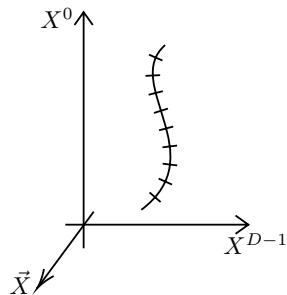
Strings can either be loops (closed strings) or have endpoints (open strings). Closed strings give rise to gravity (graviton) while open strings can give rise to gauge fields and matter. This provides a unified description where all particles arise as different states of the same fundamental entity. One of the notable outcomes of bosonic string theory is that it requires 26 spacetime dimensions for the theory to be consistent.

One of the problematic features of bosonic string theory is the existence of tachyons, particles with imaginary mass. The presence of tachyons signals an instability in the theory. Another limitation is that bosonic string theory does not include fermions. Later, superstring theory will come to address these problems, which incorporates both bosons and fermions. (See §7.) Nonetheless, bosonic string theory plays a role in the construction of Heterotic string theory. (See §11.)

In this section, let us first quantize bosonic string theory by using the light-cone gauge. For bosonic string theory, there is a concise and wonderful lecture note [Ton09] by D. Tong.

2.1 String sigma-model action

Nambu-Goto action



To begin with, let us recall the case of a relativistic point particle

$$\gamma : \tau \rightarrow x^\mu(\tau) \in \mathbb{R}^{1,D-1}.$$

The action of a relativistic point particle with mass m is

$$S = m \int_a^b ds = m \int_a^b \left(-\eta_{\mu\nu} \frac{dX^\mu}{d\tau} \frac{dX^\nu}{d\tau} \right)^{\frac{1}{2}} d\tau.$$

In a similar fashion, when a string moves through spacetime $\mathbb{R}^{1,D-1}$, it traces out a two-dimensional surface known as the worldsheet, which is described by

$$\Sigma \ni (\tau, \sigma) \rightarrow X^\mu(\tau, \sigma) \in \mathbb{R}^{1,D-1}.$$

The analogous action, called the **Nambu-Goto action**, for the string is given by the area of a string world-sheet Σ swept by a string

$$S_{\text{NG}} = -T \int_\Sigma \text{Area} = T \int_\Sigma d^2\sigma \sqrt{-\det(h_{ab})}, \quad h_{ab} = \eta_{\mu\nu} \partial_a X^\mu \partial_b X^\nu \quad (2.1)$$

where $(\sigma^0, \sigma^1) = (\tau, \sigma)$, and $T := 1/(2\pi\alpha')$ is the **string tension**. The tension is usually defined as mass per unit length. However, we do not know how to quantize the Nambu-Goto action due to the square root. Instead, for quantization of string world-sheet, we reformulate the action as follows.

String sigma-model action

Alternatively, we remove the square root and consider the following action

$$S_\sigma = -\frac{1}{4\pi\alpha'} \int d^2\sigma \sqrt{-\det h} h^{ab} \partial_a X^\mu \partial_b X^\nu \eta_{\mu\nu} \quad (2.2)$$

which is called the **string sigma-model action**.¹ Classically, the string sigma-model action is equivalent to the Nambu-Goto action.

Symmetry of S_σ

1. D -dim spacetime Poincaré invariance

The action is certainly invariant under the Poincaré transformation of the target space $\mathbb{R}^{1,D-1}$

$$X^\mu(\sigma) \rightarrow \Lambda^\mu{}_\nu X^\nu(\sigma) + V^\mu \quad \Lambda \in O(1, D-1)$$

where Λ is a Lorentz transformation and V is the translation.

2. World-sheet diffeomorphism invariance

The string sigma-model is also invariant under a diffeomorphism $f : \tilde{\Sigma} \rightarrow \Sigma$ of world-sheet two-dimensional manifolds. Namely, the action stays the same for

¹This action is called the Polyakov action in [Pol98]. However, according to [BBS06], it was discovered by Brink, Di Vecchia and Howe and by Deser and Zumino several years before Polyakov skillfully used it for path-integral quantization of the string. Therefore, let us follow the notation of the progenitor, J.H. Schwarz.

(Σ, h) and $(\tilde{\Sigma}, \tilde{h} = f^*h)$. In physics, it is most convenient to use local coordinates to express this invariance. Namely, under a local expression of the diffeomorphism

$$\begin{aligned}\tilde{X}^\mu(\tilde{\sigma}) &= X^\mu(\sigma) \quad \text{with} \quad \tilde{\sigma} = \tilde{\sigma}(\sigma), \\ \tilde{h}^{ab} &= h^{cd} \frac{\partial \tilde{\sigma}^a}{\partial \sigma^c} \frac{\partial \tilde{\sigma}^b}{\partial \sigma^d},\end{aligned}$$

the action is invariant $S_\sigma(X^\mu, h_{ab}) = S_\sigma(\tilde{X}^\mu, \tilde{h}_{ab})$.

3. Weyl invariance of world-sheet metrics

A **Weyl transformation** $h \rightarrow e^{2\omega}h$ is a local scaling of the metric of a Riemannian manifold (Σ, h) where 2ω is an arbitrary function of Σ . The action is indeed invariant under a Weyl transformation $h_{ab} \rightarrow e^{2\omega(\sigma)}h_{ab}$ of the string world-sheet Σ .

In fact, $\text{Diff} \times \text{Weyl}$ are gauge symmetries of the theory.

Energy-momentum tensor

The energy-momentum tensor is defined by

$$\begin{aligned}T_{ab} &\equiv -\frac{4\pi}{\sqrt{-h}} \frac{\delta S}{\delta h^{ab}} \\ &= -\frac{1}{\alpha'} \left[\partial_a X^\mu \partial_b X_\mu - \frac{1}{2} h_{ab} \partial_c X^\mu \partial^c X_\mu \right]\end{aligned}\tag{2.3}$$

which is symmetric and subject to

$$\begin{aligned}\text{traceless :} \quad T^a{}_a &= 0 \\ \text{conservation :} \quad \nabla_a T^{ab} &= 0\end{aligned}\tag{2.4}$$

Note that the traceless condition indeed follows from the Weyl invariance (exercise), and the conservation can be shown by using the equation of motion for X^μ in the following.

$$\begin{aligned}h^{ab} \frac{\delta S}{\delta h^{ab}} &= 0 \rightarrow T^a{}_a = 0 \\ \frac{\delta S}{\delta X^\mu} &= 0 \rightarrow \Delta X_\mu = -\frac{1}{\sqrt{-h}} \partial_a (\sqrt{-h} h^{ab} \partial_b X^\mu) = 0.\end{aligned}\tag{2.5}$$

The equation of motion with respect to the metric will be discussed in (2.11).

Gauge fixing

The string sigma-model action S_σ is equipped with the symmetries above, and we need to fix it for quantization. Physics does not depend on a choice of gauge fixing, but if we choose a clever gauge fixing, our life becomes much easier. As in [GSW87, Pol98, BBS06], we choose the light-cone gauge in the following.

The metric h_{ab} is a 2×2 symmetric matrix and the world-sheet diffeomorphism invariance tells us that there is a coordinate σ^a such that the metric is a diagonal $h_{ab} = e^\omega \eta_{ab}$. Then, the Weyl transformation brings it to the 2d Minkowski metric η_{ab} . Therefore, one can use $h_{ab} = \eta_{ab}$. If we use the light-cone coordinates on the world-sheet

$$\sigma^\pm = \tau \pm \sigma,$$

then the metric is written as

$$ds^2 = -d\sigma^+ d\sigma^- .$$

However, there are residual transformations that leave the 2d Minkowski metric η_{ab} invariant, which is called the **conformal symmetry**:

$$\sigma^+ \rightarrow f(\sigma^+) , \quad \sigma^- \rightarrow g(\sigma^-)$$

with a Weyl transformation simultaneously. To fix the conformal symmetry, we introduce the spacetime light-cone coordinates,

$$X^\pm = \sqrt{\frac{1}{2}}(X^0 \pm X^{D-1}) .$$

Then we impose

$$X^+ = x^+ + \alpha' p^+ \tau , \quad (2.6)$$

which is called the **light-cone gauge**.

Mode Expansions

After the gauge fixing, the equations of motion simply read

$$\partial_+ \partial_- X^\mu = 0 \quad \text{where} \quad \partial_\pm = \frac{1}{2} (\partial_\tau \pm \partial_\sigma) .$$

The most general solution is a factorization of left- and right-mover

$$X^\mu(\sigma, \tau) = X_L^\mu(\sigma^+) + X_R^\mu(\sigma^-) \quad (2.7)$$

for arbitrary functions X_L^μ and X_R^μ . For a closed string, we impose a periodic condition as

$$X^\mu(\sigma, \tau) = X^\mu(\sigma + 2\pi, \tau) . \quad (2.8)$$

so that the left- and right-moving fields admit the Fourier expansions

$$X_L^\mu(\sigma^+) = \frac{1}{2}x^\mu + \frac{\alpha'}{2}p^\mu\sigma^+ + i\sqrt{\frac{\alpha'}{2}} \sum_{n \neq 0} \frac{1}{n}\bar{\alpha}_n^\mu e^{-in\sigma^+} , \quad (2.9)$$

$$X_R^\mu(\sigma^-) = \frac{1}{2}x^\mu + \frac{\alpha'}{2}p^\mu\sigma^- + i\sqrt{\frac{\alpha'}{2}} \sum_{n \neq 0} \frac{1}{n}\alpha_n^\mu e^{-in\sigma^-} , \quad (2.10)$$

where the reality of X^μ requires that the coefficients of the Fourier modes obey

$$\alpha_n^\mu = (\alpha_{-n}^\mu)^* , \quad \bar{\alpha}_n^\mu = (\bar{\alpha}_{-n}^\mu)^* .$$

This mode expansion will be very important when we come to quantum theory. If we treat the world-sheet metric dynamically, we impose the equation of motion with respect to the worldsheet metric

$$T_{--} = -\frac{1}{\alpha'}\partial_- X^\mu \partial_- X_\mu = 0 , \quad T_{++} = -\frac{1}{\alpha'}\partial_+ X^\mu \partial_+ X_\mu = 0 . \quad (2.11)$$

We will see the implication of these constraints in quantum theory.

2.2 Quantizations

The momentum conjugate to X^μ is defined in this gauge

$$\Pi_\mu = \frac{\delta S}{\delta (\partial_\tau X^\mu)} = \frac{1}{2\pi\alpha'} \partial_\tau X_\mu .$$

The canonical quantization promotes X^μ and Π_μ to operators that is subject to

$$\begin{aligned} [X^\mu(\sigma, \tau), \Pi_\nu(\sigma', \tau)] &= i\delta(\sigma - \sigma') \delta_\nu^\mu \\ [X^\mu(\sigma, \tau), X^\nu(\sigma', \tau)] &= [\Pi_\mu(\sigma, \tau), \Pi_\nu(\sigma', \tau)] = 0 . \end{aligned} \quad (2.12)$$

We translate these into commutation relations for the Fourier modes x^μ , p^μ , α_n^μ and $\bar{\alpha}_n^\mu$. Using the mode expansion, we find (exercise)

$$[x^\mu, p_\nu] = i\delta_\nu^\mu \quad \text{and} \quad [\alpha_n^\mu, \alpha_m^\nu] = [\bar{\alpha}_n^\mu, \bar{\alpha}_m^\nu] = n\eta^{\mu\nu}\delta_{n+m,0} , \quad (2.13)$$

Therefore, like harmonic oscillators, the creation operators are $\alpha_{-n}^\mu, \bar{\alpha}_{-n}^\mu$ and the annihilation operators are $\alpha_n^\mu, \bar{\alpha}_n^\mu$ for $n \in \mathbb{N}$ so that the Hilbert space is spanned by

$$\alpha_{-n_1}^{\mu_1} \cdots \alpha_{-n_k}^{\mu_k} \bar{\alpha}_{-n_1}^{\mu_1} \cdots \bar{\alpha}_{-n_k}^{\mu_k} |0; k\rangle \quad \text{where} \quad p^\mu |0; k\rangle = k^\mu |0; k\rangle . \quad (2.14)$$

Let us consider the implication of the constraints (2.11). Using the mode expansions (2.9), the energy-momentum tensor can be expressed as

$$T_{--} = -\sum_n L_n^X e^{-in\sigma^-} \quad T_{++} = -\sum_n \bar{L}_n^X e^{-in\sigma^+} ,$$

where

$$L_m^X = \frac{1}{2} \sum_{n \in \mathbb{Z}} \eta_{\mu\nu} : \alpha_{m-n}^\mu \alpha_n^\nu : \quad \bar{L}_m^X = \frac{1}{2} \sum_{n \in \mathbb{Z}} \eta_{\mu\nu} : \bar{\alpha}_{m-n}^\mu \bar{\alpha}_n^\nu : \quad (2.15)$$

with $\alpha_0^\mu = \bar{\alpha}_0^\mu = \sqrt{\frac{\alpha'}{2}} p^\mu$. We will learn that L_m are generators of **Virasoro algebra**. According to the normal-ordering : : prescription, the lowering operators always appear to the right of the raising operators. However, it is easy to see that the normal ordering matters only in L_0^X and \bar{L}_0^X . Thus, in quantum theory, the constraint (2.11) on the energy-momentum tensor can be interpreted by

$$(L_n^X + A\delta_{n,0}) |\text{phys}\rangle = 0 \quad (\bar{L}_n^X + \bar{A}\delta_{n,0}) |\text{phys}\rangle = 0 \quad \text{for } n \geq 0$$

where A, \bar{A} are normal ordering constants so that

$$\langle \text{phys}' | L_n^X | \text{phys} \rangle = 0 \quad \langle \text{phys}' | \bar{L}_n^X | \text{phys} \rangle = 0 \quad \text{for } n \in \mathbb{Z} .$$

Another way to think of this constraint is as follows. A physical amplitude will not depend on the choice of gauge $h_{ab}(\sigma) + \delta h_{ab}$, i.e.

$$\delta \langle f | i \rangle = -\frac{1}{4\pi} \int d^2\sigma h(\sigma)^{1/2} \delta h_{ab}(\sigma) \left\langle f \left| T^{ab}(\sigma) \right| i \right\rangle = 0 .$$

Light-cone quantization

To see the effect of normal ordering in L_0 , let us consider the meaning of the light-cone gauge (2.6) in quantum theory. It is easy to see that (2.6) implies

$$\alpha_n^+ = 0 \quad \text{for } n \neq 0.$$

Then, the equation of motion (2.11) with respect to the worldsheet metric

$$2\partial_+ X^- \partial_+ X^+ = \sum_{i=1}^{D-2} \partial_+ X^i \partial_+ X^i$$

tells us

$$\alpha_n^- = \sqrt{\frac{1}{2\alpha'}} \frac{1}{p^+} \sum_{m=-\infty}^{\infty} \sum_{i=1}^{D-2} \alpha_{n-m}^i \alpha_m^i.$$

In particular, for $n = 0$, we have

$$M^2 = 2p^+ p^- - \sum_{i=1}^{D-2} p^i p^i = \frac{4}{\alpha'} \left(\sum_{n>0} \sum_{i=1}^{D-2} \alpha_{-n}^i \alpha_n^i + \frac{D-2}{2} \sum_{n>0} n \right).$$

The final term clearly diverges. Fortunately, we have nice regularization of this divergence

$$\begin{aligned} \sum_{n=1}^{\infty} n &\longrightarrow \sum_{n=1}^{\infty} n e^{-\epsilon n} = -\frac{\partial}{\partial \epsilon} \sum_{n=1}^{\infty} e^{-\epsilon n} \\ &= -\frac{\partial}{\partial \epsilon} (1 - e^{-\epsilon})^{-1} \\ &= \frac{1}{\epsilon^2} - \frac{1}{12} + \mathcal{O}(\epsilon). \end{aligned} \tag{2.16}$$

The first term diverges as $\epsilon \rightarrow 0$ so that we renormalize this term away. Consequently, we obtain

$$\sum_{n=1}^{\infty} n = -\frac{1}{12}. \tag{2.17}$$

Hence, we obtain

$$M^2 = \frac{4}{\alpha'} \left(N - \frac{D-2}{24} \right) = \frac{4}{\alpha'} \left(\bar{N} - \frac{D-2}{24} \right), \tag{2.18}$$

where the number operators are defined by

$$N = \sum_{i=1}^{D-2} \sum_{n>0} \alpha_{-n}^i \alpha_n^i, \quad \bar{N} = \sum_{i=1}^{D-2} \sum_{n>0} \bar{\alpha}_{-n}^i \bar{\alpha}_n^i. \tag{2.19}$$

It is easy to see from (2.18)

$$N = \bar{N} \tag{2.20}$$

which is called the **level-matching condition**. Moreover, if $D > 2$, we have the state $|0;k\rangle$ with negative [mass]²

$$M^2 = -\frac{4}{\alpha'} \frac{D-2}{24} \tag{2.21}$$

for $N = 0 = \bar{N}$ which is called the **tachyon**.

The First Excited States

Let us look at the first excited states. The level-matching condition (2.20) requires us to act both right- α_{-1}^j and left-moving $\bar{\alpha}_{-1}^i$ creation operator on $|0; k\rangle$ simultaneously. Thus, there are $(D - 2)^2$ states at $N = 1 = \bar{N}$

$$\alpha_{-1}^i \bar{\alpha}_{-1}^j |0; k\rangle , \quad (2.22)$$

whose mass

$$M^2 = \frac{4}{\alpha'} \left(1 - \frac{D-2}{24} \right) .$$

It is easy to see that the state is under a representation of the little group $\text{SO}(D-2) \subset \text{SO}(1, D-1)$ and therefore it should be massless. Interestingly, this is only the case if the dimension of spacetime is

$$D = 26 . \quad (2.23)$$

This is the critical dimension of the bosonic string. In the following sections, we will see from many different viewpoints that bosonic string theory is consistent only in this critical dimension.

Then, the states (2.22) transform in the $\mathbf{24} \otimes \mathbf{24}$ representation of $\text{SO}(24)$, which decomposes into three irreducible representations:

$$\begin{aligned} \text{Graviton } G^{\mu\nu} & \quad \alpha_{-1}^\mu \bar{\alpha}_{-1}^\nu |0; k\rangle \quad (\text{symmetric in } \mu \text{ and } \nu, \text{ and traceless}) , \\ \text{B-field } B^{\mu\nu} & \quad \alpha_{-1}^\mu \bar{\alpha}_{-1}^\nu |0; k\rangle \quad (\text{anti-symmetric in } \mu \text{ and } \nu) , \\ \text{Dilaton } \Phi & \quad \alpha_{-1}^\mu \bar{\alpha}_{-1,\mu} |0; k\rangle \quad (\text{trace part}) . \end{aligned} \quad (2.24)$$

Each irreducible representation gives rise to a massless field in spacetime. As above, the traceless symmetric, anti-symmetric, and the trace part correspond to the graviton $G_{\mu\nu}$, (Kalb-Ramond) $B_{\mu\nu}$, and the dilaton Φ , respectively. Therefore, graviton arises naturally from the quantization of closed strings! These massless fields play a pivotal role throughout the lecture notes.

Open strings

So far we have studied closed strings. Let us briefly summarize the light-cone quantization of an open string and the derivations of the results are left as an exercise to the reader. For an open string, a world-sheet spatial coordinate spans $\sigma \in [0, \pi]$ where the boundaries of the world-sheet are at $\sigma = 0, \pi$, and the action is given by

$$S = \frac{1}{4\pi\alpha'} \int_{\sigma=0}^{\sigma=\pi} d\tau d\sigma (\partial_\tau X \cdot \partial_\tau X + \partial_\sigma X \cdot \partial_\sigma X) .$$

Like the close string, the mode expansion of $X^\mu = X_L^\mu(\sigma^+) + X_R^\mu(\sigma^-)$ is given by

$$\begin{aligned} X_L^\mu(\sigma^+) &= \frac{1}{2} x^\mu + \alpha' p^\mu \sigma^+ + i \sqrt{\frac{\alpha'}{2}} \sum_{n \neq 0} \frac{1}{n} \bar{\alpha}_n^\mu e^{-in\sigma^+} , \\ X_R^\mu(\sigma^-) &= \frac{1}{2} x^\mu + \alpha' p^\mu \sigma^- + i \sqrt{\frac{\alpha'}{2}} \sum_{n \neq 0} \frac{1}{n} \alpha_n^\mu e^{-in\sigma^-} . \end{aligned} \quad (2.25)$$

where the momentum of an open string is defined as $\alpha_0^\mu = \sqrt{2\alpha'} p^\mu$. Note that the second term differs from the closed string (2.9) by a factor of two since the spatial lengths are $l_{\text{open}} = \pi$ and $l_{\text{closed}} = 2\pi$.

The variation principle of the action leads to

$$\begin{aligned} 0 = \delta S &= \frac{1}{2\pi\alpha'} \int_{\sigma=0}^{\sigma=\pi} d\tau d\sigma (-\partial_\tau X \cdot \partial_\tau \delta X + \partial_\sigma X \cdot \partial_\sigma \delta X) \\ &= \frac{1}{2\pi\alpha'} \int_{\sigma=0}^{\sigma=\pi} d\tau d\sigma [-\delta X \cdot (-\partial_\tau \partial_\tau X + \partial_\sigma \partial_\sigma X) - \partial_\tau (\partial_\tau X \cdot \delta X) + \partial_\sigma (\partial_\sigma X \cdot \delta X)] \\ &= \frac{1}{2\pi\alpha'} \int d\tau [\partial_\sigma X \cdot \delta X]_{\sigma=0}^{\sigma=\pi}, \end{aligned} \quad (2.26)$$

where we use the equation of motion and $\delta X(t = \pm\infty) = 0$ at the last equality. This compels us to impose one of the following boundary conditions:

- Neumann boundary condition: $\partial_\sigma X^\mu = 0$ at $\sigma = 0, \pi$
- Dirichlet boundary condition: $X^\mu = c^\mu$ (constant) at $\sigma = 0, \pi$

which imposes the following conditions on the modes

- Neumann boundary condition: $\alpha_n^\mu = \bar{\alpha}_n^\mu$.
- Dirichlet boundary condition: $x^\mu = c^\mu$, $p^\mu = 0$, $\alpha_n^\mu = -\bar{\alpha}_n^\mu$.

Moreover, the light-cone gauge quantization for (2.25) leads to the open string mass spectrum

$$M^2 = 2p^+ p^- - \sum_{i=1}^{D-2} p^i p^i = \frac{1}{\alpha'} \left(\sum_{n>0} \sum_{i=1}^{D-2} \alpha_{-n}^i \alpha_n^i + \frac{D-2}{2} \sum_{n>0} n \right), \quad (2.27)$$

where there is the difference in the normalization p^2 between closed and open string: $4p_{\text{open}}^2 = p_{\text{closed}}^2$. This results in a difference by a factor of four from (2.18). Again, in $D = 26$, the open string mass spectrum becomes

$$M^2 = \frac{1}{\alpha'} (N - 1). \quad (2.28)$$

so that there is the tachyon vacuum at the level $N = 0$. The massless states

$$\alpha_{-1}^i |0; k\rangle \quad i = 1, \dots, D-2 \quad (2.29)$$

at the level $N = 1$ corresponds to vector gauge bosons.

2.3 Path-integral formulations

We have obtained the bosonic string spectra, and to illustrate the interaction among string states, we will introduce **path integral**. (For more detail, see [Pol98, Appendix.A].) Figure 2 depicts an extension from quantum field theoretic (QFT) interaction to string interaction. Each end of the cylinders in the Figure corresponds to an initial/final state of a string. The action (2.2) of the string sigma model is invariant under the world-sheet diffeomorphisms, and it is a conformal field theory as a result. As we will see in the next section §3, one can make use of the **state-operator correspondence** in a 2d CFT. Hence, instead of using the Hamiltonian formalism of string states, we will use vertex operators \hat{V}_i in the path integral formalism:

$$A_n = \sum_g \int (\mathcal{D}h_{ab})_{g,n} \int \mathcal{D}X^\mu e^{-S_\sigma[X^\mu, h_{ab}]} \hat{V}_1 \cdots \hat{V}_n,$$

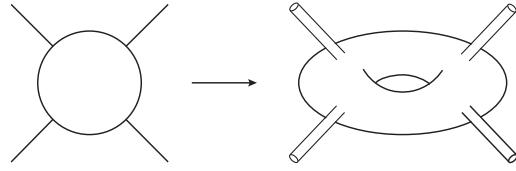


Figure 2: A Feynman diagram of QFT (left) and a string interaction diagram (right). Each end of the cylinders emanating from the torus (or a Riemann surface in general) in the string interaction diagram corresponds to an initial/final state.

where n is the number of in-coming/out-going strings, and g is the number of holes(genus) of the Riemann surface (genus is 1 in Figure 3). Here we integrate over all the metric h_{ab} and field X^μ configuration up to gauge transformations, and we also sum over all the topology (genus g) of world-sheets.

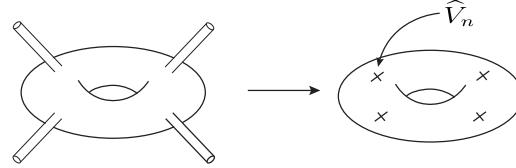


Figure 3: String interaction in state expression (left) and operator expression (right).

In the path integral method, expectation values are schematically expressed in the following form

$$\langle \mathcal{O}[X] \rangle = \int \mathcal{D}X e^{iS[X]} \mathcal{O}[X], \quad S = \frac{1}{4\pi\alpha'} \int d\sigma d\tau \left\{ (\partial_\tau X)^2 - (\partial_\sigma X)^2 \right\},$$

where \mathcal{O} is a gauge-invariant operator.

We usually **Wick rotate** ($\tau \rightarrow -it$) the theory so that it converges.

$$\langle \mathcal{O}[X] \rangle = \int \mathcal{D}X e^{-S_E[X]} \mathcal{O}[X], \quad S_E = \frac{1}{4\pi\alpha'} \int d\sigma dt \left\{ (\partial_t X)^2 + (\partial_\sigma X)^2 \right\}.$$

The subscript E will be omitted hereafter, and we will always work in the world-sheet Euclidean signature.

3 Two-dimensional conformal field theory

The dynamics of the string's worldsheet are described by a two-dimensional conformal field theory (2D CFT). Within this framework, “vertex operators” represent the various vibrational states of the string, emphasizing the significance of 2D CFT in string-theory research. A notable characteristic of 2D CFTs is their infinite-dimensional symmetry [BPZ84], which makes the story remarkably profound.

In this section, we will learn the operator analysis in 2D CFT, covering topics such as the Operator Product Expansion (OPE) and the Ward-Takahashi identity. Moreover, we will see Virasoro algebra, which is associated to 2d conformal symmetry.

However, it is important to acknowledge that what we are uncovering here only scratches the surface, and this subject could easily fill an entire semester. For readers intrigued by this fertile topic, we recommend consulting established references [[Gin88](#), [FMS12](#), [BP09](#)].

3.1 Conformal transformations

A CFT in any dimension is a quantum field theory invariant under conformal maps at quantum level. A conformal map between two Riemannian manifolds (Σ, h) and $(\tilde{\Sigma}, h')$ is a diffeomorphism $f : \Sigma \rightarrow \tilde{\Sigma}$ such that the pull-back metric f^*h' and the original metric h are related by an arbitrary function $e^{2\omega}$ of Σ

$$f^*h' = e^{2\omega}h.$$

In terms of local coordinates (σ, h_{ab}) and $(\tilde{\sigma}, h'_{ab})$, a **conformal transformation** $\sigma^a \rightarrow \tilde{\sigma}^a(\sigma)$ relates the metrics by

$$\frac{\partial \tilde{\sigma}^c}{\partial \sigma^a} \frac{\partial \tilde{\sigma}^d}{\partial \sigma^b} h'_{cd}(\tilde{\sigma}(\sigma)) = e^{2\omega(\sigma)} h_{ab}(\sigma). \quad (3.1)$$

Roughly speaking, a conformal transformation is a coordinate transformation that leaves the metric invariant up to scale and thus preserves angles. This means that the theory behaves the same at all length scales.

A conformal transformation is a diffeomorphism under which the pull-back metric is related to the original metric by a Weyl transformation. Since the string sigma model [\(2.2\)](#) is invariant under diffeomorphisms and Weyl transformations of the world-sheet, it is conformal-invariant. Hence, we will study conformal field theory of the string world-sheet.

Note that a conformal transformation is a diffeomorphism between two Riemannian manifolds, and a Weyl transformation changes the metric keeping a manifold fixed. (Keep in mind the difference!) Therefore, if the metrics (Σ, h) and $(\Sigma, e^{2\omega}h)$ are related by a Weyl transformation, then the identity map of Σ is a conformal transformation. [²](#)

2d flat space

In complex coordinates, the 2d flat metric can be written as $ds^2 = dz d\bar{z}$. Under a holomorphic map,

$$z \rightarrow f(z),$$

²Let us clarify two terminologies at this moment. A conformal transformation is a diffeomorphism under which the metric is scaled but the infinitesimal distance ds^2 is fixed. While in the previous section, the conformal symmetry appears as a residual symmetry after gauge fixing. This is a conformal transformation along with a Weyl transformation that fixes the metric but scales the infinitesimal distance ds^2 . Now the Weyl transformation is only used to undo the scaling factor due to the conformal transformation, and thus does not introduce extra degrees of freedom. These two concepts are just different ways that describe the same symmetry of our theory.

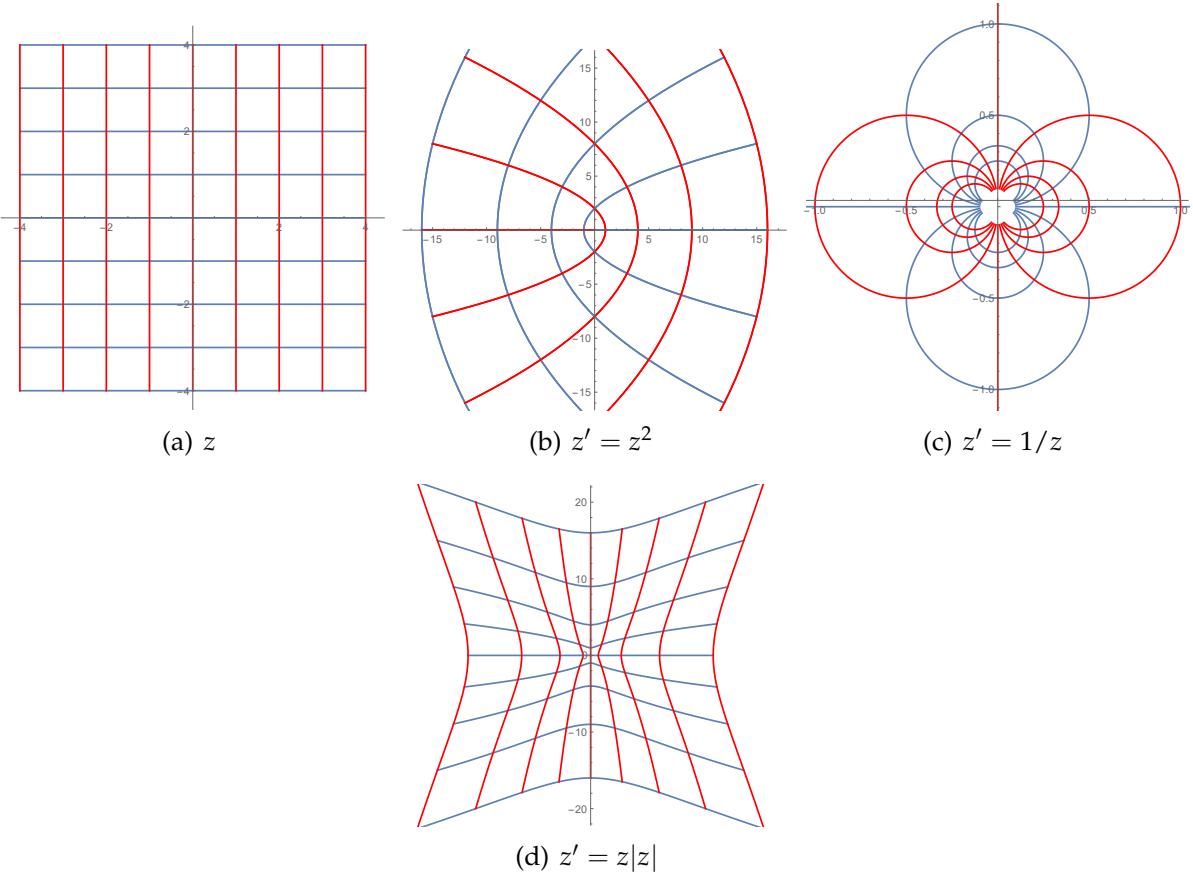


Figure 4: Coordinate transformation: The transformation from square lattice (a), onto the lattice in (b) and (c) is conformal, while the transformation onto the lattice in (d) is not.

the metric is transformed as

$$ds^2 = dz d\bar{z} \quad \rightarrow \quad ds^2 = \frac{\partial z}{\partial f} \frac{\partial \bar{z}}{\partial \bar{f}} df d\bar{f} .$$

Therefore, all the holomorphic maps are conformal transformations where $\left| \frac{\partial z}{\partial f} \right|^2$ is a conformal factor (corresponding to $e^{2\omega(\sigma)}$ in (3.1)). Moreover, it is easy to show that all 2d conformal transformations are indeed holomorphic functions (Exercise). This set is infinite-dimensional, corresponding to the coefficients of the Laurent series of holomorphic functions. Due to the infinite dimensionality, the conformal symmetry becomes so powerful in two dimensions.

3.2 State-operator correspondence

As we have seen, the action (2.2) of the string sigma model is equipped with the diffeomorphism-invariance and the Weyl-invariance. This means that it is invariant under conformal transformations. We can use this freedom to set the world-sheet to be \mathbb{R}^2

(recall that the topology of a propagating string is a cylinder).

$$\begin{aligned}\mathbb{R}^2 : \quad ds^2 &= d\sigma^1 d\sigma^1 + d\sigma^2 d\sigma^2 = dr^2 + r^2 d\theta^2 \\ &= r^2 [d(\log r)^2 + d\theta^2] , \\ \text{Cylinder} : \quad ds^2 &= dt^2 + d\sigma^2 ,\end{aligned}$$

where the overall factor r^2 in \mathbb{R}^2 is identified with the conformal factor $e^{2\omega(\sigma)}$. So we can identify

$$t = \log r , \quad \sigma = -\theta ,$$

and the reason for the minus sign will be clear later. This identification tells us that there is a correspondence between

- CFT on a cylinder and
- CFT on $\mathbb{C}^\times = \mathbb{C} \setminus \{0\}$,

provided that we choose the boundary conditions to be the same. In particular, the initial **state** of the string ($t = -\infty$) corresponds to a **local operator** inserted at the origin, which is called vertex operator (Figure 5). This is the state-operator correspondence in a CFT,

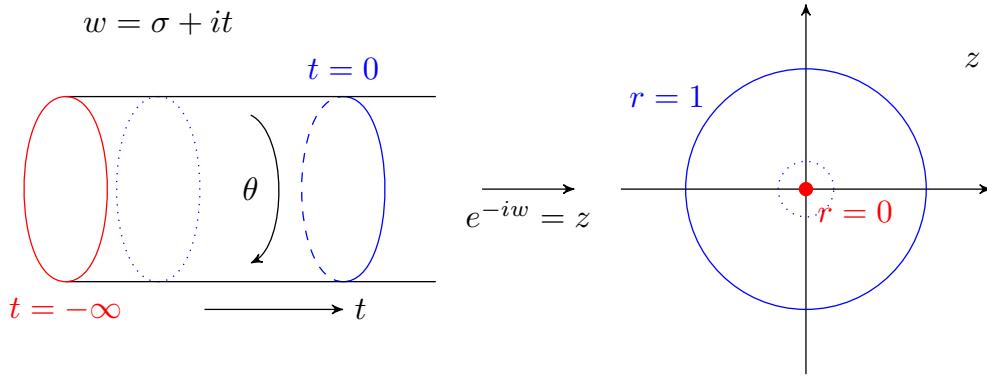


Figure 5: State-operator correspondence. The initial state (red circle) is mapped to a local operator (red dot) at the origin.

which plays a crucial role. We can express string states by the vertex operators (Figure 3).

Now the reader may wonder if there is any way to express commutation relations in terms of operators so that we can do parallel procedures of the canonical quantization in terms of operators.

3.3 Operator product expansions in free scalar theory

We study massless free scalar fields as the easiest example. Here we adopt the Euclidean metric $\delta_{ab} = \text{diag}(+, +)$ though the procedure is parallel for the Lorentz metric. We write the action in complex coordinates

$$S_X = \frac{1}{2\pi\alpha'} \int d^2z \partial X^\mu \bar{\partial} X_\mu . \quad (3.2)$$

Then, the equation of motion is given by $\partial\bar{\partial}X^\mu = 0$, which implies that X is holomorphically factorized

$$\begin{aligned} X^\mu(z, \bar{z}) &= X^\mu(z) + \bar{X}^\mu(\bar{z}), \\ X^\mu(z) &= \frac{1}{2}x^\mu - i\frac{\alpha'}{2}p^\mu \log z + i\sqrt{\frac{\alpha'}{2}} \sum_{n \neq 0} \frac{1}{n} \frac{\alpha_n^\mu}{z^n}, \\ \bar{X}^\mu(\bar{z}) &= \frac{1}{2}x^\mu - i\frac{\alpha'}{2}p^\mu \log \bar{z} + i\sqrt{\frac{\alpha'}{2}} \sum_{n \neq 0} \frac{1}{n} \frac{\bar{\alpha}_n^\mu}{\bar{z}^n}, \end{aligned} \quad (3.3)$$

where

$$z = re^{i\theta} = e^{\log r + i\theta} = e^{i\tau - i\sigma} = e^{i\sigma^-}, \quad \bar{z} = e^{i\sigma^+}.$$

This is the Euclidean version of (2.9). The (anti-)holomorphic part corresponds to the right (left) movers.

Let us consider a quantum version of the equation of motion by using a path integral formulation. Like a normal integral, we assume that the “total derivative” vanishes in the path integral. For example, in the massless free scalar case, we have

$$0 = \int \mathcal{D}X \frac{\delta}{\delta X} e^{-S[X]} = \int \mathcal{D}X e^{-S[X]} \frac{1}{\pi\alpha'} \partial\bar{\partial}X = \frac{1}{\pi\alpha'} \langle \partial\bar{\partial}X \rangle,$$

which is Ehrenfest’s theorem. The equation of motion is satisfied as an operator equation if there is no other operator nearby.

Furthermore, the same procedure can be done with operator insertion:

$$0 = \int \mathcal{D}X \frac{\delta}{\delta X_\mu(z, \bar{z})} \left(e^{-S[X]} X^\nu(w, \bar{w}) \right) = \left\langle \frac{1}{\pi\alpha'} \partial_z \partial_{\bar{z}} X^\mu(z, \bar{z}) X^\nu(w, \bar{w}) + \eta^{\mu\nu} \delta^2(z - w) \right\rangle,$$

yielding

$$\left\langle \partial\bar{\partial}X^\mu(z, \bar{z}) X^\nu(w, \bar{w}) \right\rangle = -\pi\alpha' \eta^{\mu\nu} \delta^2(z - w). \quad (3.4)$$

Now we can use the **Stokes’ theorem**

$$\int_D d^2z \left(\partial\bar{V} + \bar{\partial}V \right) = \frac{1}{i} \oint_{\partial D} (V dz - \bar{V} d\bar{z}),$$

where D is an arbitrary domain (typically a disk) and ∂D is its boundary. Using Stokes’ theorem, we obtain

$$\partial\bar{\partial} \log |z|^2 = 2\pi\delta^2(z). \quad (3.5)$$

This allows us to rewrite (3.4) into

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

where we introduced a normal ordering $:O:$. This is an operator equation, and we simply omit $\langle \rangle$ symbols. This expression describes the singular behavior when $X^\mu(z, \bar{z})$ approaches to $X^\nu(w, \bar{w})$.

We also write it as follows:

$$:X^\mu(z, \bar{z})X^\nu(w, \bar{w}): = X^\mu(z, \bar{z})X^\nu(w, \bar{w}) - \eta^{\mu\nu}G(z, w) ,$$

where $G(z, w) = -\frac{\alpha'}{2} \log |z - w|^2$. Notice that

$$\partial\bar{\partial} :X(z, \bar{z})X(w, \bar{w}): = 0 . \quad (3.6)$$

As we will see later, the divergent term is important and meaningful. This is the reason why it is convenient to introduce the normal ordering so that we can separate divergent terms from the other.

The normal ordering of an arbitrary functional of operators in the free scalar theory can be expressed as follows:

$$:f[X]: = \exp \left[-\frac{1}{2} \int d^2z d^2w G(z, w) \frac{\delta}{\delta X^\mu(z, \bar{z})} \frac{\delta}{\delta X_\mu(w, \bar{w})} \right] f[X] .$$

For example,

$$\begin{aligned} :X_1 X_2 X_3 X_4 X_5: &= X_1 X_2 X_3 X_4 X_5 - (G_{12} X_3 X_4 X_5 + \dots)_{10 \text{ terms}} \\ &\quad + (G_{12} G_{34} X_5 + \dots)_{15 \text{ terms}} , \end{aligned}$$

where $X_i = X(z_i)$ and $G_{ij} = G(z_i, z_j)$.³

Moreover, a product of normal ordered operators is given as follows:

$$:f[X]g[X]: = \exp \left[\int d^2z d^2w G(z, w) \left. \frac{\delta}{\delta X^\mu(z, \bar{z})} \right|_f \left. \frac{\delta}{\delta X_\mu(w, \bar{w})} \right|_g \right] :f[X]g[X]: . \quad (3.7)$$

For instance, we have

$$:\partial X(z, \bar{z})X(w, \bar{w}): = :\partial X(z, \bar{z})X(w, \bar{w}): + \partial_z G(z, w) = :\partial X(z, \bar{z})X(w, \bar{w}): - \frac{\alpha'}{2} \frac{1}{z - w} .$$

As another example, we can consider the OPE of vertex operators (See §3.7)

$$\begin{aligned} :e^{ik \cdot X(z, \bar{z})}e^{ik' \cdot X(w, \bar{w})}: &= \exp \left[(ik) \cdot (ik') \left(-\frac{\alpha'}{2} \log |z - w|^2 \right) \right] :e^{ik \cdot X(z, \bar{z}) + ik' \cdot X(w, \bar{w})}: \\ &= |z - w|^{\alpha' k \cdot k'} :e^{ik \cdot X(z, \bar{z}) + ik' \cdot X(w, \bar{w})}: . \end{aligned}$$

In a general field theory, a product of a pair of fields can be expanded by a single operator

$$\Phi^i(z)\Phi^j(w) = \sum_k C_k^{ij}(z - w)\Phi^k(w) .$$

This is called **operator product expansion (OPE)**, and it describes the behavior when the two operators approach each other.

³Here the normal ordering is used to cut off UV divergence in a particular renormalization scheme. It is not necessarily equivalent to the normal ordering defined by the creation and annihilation operators in the previous section.

Typically, for a massless free scalar field theory, we have

$$\begin{aligned} X^\mu(z, \bar{z})X^\nu(w, \bar{w}) &= -\frac{\alpha'}{2}\eta^{\mu\nu}\log|z-w|^2 + :X^\mu X^\nu(w, \bar{w}): \\ &+ \sum_{k=1}^{\infty} \frac{1}{k!} \left\{ (z-w)^k : \partial^k X^\mu X^\nu(w, \bar{w}): + (\bar{z}-\bar{w})^k : \bar{\partial}^k X^\mu X^\nu(w, \bar{w}): \right\}. \end{aligned} \quad (3.8)$$

The second and the third terms of RHS can be understood as the Taylor expansions of $:X(z)X(w):$ in terms of z . Note that mixing terms ($:\partial^m \bar{\partial}^n XX:$) vanish due to the “equation of motion” (3.6).

As we will see below, divergent terms in OPE are important. Thus, we write

$$X^\mu(z, \bar{z})X^\nu(w, \bar{w}) \sim -\frac{\alpha'}{2}\eta^{\mu\nu}\log|z-w|^2.$$

The symbol \sim means the regular terms are ignored.

3.4 Conformal Ward-Takahashi identity

When an action is invariant under a certain transformation δ (namely, $\delta S = 0$) we say the theory has a (classical) symmetry. Furthermore, the measure of the path integral is also invariant under the transformation, the theory has the symmetry at quantum level. On the other hand, if the measure is not invariant, then the symmetry is anomalous. It is non-trivial to see if the theory has an anomaly or not.

If the theory has a symmetry, Noether’s theorem states that there is the corresponding conserved current j^a . The space integral of its time component is the conserved charge $Q = \int_{\text{space}} j^0$, which generates the symmetry $\delta\phi = [Q, \phi]$. Let δ be a symmetry $\delta S = 0$ and assume it acts on a field as follows: $\delta\phi(z) = \epsilon(\dots)$, where ϵ is a small parameter. If we promote the parameter ϵ to be world-sheet coordinate dependent (i.e. $\hat{\delta}\phi = \epsilon(z, \bar{z})(\dots)$), then $\hat{\delta}S$ is no longer zero but has to take the following form:

$$\hat{\delta}S = \int \frac{d^2\sigma}{2\pi} \epsilon(x, y) \partial_a j^a = \int \frac{d^2z}{2\pi} \epsilon(z, \bar{z}) \left(\partial \bar{j} + \bar{\partial} j \right),$$

where we introduced $j = j_z$ and $\bar{j} = j_{\bar{z}}$. If the parameter is constant, $\delta S = (\text{total derivative}) = 0$ so that j^a is the **Noether current**. The conservation of the Noether current $\partial_a j^a = 0$ in the flat 2d implies that j (\bar{j}) is a holomorphic (anti-holomorphic) function of z .

Ward-Takahashi identity

Now let us study the transformation of a point operator $\mathcal{O}(w, \bar{w})$ under the symmetry δ . For this purpose, we set

$$\epsilon(z, \bar{z}) = \begin{cases} \epsilon & (\text{const.}) \quad \text{for } z \in D_w, \\ 0 & \quad \quad \quad \text{for } z \notin D_w, \end{cases}$$

where D_w is a disk containing w . Then, the variation of the path integral

$$0 = \int \mathcal{D}X \hat{\delta} \left[e^{-S} \mathcal{O}(w, \bar{w}) \right] = \int \mathcal{D}X e^{-S} \left[\delta \mathcal{O}(w, \bar{w}) - \hat{\delta}S \cdot \mathcal{O}(w, \bar{w}) \right].$$

Therefore, we have

$$\begin{aligned}\delta\mathcal{O}(w, \bar{w}) &= \int \frac{d^2z}{2\pi} \epsilon(z, \bar{z}) \left(\partial \bar{j}(\bar{z}) + \bar{\partial} j(z) \right) \mathcal{O}(w, \bar{w}) \\ &= \frac{\epsilon}{2\pi i} \oint_{\partial D_w} \left(dz j(z) - d\bar{z} \bar{j}(\bar{z}) \right) \mathcal{O}(w, \bar{w}),\end{aligned}$$

This is called the **Ward-Takahashi identity**.

Let us see an example of the free scalar field (3.2). It is easy to see that the action of the free scalar field is invariant under the transformation $X^\mu(z, \bar{z}) \rightarrow X^\mu(z, \bar{z}) + \epsilon^\mu$. If $\epsilon^\mu(z, \bar{z})$ has a function, then we have

$$\begin{aligned}\widehat{\delta}S &= \frac{-1}{2\pi\alpha'} \int d^2z \epsilon_\mu(z, \bar{z}) \left[\bar{\partial}(\partial X^\mu) + \partial(\bar{\partial} X^\mu) \right] \\ j^\mu &= -\frac{1}{\alpha'} \partial X^\mu, \quad \bar{j}^\mu = -\frac{1}{\alpha'} \bar{\partial} X^\mu.\end{aligned}\tag{3.9}$$

It is straightforward from (3.8) to check that

$$\delta X^\mu(w) = \epsilon_\nu \oint_{\partial D} \frac{dz}{2\pi i} j^\nu(z) X^\mu(w)$$

is consistent.

Now, let us consider the Ward-Takahashi identity for conformal symmetry. For an infinitesimal conformal transformation $\sigma^a \rightarrow \sigma^a + \epsilon^a(\sigma)$, the metric is transformed as

$$\delta_{ab} \rightarrow \delta_{ab} + \partial_a \epsilon_b + \partial_b \epsilon_a.$$

Since this is a conformal transformation, it is proportional to δ_{ab} so that we have

$$\partial_a \epsilon_b + \partial_b \epsilon_a = (\partial_\rho \epsilon^\rho) \delta_{ab}.\tag{3.10}$$

A solution to this equation is called a **conformal Killing vector**. The current for the conformal transformation can be written as

$$j_a = T_{ab} \epsilon^b,$$

where the straightforward calculation provides the energy-momentum tensor

$$T_{ab} = -2\pi \left[\frac{\partial L}{\partial(\partial^a \phi)} \partial_b \phi - \delta_{ab} L \right].\tag{3.11}$$

If we assume ϵ is constant, it is easy to see that the conservation of the current implies the conservation of the energy-momentum tensor:

$$\partial_a j^a = 0 \rightarrow \partial^a T_{ab} = 0.\tag{3.12}$$

For general $\epsilon^a(\sigma)$, the conservation of the current gives the traceless condition of T_{ab} :

$$0 = \partial^a j_a = \frac{1}{2} T_{ab} (\partial^a \epsilon^b + \partial^b \epsilon^a) = \frac{1}{2} T^a{}_a (\partial_\rho \epsilon^\rho) \rightarrow T^a{}_a = 0.\tag{3.13}$$

In the complex coordinate $z = x^1 + ix^2$, the traceless condition can be written as

$$T_{z\bar{z}} = T_{\bar{z}z} = 0$$

and the conservation of the energy-momentum tensor can be written as

$$\partial_{\bar{z}} T_{zz} = 0, \quad \partial_z T_{\bar{z}\bar{z}} = 0.$$

Thus, the non-vanishing components of the energy-momentum tensor factorize to a chiral and anti-chiral field,

$$T(z) := T_{zz} \quad \text{and} \quad \bar{T}(\bar{z}) := T_{\bar{z}\bar{z}}.$$

As a result, the Noether currents for conformal transformations $z \rightarrow z + \epsilon(z)$ and $\bar{z} \rightarrow \bar{z} + \bar{\epsilon}(\bar{z})$ are

$$j(z) = \epsilon(z) T(z), \quad \bar{j}(\bar{z}) = \bar{\epsilon}(\bar{z}) \bar{T}(\bar{z}).$$

The application to the Ward-Takahashi identity leads to **conformal Ward-Takahashi identity**

$$\delta_{\epsilon, \bar{\epsilon}} \mathcal{O}(w, \bar{w}) = \frac{1}{2\pi i} \oint_{C_w} dz \epsilon(z) T(z) \mathcal{O}(w, \bar{w}) + \frac{1}{2\pi i} \oint_{C_{\bar{w}}} d\bar{z} \bar{\epsilon}(\bar{z}) \bar{T}(\bar{z}) \mathcal{O}(w, \bar{w}), \quad (3.14)$$

where the contour integral is taken as a counter-clockwise circle both in z and in \bar{z} (thereby explaining the sign difference of the second term).

3.5 Primary fields

Let us first introduce some terminologies in 2d CFT. Fields depending only on z , i.e. $\phi(z)$, are called **chiral or holomorphic fields** and fields $\bar{\phi}(\bar{z})$ only depending on \bar{z} are called **anti-chiral or anti-holomorphic fields**. If a field ϕ transforms under the scaling transformation $z \rightarrow \lambda z$ as

$$\phi(z, \bar{z}) \rightarrow \phi'(\lambda z, \bar{\lambda} \bar{z}) = \lambda^{-h} \bar{\lambda}^{-\bar{h}} \phi(z, \bar{z}), \quad (3.15)$$

it is said to have **conformal dimension** (h, \bar{h}) . If a field transforms under a conformal transformation $z \rightarrow f(z)$ as

$$\phi(z, \bar{z}) \rightarrow \phi'(f(z), \bar{f}(\bar{z})) = \left(\frac{\partial f}{\partial z} \right)^{-h} \left(\frac{\partial \bar{f}}{\partial \bar{z}} \right)^{-\bar{h}} \phi(z, \bar{z}) \quad (3.16)$$

it is called a **primary field** of conformal dimensions (h, \bar{h}) . If (3.16) is true only for $f \in \mathrm{SL}(2, \mathbb{C})/\mathbb{Z}_2$, then it is called a **quasi-primary field**. Note that $\mathrm{SL}(2, \mathbb{C})/\mathbb{Z}_2$ group acts on the holomorphic coordinate as

$$z \mapsto \frac{az + b}{cz + d} \quad \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}(2, \mathbb{C})/\mathbb{Z}_2.$$

How do primary fields transform infinitesimally? Under the infinitesimal conformal transformation $z \rightarrow f(z) = z - \epsilon(z)$, we know that

$$\begin{aligned} \left(\frac{\partial f}{\partial z} \right)^{-h} &= 1 + h \partial_z \epsilon(z) + O(\epsilon^2), \\ \phi(z - \epsilon(z), \bar{z}) &= \phi(z) - \epsilon(z) \partial_z \phi(z, \bar{z}) + O(\epsilon^2). \end{aligned} \quad (3.17)$$

Hence, under an infinitesimal conformal transformation, the variation of a primary field is given by

$$\delta_\epsilon \phi(z, \bar{z}) = \left(h\partial_z \epsilon + \epsilon \partial_z + \bar{h}\bar{\partial}_{\bar{z}}\bar{\epsilon} + \bar{\epsilon}\bar{\partial}_{\bar{z}} \right) \phi(z, \bar{z}). \quad (3.18)$$

Consequently, using simple complex analysis

$$\begin{aligned} (\partial_w \epsilon(w)) \phi(w, \bar{w}) &= \frac{1}{2\pi i} \oint_{C_w} dz \frac{\epsilon(z) \phi(w, \bar{w})}{(z - w)^2} \\ \epsilon(w) (\partial_w \phi(w, \bar{w})) &= \frac{1}{2\pi i} \oint_{C_w} dz \frac{\epsilon(z) \partial_w \phi(w, \bar{w})}{z - w}, \end{aligned} \quad (3.19)$$

one can read off the OPE of a primary operator ϕ of conformal dimension (h, \bar{h}) with the energy-momentum tensor T (anti-chiral part \bar{T} can be obtained by complex conjugate)

$$T(z) \phi(w, \bar{w}) = h \frac{\phi(w, \bar{w})}{(z - w)^2} + \frac{\partial_w \phi(w, \bar{w})}{z - w} + \text{regular terms} \dots \quad (3.20)$$

In general, the OPE of an operator \mathcal{O} of conformal dimension (h, \bar{h}) with the energy-momentum tensor T and \bar{T} takes the form

$$T(z) \mathcal{O}(w, \bar{w}) = \dots + h \frac{\mathcal{O}(w, \bar{w})}{(z - w)^2} + \frac{\partial \mathcal{O}(w, \bar{w})}{z - w} + \dots$$

One of the main interests in a CFT is to calculate correlation functions of primary fields. Indeed, the conformal Ward-Takahashi identity can be applied to a correlation function of primary fields

$$\langle T(z) \phi_1(w_1, \bar{w}_1) \dots \phi_n(w_n, \bar{w}_n) \rangle = \sum_{i=1}^n \left(\frac{h_i}{(z - w_i)^2} + \frac{\partial_{w_i}}{z - w_i} \right) \langle \phi_1(w_1, \bar{w}_1) \dots \phi_n(w_n, \bar{w}_n) \rangle.$$

Moreover, conformal symmetry is so powerful that it determines the forms of two-point and three-point functions of primary fields (Exercise).

• 2-point function

For chiral primary operators ϕ_i with conformal dimension h_i ($i = 1, 2$), their 2-point function is of form

$$\langle \phi_1(z_1) \phi_2(z_2) \rangle = \delta_{h_1 h_2} \frac{d_{12}}{(z_1 - z_2)^{2h_1}} \quad (3.21)$$

If d_{12} is non-degenerate, the fields can be normalized such that $d_{12} = \delta_{12}$.

• 3-point function

A 3-point function is also completely fixed by conformal invariance up to the appearance of a **structure constant** C_{123} (exercise) C_{ijk} ,

$$\langle \phi_1(z_1) \phi_2(z_2) \phi_3(z_3) \rangle = \frac{C_{123}}{(z_1 - z_2)^{h_1+h_2-h_3} (z_2 - z_3)^{h_2+h_3-h_1} (z_3 - z_1)^{h_3+h_1-h_2}} \quad (3.22)$$

The structure constant depends on a CFT and, in general, it is not easy to determine it.

• Multi-point function

The computation of multi-point functions involves **conformal blocks** with the 3-point function. The details are explained in [FMS12, BP09].

Free scalar field

Now let us study conformal Ward-Takahashi identity in the simplest example, the free scalar field (3.2). Let us recall that the energy-momentum tensor in 2d free scalar theory is

$$T_{ab}^X = -\frac{1}{\alpha'} \left(\partial_a X^\mu \partial_b X_\mu - \frac{1}{2} \delta_{ab} (\partial_c X^\mu \partial^c X_\mu) \right), \quad (3.23)$$

As in (2.7), since the equation of motion for X^μ is $\partial_z \bar{\partial}_z X^\mu = 0$, the classical solution holomorphically factorizes as

$$X^\mu(z, \bar{z}) = X^\mu(z) + \bar{X}^\mu(\bar{z}).$$

In (3.9), we find the conserved holomorphic and anti-holomorphic U(1) current

$$j^\mu(z) := -\partial X^\mu(z)/\alpha', \quad \text{and} \quad \bar{j}^\mu(\bar{z}) := -\bar{\partial} \bar{X}^\mu(\bar{z})/\alpha'. \quad (3.24)$$

Moreover, the energy-momentum tensor becomes much simpler in complex coordinates. It is simple to check that $T_{z\bar{z}}^X = 0$ while

$$T^X(z) = -\frac{1}{\alpha'} \partial X^\mu(z) \partial X_\mu(z), \quad \bar{T}^X(\bar{z}) = -\frac{1}{\alpha'} \bar{\partial} \bar{X}^\mu(\bar{z}) \bar{\partial} \bar{X}_\mu(\bar{z}). \quad (3.25)$$

From the definition (3.16), one can see that $X(z, \bar{z})$ is a primary field of conformal dimension $(0, 0)$. However, since the conformal dimension is of $(0, 0)$, the two-point function does not exactly take the form (3.21). Indeed, the OPE XX tells us that the propagator takes the form

$$\langle X^\mu(z, \bar{z}) X^\nu(w, \bar{w}) \rangle = -\frac{\alpha'}{2} \eta^{\mu\nu} \log |z-w|^2.$$

Also, for each μ , the currents $j^\mu(z), \bar{j}^\mu(\bar{z})$ are primary fields of conformal dimension $(1, 0)$ and $(0, 1)$, respectively. Focusing only on the holomorphic part, an immediate check is their correlation function

$$\langle \partial X^\mu(z) \partial X^\nu(w) \rangle = -\frac{\alpha'}{2} \frac{\eta^{\mu\nu}}{(z-w)^2},$$

which takes the form (3.21). To convince ourselves completely, we need to compute the OPE with the energy-momentum tensor by Wick's theorem

$$\begin{aligned} T^X(z) \partial X^\mu(w) &= -\frac{1}{\alpha'} : \partial X^\nu(z) \partial X_\nu(z) : \partial X^\mu(w) \\ &= \frac{\partial X^\mu(w)}{(z-w)^2} + \frac{\partial^2 X^\mu(w)}{z-w} + \text{regular terms} \dots \end{aligned} \quad (3.26)$$

This is indeed the OPE for a primary operator of conformal dimension $h = 1$.

Finally, let us check to see the TT OPE of the energy-momentum tensors. This can be done by applying the Wick contractions, and the result is

$$\begin{aligned} T^X(z) T^X(w) &= \frac{1}{\alpha'^2} : \partial X^\mu(z) \partial X_\mu(z) : : \partial X^\nu(w) \partial X_\nu(w) : \\ &= \frac{D/2}{(z-w)^4} + \frac{2T^X(w)}{(z-w)^2} + \frac{\partial T^X(w)}{z-w} + \dots \end{aligned} \quad (3.27)$$

Therefore, the energy-momentum tensor is an operator of conformal dimension $(h, \bar{h}) = (2, 0)$. Because there is a higher singular term proportional to $(z-w)^{-4}$, the energy-momentum tensor is not a primary field. In fact, this is a general property of the TT OPE in all 2d CFTs.

3.6 Virasoro algebra

For the free scalar field, we have already seen that T has conformal dimension $(h, \tilde{h}) = (2, 0)$. This remains true in any CFT. The reason for this is simple: T_{ab} has dimension $\Delta = 2$ because we obtain the energy by integrating over space. It has spin $s = 2$ because it is a symmetric two-tensor. However, these two pieces of information are equivalent to the statement that T is an operator of conformal dimension $(2, 0)$. This means that the TT OPE takes the form,

$$T(z) T(w) = \dots + \frac{2T(w)}{(z-w)^2} + \frac{\partial T(w)}{z-w} + \dots$$

and a similar one for $\overline{T}\overline{T}$. What other terms could we have in this expansion? Since each term has dimension $\Delta = 4$, the unitarity indeed tells us that the singular part of the OPE takes

$$T(z) T(w) = \frac{c/2}{(z-w)^4} + \frac{2T(w)}{(z-w)^2} + \frac{\partial T(w)}{z-w} + \dots$$

From the OPE of the energy-momentum tensor, one can see its variation under an infinitesimal conformal transformation $z \rightarrow z - \epsilon(z)$

$$\begin{aligned} \delta_\epsilon T(w) &= \frac{1}{2\pi i} \oint_{C_w} dz \epsilon(z) T(z) T(w) \\ &= \epsilon(w) \partial T(w) + 2\epsilon'(w) T(w) + \frac{c}{12} \epsilon'''(w) \end{aligned} \quad (3.28)$$

One can verify by a straightforward computation that this is the infinitesimal version of the following transformation under finite transformation $z \rightarrow w(z)$:

$$T'(w) = \left(\frac{\partial w}{\partial z} \right)^{-2} \left[T(z) - \frac{c}{12} S(w, z) \right], \quad (3.29)$$

where the **Schwarzian derivative** S is defined as

$$S(w, z) := \frac{1}{(\partial_z w)^2} \left((\partial_z w)(\partial_z^3 w) - \frac{3}{2} (\partial_z^2 w)^2 \right). \quad (3.30)$$

Using the mapping from the plane to the cylinder $z = e^{-iw}$ ($w = \sigma + it$) as in Figure 5, the energy-momentum tensor is transformed as

$$T_{cyl}(w) = -z^2 T(z) + \frac{c}{24}. \quad (3.31)$$

The Laurent mode expansion of the energy-momentum tensor on the cylinder is therefore

$$T_{cyl}(w) = - \sum_{n \in \mathbb{Z}} \left(L_n - \frac{c}{24} \delta_{n,0} \right) e^{inw}. \quad (3.32)$$

Including the contribution from the anti-holomorphic sector, the Hamiltonian is defined by

$$H \equiv \int d\sigma T_{\tau\tau} = - \int d\sigma (T_{ww} + \overline{T}_{\overline{w}\overline{w}}) = L_0 + \overline{L}_0 - \frac{c}{12}.$$

This tells us that the ground state energy on the cylinder is

$$E = -\frac{c}{12}.$$

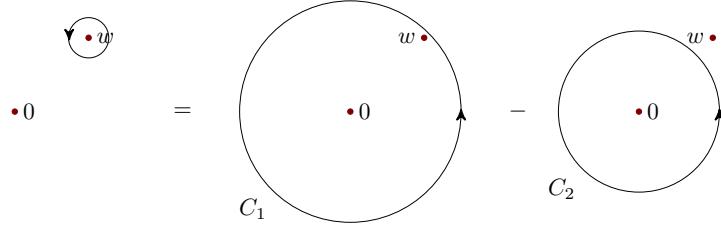


Figure 6:

This is indeed the (negative) Casimir energy on a cylinder. In the string sigma model, each coordinate X^μ of the target space gives rise to a free scalar theory, and it is easy to see from the computation (3.27) that each coordinate X^μ contributes to the central charge by $c = 1$. Thus, each target dimension yields the energy density $E = -1/12$ as we have seen in the quantization of bosonic string theory (2.17). Moreover, we have the central charge

$$c^X = D \quad (3.33)$$

for the string sigma model with the target space $\mathbb{R}^{1,D-1}$.

In the Euclidean flat space, the mode expansion of the energy-momentum tensor is expressed as

$$T(z) = \sum_{n \in \mathbb{Z}} \frac{L_n}{z^{n+2}},$$

where the shift by two in the exponent of z is due to the conformal dimension of the energy-momentum tensor. It is natural to find the commutation relation of the generators L_m . The commutator can be computed by

$$[L_m, L_n] = \left(\oint \frac{dz}{2\pi i} \oint \frac{dw}{2\pi i} - \oint \frac{dw}{2\pi i} \oint \frac{dz}{2\pi i} \right) z^{m+1} w^{n+1} R(T(z) T(w))$$

Here we impose the radial ordering

$$R(A(z)B(w)) := \begin{cases} A(z)B(w) & \text{for } |z| > |w| \\ B(w)A(z) & \text{for } |w| > |z| \end{cases}$$

according to which the contour is chosen as in Figure 6. Clever manipulation of the contour makes life easier

$$[L_m, L_n] = \oint \frac{dw}{2\pi i} \oint_w \frac{dz}{2\pi i} z^{m+1} w^{n+1} T(z) T(w) \quad (3.34)$$

$$= \oint \frac{dw}{2\pi i} \text{Res} \left[z^{m+1} w^{n+1} \left(\frac{c/2}{(z-w)^4} + \frac{2T(w)}{(z-w)^2} + \frac{\partial T(w)}{z-w} + \dots \right) \right] \quad (3.35)$$

A simple computation (Exercise) leads to the **Virasoro algebra**

$$[L_m, L_n] = (m-n)L_{m+n} + \frac{c}{12}m(m^2-1)\delta_{m+n,0}$$

For instance, the Virasoro generators of the free scalar theory are expressed in terms of the modes as in (2.15).

3.7 Vertex operators

We shall connect 2d conformal field theory to string theory, considering vertex operators. The state-operator correspondence tells us that there are corresponding local operators for the tachyon and the massless states (2.24) in the bosonic closed string theory. The tachyon state is just a vacuum state with a certain momentum k^μ . Therefore, the corresponding operator is

$$|0;k\rangle \leftrightarrow :e^{ikX(0,0)}: . \quad (3.36)$$

This can be easily verified from the operator analogue of $p^\mu|0;k\rangle = k^\mu|0;k\rangle$, namely,

$$\partial X^\mu(z) :e^{ik\cdot X(0)}: \sim \frac{-i\alpha' k^\mu}{2z} :e^{ik\cdot X(0)}: .$$

The first excited states are obtained by acting $\alpha_{-1}^\mu \tilde{\alpha}_{-1}^\nu$ on the vacuum $|0;k\rangle$. Each mode in (3.3) can be extracted by the Fourier transformation of $\partial X^\mu(z)$ and the corresponding operator is

$$\alpha_{-m}^\mu = i\sqrt{\frac{2}{\alpha'}} \oint \frac{dz}{2\pi i} z^{-m} \partial X^\mu(z) \rightarrow i\sqrt{\frac{2}{\alpha'}} \frac{1}{(m-1)!} \partial^m X(0) .$$

Thus, we have operators corresponding to the massless states

$$\zeta_{\mu\nu} \alpha_{-1}^\mu \bar{\alpha}_{-1}^\nu |0;k\rangle \leftrightarrow \zeta_{\mu\nu} \partial X^\mu \bar{\partial} X^\nu :e^{ikX}:$$

where $\zeta_{\mu\nu}$ are the polarization tensors subject to $k^\mu \zeta_{\mu\nu} = 0$.

The string amplitude is

$$A_n = \sum_g \int (\mathcal{D}h_{ab})_{g,n} \int \mathcal{D}X^\mu e^{-S_\sigma[X^\mu, h_{ab}]} \prod_{i=1}^n \int d^2z \sqrt{h} V_i . \quad (3.37)$$

Here $\hat{V}_i = \int d^2z \sqrt{h} V_i$ is an operator corresponding to a string state. It is integrated out over the world-sheet to be Diff \times Weyl invariant.

Mass from vertex operator

Now, let us consider a constant scaling $z \rightarrow \lambda z$ and $\bar{z} \rightarrow \bar{\lambda} \bar{z}$. Under the scaling, a field transforms as $\phi(z, \bar{z}) \rightarrow \lambda^{-h} \bar{\lambda}^{-\bar{h}} \phi(z, \bar{z})$, which should compensate for the scaling of the measure $dz d\bar{z} \rightarrow \lambda \bar{\lambda} dz d\bar{z}$. Namely, $h = \bar{h} = 1$. Conformal dimensions of the vertex operators are

Name	\mathcal{O}	(h, \bar{h})
Tachyon	$e^{ik\cdot X}$	$(\frac{\alpha' k^2}{4}, \frac{\alpha' k^2}{4})$
1st excited states	$\zeta_{\mu\nu} \partial X^\mu \bar{\partial} X^\nu e^{ikX}$	$(1 + \frac{\alpha' k^2}{4}, 1 + \frac{\alpha' k^2}{4})$

The consistency condition for the tachyon leads to

$$M^2 = -k_{\text{Tachyon}}^2 = -\frac{4}{\alpha'} .$$

Similarly, the first excited states lead to the massless condition

$$M^2 = -k_{\text{1st}}^2 = 0 .$$

Both results are consistent with the analysis in (2.21) and (2.24). Here one may consider that one can obtain the string spectrum in arbitrary spacetime dimensions. However, as we will show in the next section, the on-shell condition only holds in $D = 26$ where the Weyl transformation is a quantum symmetry.

4 Weyl anomaly

The classical action (2.2) of the string sigma model is invariant under the Weyl symmetry, and the Weyl invariance implies the traceless condition (2.4) of the energy-momentum tensor. However, we will see that there can be an anomaly in the Weyl symmetry if the string world-sheet is curved. As we will see, it is characterized by the central charge $T_a^a = -\frac{c}{12}R^{(2)}$ where $R^{(2)}$ is the world-sheet Ricci scalar curvature.

Moreover, in string theory, a target space can also be a non-trivial curved spacetime where the action becomes

$$S = \frac{1}{4\pi\alpha'} \int d^2z \sqrt{h} h^{ab} \partial_a X^\mu \partial_b X^\nu G_{\mu\nu}(X) + \dots$$

This is no longer a free theory, and it describes an interacting non-linear theory, called the **(string) non-linear sigma model**. In an interacting quantum field theory, the breakdown of scale invariance is described in terms of a β function. In the non-linear sigma model, the anomaly of the Weyl invariance due to the curved target spacetime is also characterized by β -functions.

4.1 Weyl anomaly from curved world-sheet

Although the string sigma model is classically Weyl-invariant, in quantum theory, the trace of the energy-momentum tensor can be nonzero due to anomaly. Nonetheless, it is

- world-sheet diff invariant,
- zero on a flat world-sheet \mathbb{R}^2 ,
- world-sheet mass dimension two.

Hence, they constrain the form of the trace of the energy-momentum tensor as

$$T_a^a = kR^{(2)},$$

where $R^{(2)}$ is the world-sheet Ricci scalar curvature. In this subsection, we shall determine the coefficient k .

Given a curved world-sheet Riemann surface (Σ, h_{ab}) with non-trivial metric, we can always take a local coordinate (called **conformal gauge**) of the world-sheet Riemann surface in which the metric is a conformally flat

$$ds^2 = e^{2\omega(\sigma^1, \sigma^1)} (d\sigma^1 d\sigma^1 + d\sigma^2 d\sigma^2) = e^{2\omega(z, \bar{z})} dz d\bar{z}. \quad (4.1)$$

This local coordinate is called an **isothermal coordinate**. In the isothermal coordinate, the scalar curvature and non-trivial Christoffel symbols are expressed as

$$R^{(2)} = -8e^{-2\omega} \partial \bar{\partial} \omega, \quad \Gamma_{zz}^z = 2\partial\omega, \quad \Gamma_{z\bar{z}}^z = 2\bar{\partial}\omega \quad (4.2)$$

and the other Christoffel symbols are zero $\Gamma_{bc}^a = 0$. Therefore, the diagonal element of the energy-momentum tensor becomes

$$T_{z\bar{z}} = \frac{1}{4} e^{2\omega} T_a^a = -2k\partial\bar{\partial}\omega.$$

Also, the conservation $\nabla_a T_b^a = 0$ of the energy-momentum tensor now becomes

$$\begin{aligned} 0 &= \nabla^z T_{zz} + \nabla^{\bar{z}} T_{\bar{z}\bar{z}} = g^{\bar{z}\bar{z}} [\nabla_{\bar{z}} T_{zz} + \nabla_z T_{\bar{z}\bar{z}}] \\ &= g^{\bar{z}\bar{z}} [\bar{\partial} T_{zz} + (\partial - 2\partial\omega)(-2k\partial\bar{\partial}\omega)] \\ &= g^{\bar{z}\bar{z}} [\bar{\partial} (T_{zz} - 2k(\partial\partial\omega - \partial\omega\partial\omega))] . \end{aligned}$$

This implies that T_{zz} deviates from a holomorphic operator $T(z)$

$$T_{zz} = 2k(\partial\partial\omega - \partial\omega\partial\omega) + T(z) . \quad (4.3)$$

On a curved world-sheet, the conformal Ward-Takahashi identity (3.14) becomes

$$\begin{aligned} \delta_{\epsilon, \bar{\epsilon}} \mathcal{O}(w, \bar{w}) &= \int_{\Sigma} \frac{d^2 z}{2\pi i} \sqrt{h} \left\{ (\nabla^z \epsilon^z T_{zz} + \nabla^{\bar{z}} \bar{\epsilon}^{\bar{z}} T_{\bar{z}\bar{z}}) + (\nabla^{\bar{z}} \bar{\epsilon}^{\bar{z}} T_{\bar{z}\bar{z}} + \nabla^z \epsilon^z T_{zz}) \right\} \mathcal{O}(w, \bar{w}) \\ &= \int_{\Sigma} \frac{d^2 z}{2\pi i} \left\{ (\nabla_{\bar{z}} \epsilon(z) T_{zz} + \nabla_z \epsilon(z) T_{\bar{z}\bar{z}}) + (\nabla_z \bar{\epsilon}(\bar{z}) T_{\bar{z}\bar{z}} + \nabla_{\bar{z}} \bar{\epsilon}(\bar{z}) T_{zz}) \right\} \mathcal{O}(w, \bar{w}) \\ &= \int_{\Sigma} \frac{d^2 z}{2\pi i} \left\{ \bar{\partial}(\epsilon(z) T(z)) + \partial(\bar{\epsilon}(\bar{z}) \bar{T}(\bar{z})) \right\} \mathcal{O}(w, \bar{w}) , \end{aligned}$$

where we use the current conservation

$$T_{z\bar{z}} (\nabla^z \epsilon^{\bar{z}} + \nabla^{\bar{z}} \epsilon^z) = 0$$

from the second and third line. This means that $T(z)$ obeys the flat version of the conformal Ward-Takahashi identity (3.14). This gives the transform property of $T(z)$ as in the flat space (3.28)

$$\delta_{\epsilon} T(z) = \epsilon(z) \partial T(z) + 2\partial\epsilon(z) T(z) + \frac{c}{12} \partial^3 \epsilon(z) .$$

On the other hand, (4.3) yields

$$\delta_{\epsilon} T(z) = \delta_{\epsilon} T_{zz} - 2k(\partial\partial\delta_{\epsilon}\omega - 2\partial\omega\partial\delta_{\epsilon}\omega) . \quad (4.4)$$

Using (finite) transformations,

$$\begin{aligned} z &\rightarrow \tilde{z} = z - \epsilon(z) , \\ T_{zz} &\rightarrow \tilde{T}_{\bar{z}\bar{z}} = (\partial_z \tilde{z})^{-2} T_{zz} , \\ \omega(z) &\rightarrow \tilde{\omega}(\tilde{z}) = \omega(z) - \frac{1}{2} \log |\partial_z \tilde{z}|^2 , \end{aligned}$$

(4.4) can be expressed as

$$\delta_{\epsilon} T(z) = \epsilon(z) \partial T(z) + 2\partial\epsilon(z) T(z) - k\partial^3 \epsilon(z) . \quad (4.5)$$

Finally, we see that the Weyl anomaly is proportional to the central charge

$$T_a^a = kR^{(2)} = -\frac{c}{12} R^{(2)} . \quad (4.6)$$

We have seen in (3.33) that the central charge of the string sigma model is equal to the dimension D of the target space. Hence, naively looking, the bosonic string theory would suffer from the Weyl anomaly. However, as we will see in §5.1, we introduce “ghost CFT” to fix gauge in the world-sheet path integral, and the ghost CFT has $c^{\text{gh}} = -26$. As a result, the bosonic string theory is Weyl invariant at quantum level only if $D = 26$.

4.2 Non-linear sigma model

In §2, we start with a flat D -dimensional target space, and the quantization of bosonic string naturally leads to the graviton (2.24). Therefore, in general, we can consider that the target spacetime is curved with a non-trivial metric so that the action is

$$S[X^\mu, h_{ab}] = \frac{1}{4\pi\alpha'} \int d^2\sigma \left(\sqrt{h} h^{ab} \partial_a X^\mu \partial_b X^\nu G_{\mu\nu}(X) \right).$$

This is called the non-linear sigma model. In this subsection, we provide a concise overview of the non-linear sigma model where the massless fields (2.24) receive quantum corrections. Although we present key results, many of their derivations are not included here. For more details, we direct the reader to [GSW87, §3.4].

If we consider a perturbation from the flat-metric in this action

$$G_{\mu\nu}(X) = \eta_{\mu\nu} + f_{\mu\nu}(X),$$

then the partition function becomes

$$\begin{aligned} Z &= \int \mathcal{D}h_{ab} \mathcal{D}X^\mu e^{-S} \\ &= \int \mathcal{D}h_{ab} \mathcal{D}X^\mu e^{-S_0} \left(1 + \frac{1}{4\pi\alpha'} \int d^2\sigma \left(\sqrt{h} h^{ab} \partial_a X^\mu \partial_b X^\nu f_{\mu\nu}(X) \right) + \dots \right). \end{aligned}$$

Notice that the perturbative part is the graviton operator with wave function $f_{\mu\nu}(X) = \zeta_{\mu\nu} e^{ik \cdot X}$.

In a curved target, the theory is no longer free, so that we need to take into account various quantum effects. We will see that β -functions encode the anomaly of scaling invariance at quantum level. Note that the discussion here is brief, and the reader is referred to [CT89] for the details. As a 2d field theory, we can consider the vacuum expectation value (VEV) for X , which we set to X_0 :

$$\hat{X}(\sigma, \tau) = X_0 + X(\sigma, \tau).$$

On the other hand, X_0 is a certain point in spacetime and we will expand the metric around this point:

$$G_{\mu\nu}(X) = G_{\mu\nu} - \frac{1}{3} R_{\mu\lambda\nu\rho} X^\lambda X^\rho + \mathcal{O}(X^3),$$

where $G_{\mu\nu}$ and $R_{\mu\lambda\nu\rho}$ are a metric and a Riemann curvature tensor of the target spacetime at X_0 , respectively. In a field-theoretic sense, these can be understood as coupling constants

$$\begin{aligned} S &= \frac{1}{4\pi\alpha'} \int d^2\sigma \partial^a X^\mu \partial_a X^\nu \left(G_{\mu\nu} - \frac{1}{3} R_{\mu\lambda\nu\rho} X^\lambda X^\rho + \dots \right) \\ &\rightarrow \frac{1}{2} \int d^2\sigma \partial^a X^\mu \partial_a X^\nu \left(G_{\mu\nu} - \frac{2\pi\alpha'}{3} R_{\mu\lambda\nu\rho} X^\lambda X^\rho + \mathcal{O}(\alpha'^2) \right). \end{aligned}$$

Here we rescale the field $X \rightarrow \sqrt{2\pi\alpha'} X$ so that the expansion looks like a “stringy expansion”.

Perturbation theory for non-linear sigma model

We want to check if the theory (non-linear sigma model) has Weyl anomaly. As we briefly saw, it is an interacting theory, and interacting theories have generally non-trivial β -functions:

$$\beta[g] \equiv E \frac{\partial}{\partial E} g(E) = \frac{\partial}{\partial(\log E)} g(E) ,$$

where g is a coupling constant and E is a characteristic energy scale. When we consider a global scaling of coordinate: $z \rightarrow \tilde{z} = \lambda z = (1 - \epsilon)z = e^{-\epsilon}z$, energy scales oppositely: $E \rightarrow \tilde{E} = \frac{1}{\lambda}E = e^{\epsilon}E$. So the β -function can be written as

$$\beta[g] = \frac{\partial}{\partial \epsilon} g(\epsilon) .$$

The variation of the action is expressed in two ways:

$$\delta_\epsilon S = \begin{cases} \int \frac{d^2\sigma}{2\pi} \sqrt{h} \delta_\epsilon h^{ab} T_{ba} = -\epsilon \int \frac{d^2\sigma}{2\pi} T_a^a , \\ \frac{1}{4\pi\alpha'} \int d^2\sigma \partial^a X^\mu \partial_a X^\nu \left(\epsilon \frac{\partial}{\partial \epsilon} G_{\mu\nu}(\epsilon) + \dots \right) , \end{cases}$$

where the first variation is a formal transformation of the theory, and in quantum regime, it should be proportional to the trace part of the energy-momentum tensor. On the other hand, the second variation is the actual theory with an assumption that ϵ dependence of the theory is only in the coupling constants. Identifying them, we have

$$T_a^a = -\frac{1}{2\alpha'} \beta[G_{\mu\nu}] \partial^b X^\mu \partial_b X^\nu + \dots .$$

This shows that the anomaly is parametrized by β -functions.

Let us consider perturbation theory, namely loop corrections to two-point function etc, so that we can see if the theory is anomalous.

$$\begin{aligned} & \langle X^\mu(\sigma_1) X^\nu(\sigma_2) \rangle \\ &= \int \frac{d^2k}{(2\pi)^2} \frac{2\pi\alpha'}{k^2} e^{ik \cdot (\sigma_1 - \sigma_2)} \left\{ G^{\mu\nu} + \frac{2\pi\alpha'}{3} R^{\mu\nu} \left(\int \frac{d^2p}{(2\pi)^2} \frac{1}{p^2} + \frac{1}{k^2} \int \frac{d^2p}{(2\pi)^2} \right) + \dots \right\} . \end{aligned}$$

Among terms proportional to $R^{\mu\nu}$ in the integrand, the first integral provides a logarithmic and the second does quadratic divergence. Since the quadratic divergence is discarded by dimensional regularization, we focus on the logarithmic divergence and introduce regularization parameters:

$$\int_E^\Lambda \frac{d^2p}{(2\pi)^2} \frac{1}{p^2} = \frac{1}{2\pi} \log \left(\frac{\Lambda}{E} \right) ,$$

where Λ is an ultra-violet (UV) energy scale supposed to be ∞ , and E is an infra-red (IR) energy scale supposed to be our life energy scale, which is very low (~ 0).

The divergence can be subtracted by counter terms as follows. The bare action $\hat{S} = S + S_{ct}$ describes UV physics of energy scale Λ whereas the physical action is S describes IR physics of energy scale E . They are related by renormalization

$$\hat{S} = \frac{1}{4\pi\alpha'} \int d^2\sigma \partial^a \hat{X}^\mu \partial_a \hat{X}^\nu \left(\hat{G}_{\mu\nu} - \frac{1}{3} \hat{R}_{\mu\lambda\nu\rho} \hat{X}^\lambda \hat{X}^\rho + \dots \right)$$

with

$$\begin{aligned}\widehat{X}^\mu &= Z_v^\mu X^v, \quad Z_v^\mu = \delta_v^\mu + \sum_{n=1}^{\infty} \alpha'^n Z_{v,(n)}^\mu(\Lambda/E), \\ \widehat{G}^{\mu\nu} &= G^{\mu\nu} + \sum_{n=1}^{\infty} \alpha'^n G_{(n)}^{\mu\nu}(\Lambda/E), \text{etc.}\end{aligned}$$

Note that the bare action only depends on Λ (not on E), and hence, the bare coupling constants ($\widehat{G}_{\mu\nu}$ etc) only depend on high energy Λ . The counter terms lead to other contributions

$$\langle X^\mu(\sigma_1)X^\nu(\sigma_2) \rangle \sim \left\{ G^{\mu\nu} + \frac{\alpha'}{3} R^{\mu\nu} \log\left(\frac{\Lambda}{E}\right) - \alpha' \left(G_{(1)}^{\mu\nu} + Z_{(1)}^{\mu\nu} + Z_{(1)}^{v\mu} \right) + \dots \right\}. \quad (4.7)$$

Unfortunately, the equation above cannot fix the ratio between $G_{(1)}$ and $Z_{(1)}$. We need further information like a 4-point function to determine the ratio. We simply list the result:

$$G_{(1)}^{\mu\nu} = R^{\mu\nu} \log\left(\frac{\Lambda}{E}\right), \quad Z_{(1)}^{\mu\nu} = -\frac{1}{3} R^{\mu\nu} \log\left(\frac{\Lambda}{E}\right),$$

which does cancel the divergent term in (4.7). From the result we can derive the β -function:

$$\begin{aligned}G_{\mu\nu}(E, \Lambda) &= \widehat{G}_{\mu\nu}(\Lambda) - \alpha' R_{\mu\nu} \log\left(\frac{\Lambda}{E}\right), \\ \beta[G_{\mu\nu}] &= \frac{\partial}{\partial(\log E)} G_{\mu\nu}(E, \Lambda) = \alpha' R_{\mu\nu}.\end{aligned}$$

Therefore, for the theory to be anomaly-free, the spacetime is required to be Ricci-flat ($R_{\mu\nu} = 0$).

Non-linear sigma model (general)

If we incorporate all the massless states (2.24), we can write the most general form of the non-linear sigma model for a closed string as

$$S_{\text{closed}} = \frac{1}{4\pi\alpha'} \int d^2\sigma \left(\sqrt{h} h^{ab} \partial_a X^\mu \partial_b X^\nu G_{\mu\nu}(X) + i\epsilon^{ab} \partial_a X^\mu \partial_b X^\nu B_{\mu\nu}(X) + \alpha' \sqrt{h} R^{(2)} \Phi(X) \right). \quad (4.8)$$

Note that the B -field is a higher dimensional analogue of gauge fields (mathematically called “gerbe”). Its gauge transformation is given by

$$\delta B_{\mu\nu} = \partial_\mu \Lambda_\nu - \partial_\nu \Lambda_\mu, \quad (4.9)$$

and the field strength $H_{\mu\nu\lambda}$ is defined as

$$H_{\mu\nu\lambda} = \partial_\mu B_{\nu\lambda} + \partial_\nu B_{\lambda\mu} + \partial_\lambda B_{\mu\nu}.$$

The B -field often is the source of “stringy” effects, and it plays an important role on many occasions.

Again, the Weyl anomaly is characterized by β -functions for the massless fields

$$T_a^a = -\frac{1}{2\alpha'}\beta[G_{\mu\nu}]\partial_a X^\mu \partial^a X^\nu - \frac{i}{2\alpha'}\beta[B_{\mu\nu}]\epsilon^{ab}\partial_a X^\mu \partial_b X^\nu - \frac{1}{2}\beta[\Phi]R^{(2)}$$

where β -functions are expressed as

$$\begin{aligned}\beta[G_{\mu\nu}] &= \alpha'R_{\mu\nu} + 2\alpha'\nabla_\mu\nabla_\nu\Phi - \frac{\alpha'}{4}H_{\mu\lambda\rho}H_\nu^{\lambda\rho} + \mathcal{O}(\alpha'^2), \\ \beta[B_{\mu\nu}] &= -\frac{\alpha'}{2}\nabla^\lambda H_{\lambda\mu\nu} + \alpha'\nabla^\lambda\Phi H_{\lambda\mu\nu} + \mathcal{O}(\alpha'^2), \\ \beta[\Phi] &= \frac{D-26}{6} - \frac{\alpha'}{2}\nabla^2\Phi + \alpha'\nabla^\lambda\Phi\nabla_\lambda\Phi - \frac{\alpha'}{24}H_{\mu\lambda\rho}H^{\mu\lambda\rho} + \mathcal{O}(\alpha'^2).\end{aligned}$$

In order for the theory to be non-anomalous, all the β -functions must be zero, and the vanishing of $\beta[\Phi]$ again requires $D = 26$. The perturbative quantum field theory on the world-sheet can describe the classical gravity from the vanishing of β function.

1. The equation $\beta[G_{\mu\nu}] = 0$ is analogous to Einstein's equation with source terms from the B -field and the dilaton.
2. The equation $\beta[B_{\mu\nu}] = 0$ is the generalization of the Maxwell's equation.

The condition that β -functions vanish ($\beta = 0$) is equivalent to the equation of motion for the following spacetime effective action⁴

$$S_{\text{eff}} = \frac{1}{2\kappa_D^2} \int d^D X \sqrt{-G} e^{-2\Phi} \left[\frac{2(D-26)}{3\alpha'} + R - \frac{1}{12}H_{\mu\lambda\rho}H^{\mu\lambda\rho} + 4\nabla^\lambda\Phi\nabla_\lambda\Phi + \mathcal{O}(\alpha'^2) \right]. \quad (4.10)$$

The action (4.8) tells us another important fact in string theory. The dilaton field is coupled to the world-sheet Ricci scalar. If the dilaton takes an expectation value $\langle\Phi(X)\rangle = \Phi$, the dilaton action gives the Euler characteristic of the world-sheet Riemann surface

$$\begin{aligned}S_{\text{dilaton}} &= \frac{1}{4\pi\alpha'} \int d^2\sigma \left(\alpha' \sqrt{h} R^{(2)} \Phi(X) \right) \\ &\rightarrow \frac{1}{4\pi} \int d^2\sigma \left(\sqrt{h} R^{(2)} \Phi \right) = \Phi(2-2g)\end{aligned}$$

In other words, the 2d gravity is topological with no physical degrees of freedom. Defining $g_s = e^\Phi$, and the dilaton part of the action becomes

$$e^{-S_{\text{dilaton}}} = g_s^{2g-2}.$$

This implies that g_s can be understood as a string coupling. An n -point string tree amplitude can be understood as $n-2$ cylinders attaching to a cylinder. Hence, the amplitude should be proportional to g_s^{n-2} . A higher loop (higher genus) amplitude can be derived by attaching g cylinders to the tree amplitude, and the amplitude should be $\hat{A}_{n,g} \propto g_s^{n-2+2g}$. (See Figure 7.) Usually, vertex operators are re-normalized so that g_s^n is included in the definition of $V_1 \cdots V_n$. Therefore, the re-normalized amplitude should be

$$A_{n,g} \propto g_s^{2g-2},$$

which coincide with the dilaton action. In conclusion, the string coupling g_s is just the expectation value of the dilaton field.

⁴We are in the regime where the perturbation theory works well. Namely, the length scale of the target

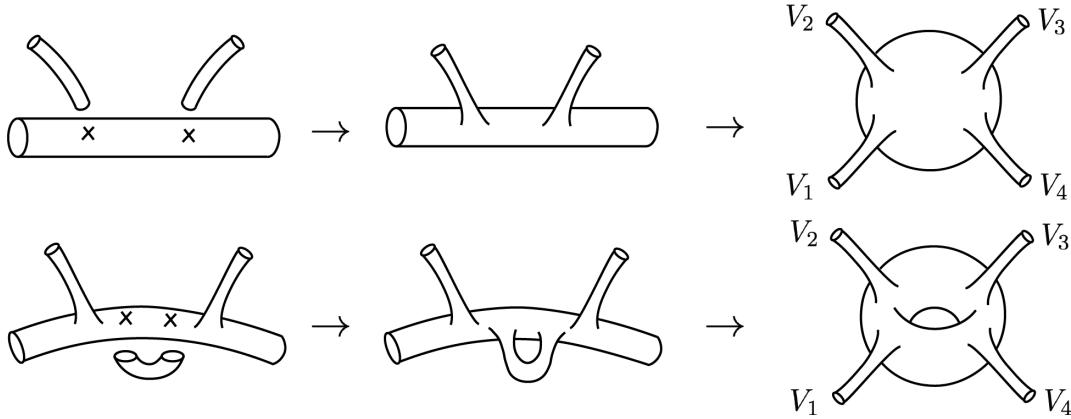


Figure 7: 4-point amplitude example. The upper one is a construction of 4pt tree amplitude from cylinders. The lower one is a construction of 4-point 1-loop amplitude from the tree amplitude.

5 BRST quantization

We have been studying bosonic string theory, but there are several caveats.

- In the light-cone quantization (2.6), Lorentz invariance is not manifest. Can we quantize strings in a way that is manifestly Lorentz invariant?
- We have seen that the Weyl symmetry of the string sigma model is anomalous (4.6) on a curved world-sheet. How can the bosonic string be anomaly-free?
- Although we learned that string amplitude is expressed via Feynman path integral (3.37), we do not know how to perform this path integral. In particular, the path integral is endowed with huge gauge symmetries, world-sheet diffeomorphism and Weyl symmetry. How can we treat integration measure and fix gauge in the path integral?

To answer these questions, we will study the quantization procedure via path integral, which is often called **modern covariant quantization**. This method uses the analog of the Faddeev-Popov method of gauge theories. Furthermore, the physical state condition is imposed via the BRST symmetry.

After gauge-fixing the reparametrization and Weyl symmetry, the integral over h_{ab} turns into a path-integral over ghost CFT, which has $c^g = -26$. Therefore, in order for the theory to be Weyl anomaly-free, the matter part of the theory has to be a CFT with $c^X = 26$.

5.1 Quantization via path integral

Faddeev-Popov gauge fixing

As expressed in (3.37), the integration measure can be written as $[\mathcal{D}X_\mu][\mathcal{D}h_{ab}]_{g,n}$ where the scalar fields and the metric on 2d surfaces with genus g and n marked points are

space is long compared to the string scale. We thus can ignore the internal structure of the string and write the following low energy effective theory for quantum gravity.

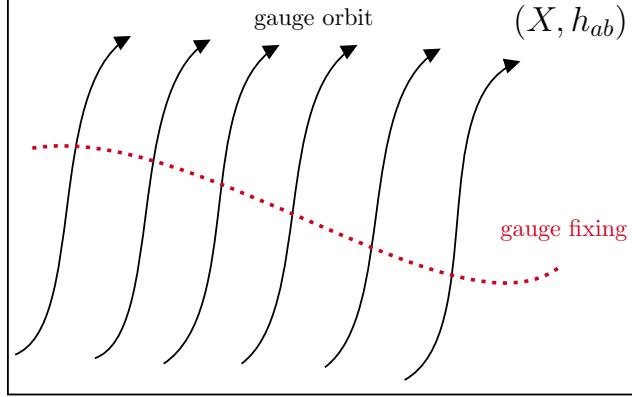


Figure 8: We need to fix a gauge in the field configuration space. Physically inequivalent configurations are depicted by the dotted line.

integrated out. However, if there is no anomaly, this integral has the world-sheet diffeomorphisms and Weyl symmetry under which the field configurations are transformed as

$$\begin{aligned} X_\mu^\zeta(\tilde{\sigma}) &= X_\mu(\sigma) \\ h_{ab}^\zeta(\tilde{\sigma}) &= e^{2\omega(\sigma)} \frac{\partial \sigma^c}{\partial \tilde{\sigma}^a} \frac{\partial \sigma^d}{\partial \tilde{\sigma}^b} h_{cd}(\sigma), \end{aligned} \tag{5.1}$$

where ζ indicates $\text{Diff} \times \text{Weyl}$. These symmetries are redundant and integrating along these directions just gives rise to the volume of the symmetry group. (In Figure 8, the solid arrows schematically draw gauge redundancy and the dotted line shows physically distinct configurations.) Thus, we need to carry out gauge-fixing. Thankfully, there is a standard method to fix gauge, introduced by Faddeev and Popov. A basic idea is to insert the identity of the following form in the path integral:

$$1 = \int \mathcal{D}\zeta \delta(h - \hat{h}^\zeta) \det \left(\frac{\delta \hat{h}^\zeta}{\delta \zeta} \right) \tag{5.2}$$

where the Jacobian factor $\det \left(\frac{\delta \hat{h}^\zeta}{\delta \zeta} \right)$ is called **Faddeev-Popov determinant** and we denote it by $\Delta_{FP}[\hat{h}]$. Since $\mathcal{D}\zeta$ is a gauge-invariant measure, $\Delta_{FP}[\hat{h}]$ is independent of ζ so that it can be factored out of the integral above. The insertion of this identity into the path integral fixes the metric as \hat{h} because of the delta functional. Because $\int \mathcal{D}\zeta$ integral only contributes an infinite multiplicative factor, we can discard this integral. Therefore, after Faddeev-Popov gauge fixing, the form of the path integral can be schematically written as

$$Z[\hat{h}] = \int \mathcal{D}X \Delta_{FP}[\hat{h}] e^{-S_\sigma[X, \hat{h}]} \tag{5.3}$$

We should make several remarks about this procedure.

- First, the conformal transformations are residual gauge symmetries not fixed above. We have to throw away these residual gauge symmetries in the path integral in order to avoid over-counting. Indeed we will be careful to fix this extra residual gauge freedom when computing string amplitudes.

- Second, there are caveats related to global properties of the world-sheet Riemann surface $\Sigma_{g,n}$. In fact, metrics on a Riemann surface encode the information of “shape” of the Riemann surface $\Sigma_{g,n}$ called **complex moduli**, which is not accounted for by local gauge transformations ζ . The space which parametrizes “shape” of the Riemann surface is called **moduli space of Riemann surface** $\Sigma_{g,n}$ which is $(6g - 6 + 2n)$ -dim $_{\mathbb{R}}$:

$$\mathcal{M}_{g,n} := \frac{[\mathcal{D}h_{ab}]_{g,n}}{\text{Diff} \times \text{Weyl}}$$

Therefore, the path integral involves integrals over $\mathcal{M}_{g,n}$ as well. At this moment, we postpone both the issues and will come back to them in §6 on string amplitudes.

Now let us take an infinitesimal version of (5.1) where a Weyl transformation is parameterized by $\omega(\sigma)$ and an infinitesimal diffeomorphism by $\delta\sigma^\alpha = \epsilon^\alpha(\sigma)$. Subsequently, the change of the metric under a gauge transformation is read off

$$\delta\hat{h}_{ab} = 2\omega\hat{h}_{ab} + \nabla_a\epsilon_b + \nabla_b\epsilon_a := 2\tilde{\omega}\hat{h}_{ab} + (P \cdot \epsilon)_{ab}, \quad (5.4)$$

where we decompose it into

$$\begin{aligned} (P \cdot \epsilon)_{ab} &= \nabla_a\epsilon_b + \nabla_b\epsilon_a - h_{ab}(\nabla \cdot \epsilon) \\ \tilde{\omega} &= \omega + \frac{1}{2}(\nabla \cdot \epsilon). \end{aligned} \quad (5.5)$$

Indeed, the operator P maps vectors ϵ_a to symmetric traceless two-tensors $(P \cdot \epsilon)_{ab}$. Thus, the Faddeev-Popov determinant can be written as

$$\Delta_{FP}[\hat{h}] = \det \frac{\delta(P \cdot \epsilon, \tilde{\omega})}{\delta(\epsilon, \omega)} = \det \begin{vmatrix} P & 0 \\ * & 1 \end{vmatrix} = \det P.$$

To compute $\det P$, we use **Faddeev-Popov ghosts**, which can be understood as an infinite-dimensional version of the following integral. Given a matrix M_{ij} , its determinant can be expressed as a Grassmann integral

$$\int \prod_{i=1}^n d\psi_i d\theta_i \exp(\theta_i M_{ij} \psi_j) = \det M.$$

where θ, ψ are Grassmann variables. Accordingly, we introduce anti-commuting fermionic fields, c^a (ghosts) and b_{ab} (anti-ghost) where b_{ab} transforms as a symmetric traceless tensor and c^a as a vector. Then, we can express

$$\Delta_{FP}[\hat{h}] = \int \mathcal{D}c \mathcal{D}b \exp \left(\frac{i}{2\pi} \int d^2\sigma \sqrt{-\hat{h}} b^{ab} (P \cdot c)_{ab} \right) := \int \mathcal{D}c \mathcal{D}b \exp [iS_g],$$

where the ghost action can be written as

$$S_g = \frac{1}{2\pi} \int d^2\sigma \sqrt{-\hat{h}} b^{ab} \nabla_a c_b. \quad (5.6)$$

Even though the bc ghost fields were introduced to fix a gauge, they look like dynamical fields with the action above. Consequently, the Faddeev-Popov gauge fixing results in a fermionic 2d CFT, usually called **bc ghost CFT**.

Let us make some remarks about the equations of motion of S_g :

- The equation of motion for c_a is given by

$$P \cdot c = \nabla_a c_b + \nabla_b c_a - h_{ab}(\nabla \cdot c) = 0. \quad (5.7)$$

Therefore the solutions for c are in one-to-one correspondence with the **conformal Killing vectors**, which are the generators of the residual symmetry.

- The equation of motion for b_{ab} is

$$\nabla_a b^{ab} = 0. \quad (5.8)$$

We will understand the geometric meaning of these equations when discussing the moduli space of Riemann surfaces.

To understand the properties of the bc ghost CFT, it is convenient to use the Euclidean signature so that we will perform Wick rotation in what follows. Then, the factor of i in the action disappears. The expression for the full partition function (5.3) is

$$Z[\hat{h}] = \int \mathcal{D}X \mathcal{D}c \mathcal{D}b \exp \left(-S_\sigma[X, \hat{h}] - S_g[b, c, \hat{h}] \right). \quad (5.9)$$

bc ghost CFT

Now let us study the bc ghost CFT more in detail. For this purpose, we pick the conformal gauge (4.1) for a world-sheet metric. Using (4.2), the ghost action can be written as

$$\begin{aligned} S_g &= \frac{1}{2\pi} \int d^2z (b_{zz} \nabla_{\bar{z}} c^z + b_{\bar{z}\bar{z}} \nabla_z c^{\bar{z}}) \\ &= \frac{1}{2\pi} \int d^2z b_{zz} \partial_{\bar{z}} c^z + b_{\bar{z}\bar{z}} \partial_z c^{\bar{z}} \end{aligned} \quad (5.10)$$

For the sake of simplicity, let us define

$$b = b_{zz}, \quad \bar{b} = b_{\bar{z}\bar{z}}, \quad c = c^z, \quad \bar{c} = c^{\bar{z}}.$$

Then, the action simplifies to

$$S_g = \frac{1}{2\pi} \int d^2z (b \bar{\partial} c + \bar{b} \partial \bar{c}), \quad (5.11)$$

which gives the equations of motion

$$\bar{\partial}b = \partial\bar{b} = \bar{\partial}c = \partial\bar{c} = 0 \quad (5.12)$$

Thus, we see that b and c are holomorphic fields, while \bar{b} and \bar{c} are anti-holomorphic.

To compute the OPEs of the bc ghost fields, we use the path integral techniques in §3.3:

$$0 = \int \mathcal{D}c \mathcal{D}b \frac{\delta}{\delta b(z)} \left[e^{-S_g} b(w) \right] = \int \mathcal{D}c \mathcal{D}b e^{-S_g} \left[-\frac{1}{2\pi} \bar{\partial}c(z)b(w) + \delta(z-w) \right],$$

which tells us that

$$\bar{\partial}c(z)b(w) = 2\pi \delta(z-w).$$

We can perform a similar computation for $c(z)$, which yields

$$\bar{\partial}b(z)c(w) = 2\pi \delta(z-w).$$

We can integrate both of these equations using (3.5). Then, we obtain the bc OPE

$$\begin{aligned} b(z)c(w) &= \frac{1}{z-w} + \dots \\ c(w)b(z) &= \frac{1}{w-z} + \dots \end{aligned}$$

In fact, the second equation follows from the first equation and Fermi statistics. Hence, the OPEs of $b(z)b(w)$ and $c(z)c(w)$ are trivial for the obvious reason.

In any CFT, it is of most importance to find the form of the energy-momentum tensor. The energy-momentum tensor is obtained via Noether's theorem with respect to worldsheet transformations $\delta z = \epsilon(z)$, under which

$$\delta b = (\epsilon\partial + 2(\partial\epsilon))b, \quad \delta c = (\epsilon\partial - (\partial\epsilon))c.$$

Indeed both b and c are primary fields with conformal dimensions $h = 2$ and $h = -1$, respectively, which can be easily seen from their index structure b_{zz} and c^z . From these rules, one can deduce the form of the energy-momentum tensor

$$T^g(z) = -2 :b(z)\partial c(z): + :c(z)\partial b(z):. \quad (5.13)$$

In fact, this form can be obtained from the first principle, namely the variation of the action under the metric (Exercise).

The OPEs of b and c with the stress tensor are

$$\begin{aligned} T^g(z)c(w) &= -\frac{c(w)}{(z-w)^2} + \frac{\partial c(w)}{z-w} + \dots \\ T^g(z)b(w) &= \frac{2b(w)}{(z-w)^2} + \frac{\partial b(w)}{z-w} + \dots \end{aligned}$$

so that b and c are primary fields of conformal dimension 2 and -1 , respectively. Consequently, they admit the mode expansions

$$b(z) = \sum_{m \in \mathbb{Z}} \frac{b_m}{z^{m+2}} \quad c(z) = \sum_{m \in \mathbb{Z}} \frac{c_m}{z^{m-1}}.$$

The ghost OPEs give the commutation relations

$$\{b_m, c_n\} = \delta_{m+n,0}, \quad \{c_m, c_n\} = \{b_m, b_n\} = 0. \quad (5.14)$$

Also, the Virasoro generators of the bc ghost are expressed as

$$L_m^g = \sum_{n \in \mathbb{Z}} (2m-n) :b_n c_{m-n}: + a^g \delta_{m,0}. \quad (5.15)$$

The constant a^g in the zero mode L_0^g is determined by the Casimir energy $\frac{c^g}{24}$ in (3.32) and the commutation relation of the bc modes $\frac{1}{12} = -\sum_{n>0} n$

$$a^g = \frac{-26}{24} + \frac{1}{12} = -1.$$

This is consistent with the commutation relation $L_0 = [L_1, L_{-1}]$ ([[Pol98](#), (2.7.20)]).

Finally, we can compute the TT OPE (Exercise)

$$T^g(z) T^g(w) = \frac{-13}{(z-w)^4} + \frac{2T^g(w)}{(z-w)^2} + \frac{\partial T^g(w)}{z-w} + \dots$$

Now one can read off the central charge of the bc ghost system, which is

$$c^g = 2(-13) = -26.$$

We learned that the Weyl symmetry is anomalous unless $c = 0$. Since the Weyl symmetry is a gauge symmetry, the theory must be Weyl anomaly-free. Since the total central charge of the string sigma model and ghost theory ([5.9](#)) is given by $c = c^X + c^g$, the dimension of the target space must be $D = 26$. Again, we obtain the critical dimension of the bosonic string theory!

5.2 BRST quantization

In 4d Yang-Mills theory, the Lagrangian with Faddeev-Popov ghosts has the continuous symmetry, called **BRST symmetry** (Becchi-Rouet-Stora-Tyupin). The BRST symmetry is generated by a nilpotent charge Q_B ($Q_B^2 = 0$) that commutes with the Hamiltonian. The nilpotency of the BRST charge has substantial consequences. The BRST transformations come from the gauge symmetry, all physical states must be BRST-invariant. Hence, we require a physical state to be annihilated by Q_B

$$Q_B |\text{phys}\rangle = 0.$$

However, one can always add a state of the form $Q_B |\chi\rangle$ since this state will be annihilated by Q_B because of the nilpotent BRST charge. However, this state is orthogonal to all physical states, and therefore it is a **null state**. Thus, two states related by

$$|\psi'\rangle = |\psi\rangle + Q_B |\chi\rangle$$

have the same inner products with all the physical states and they are thus indistinguishable. This is the remnant in the gauge-fixed version of the original gauge symmetry. As a result, the Hilbert space of physical states is isomorphic to the Q_B -cohomology, i.e.

$$\mathcal{H}^{\text{phys}} \cong \frac{\mathcal{H}^{Q_B\text{-closed}}}{\mathcal{H}^{Q_B\text{-exact}}}.$$
 (5.16)

This covariant way of determining the physical Hilbert space with ghosts is known as **BRST quantization**.

We are now ready to apply this formalism to the bosonic string [[Pol98](#), §4]. Combining ([3.2](#)) and ([5.11](#)), the action we consider is

$$S_X + S_g = \frac{1}{2\pi} \int d^2z \left(\frac{1}{\alpha'} \partial X^\mu \bar{\partial} X_\mu + b \bar{\partial} c + \bar{b} \partial \bar{c} \right).$$

This is invariant under the following BRST transformations:

$$\begin{aligned} \delta_B X^\mu &= i\epsilon(c\partial + \bar{c}\bar{\partial})X^\mu, \\ \delta_B c &= i\epsilon c\partial c \quad \delta_B \bar{c} = i\epsilon \bar{c}\bar{\partial}\bar{c}, \\ \delta_B b &= i\epsilon(T^X + T^g) \quad \delta_B \bar{b} = i\epsilon(\bar{T}^X + \bar{T}^g), \end{aligned}$$
 (5.17)

where the explicit forms of the energy-momentum tensors can be found in (3.25) and (5.13). Note that we impose (5.12) here. This exhibits typical features of the BRST transformation: a bosonic field is transformed into a fermionic field and vice versa, and the ghost field b is transformed into the energy-momentum tensor. Noether's theorem tells us that there is a classical current associated to the BRST symmetry, and the holomorphic part of the BRST current takes the form (Exercise)

$$\begin{aligned} j_B &= c(z)T^X(z) + \frac{1}{2} :c(z)T^g(z): + \frac{3}{2} :\partial^2 c(z): \\ &= c(z)T^X(z) + :b(z)c(z)\partial c(z): + \frac{3}{2} :\partial^2 c(z): . \end{aligned} \quad (5.18)$$

and the BRST charge is defined by

$$Q_B = \oint \frac{dz}{2\pi i} j_B .$$

It follows from (5.17) that

$$\{Q_B, b(z)\} = (T^X + T^g) \rightarrow \{Q_B, b_m\} = L_m^X + L_m^g . \quad (5.19)$$

Computing $j_B j_B$ OPE, one can convince oneself that the BRST charge is nilpotent $Q_B^2 = 0$ if and only if $D = 26$ (Exercise). Therefore, j_B is a primary field with $h = 1$, and Q_B is a conserved charge at quantum level only when $D = 26$. Furthermore, we can express the BRST charge in terms of the X^μ Virasoro operators and the ghost oscillators as

$$Q_B = \sum_n c_n (L_{-n}^X - \delta_{n,0}) + \sum_{m,n} \frac{m-n}{2} :c_m c_n b_{-m-n}: , \quad (5.20)$$

where the normal ordering constant comes from $\{Q_B, b_0\} = L_0^X + L_0^g$. In the case of closed strings, there is the anti-holomorphic part \bar{Q}_B , and the total BRST charge is $Q_B + \bar{Q}_B$.

We will find the physical open string states in the BRST context. According to our previous discussion, the physical states will have to be annihilated by the BRST charge, and not be of the form $Q_B |\cdot\rangle$.

First, we have to describe our extended Hilbert space which includes the ghosts. As far as the X^μ oscillators are concerned, the situation is the same as in (2.14), so we need to consider only the ghost Hilbert space. Indeed the ghost commutation relations (5.14) generate a two-state spin system $|\uparrow\rangle, |\downarrow\rangle$ where

$$\begin{aligned} b_0 |\downarrow\rangle &= 0, & b_0 |\uparrow\rangle &= |\downarrow\rangle , \\ c_0 |\uparrow\rangle &= 0, & c_0 |\downarrow\rangle &= |\uparrow\rangle , \\ b_{n>0} |\uparrow\rangle &= b_{n>0} |\downarrow\rangle = c_{n>0} |\uparrow\rangle = c_{n>0} |\downarrow\rangle = 0 . \end{aligned} \quad (5.21)$$

The full Hilbert space will be a tensor product of the two $|k, \uparrow\rangle, |k, \downarrow\rangle$ by acting creation operators where $|k\rangle = |0; k\rangle$ denotes the vacuum of the matter theory. From the light-cone quantization, we know that there is only one vacuum called Tachyon. Therefore, we have to pick the ghost vacuum among the two spin states. For this purpose, we further impose one more condition, namely

$$b_0 |\text{phys}\rangle = 0 . \quad (5.22)$$

This is sometimes called the **Siegel gauge** [GSW87, §3.2]. Under this condition, (5.21) tells us that the correct ghost vacuum is $|\downarrow\rangle$. We can now create states from this vacuum

by acting with the negative modes of the ghosts b_m, c_n . Note that $c_0|\downarrow\rangle = |\uparrow\rangle$ does not satisfy the Siegel condition (5.22). Also, (5.22) yields the condition

$$0 = \{Q_B, b_0\}|\text{phys}\rangle = (L_0^X + L_0^g)|\text{phys}\rangle. \quad (5.23)$$

Therefore, if a physical state is at level N with momentum k , $|\text{phys}\rangle = |N; k\rangle$, then we have

$$k^2 = \frac{1-N}{\alpha'} \quad (5.24)$$

which is consistent with the light-cone gauge. As in (2.19), we can define the number operator with the ghost sector as

$$N = \sum_{n>0} \eta_{\mu\nu} : \alpha_{-n}^\mu \alpha_n^\nu : + n (: b_{-n} c_n : + : c_{-n} b_n :) \quad (5.25)$$

so that the Virasoro zero mode is expressed as

$$L_0 = \alpha' k^2 + N - 1. \quad (5.26)$$

Now, let us impose the physical condition (5.16) for open-string states level by level. To this end, let us explicitly write the open-string mode expansions of the zero modes of the Virasoro generators

$$\begin{aligned} L_0^X &= \alpha' p^2 + \sum_{n=1}^{\infty} \alpha_{-n} \cdot \alpha_n = \alpha' p^2 + \alpha_{-1} \cdot \alpha_1 + \dots, \\ L_0^g &= -1 + \sum_{n \in \mathbb{Z}} n : b_{-n} c_n := -1 + b_{-1} c_1 + c_{-1} b_1 + \dots, \end{aligned} \quad (5.27)$$

(Compare with (2.15) and (5.15).) At level zero, there is only the vacuum state and the BRST quantization leads to $|k, \downarrow\rangle$

$$Q_B |k, \downarrow\rangle = (L_0^X - 1)c_0 |k, \downarrow\rangle = (L_0^X - 1)|k, \uparrow\rangle = 0.$$

and it is not Q_B -exact.

At the first level, the possible operators that can act on the vacuum $|k, \downarrow\rangle$ are α_{-1}^μ , b_{-1} and c_{-1} . The most general state of this form is then

$$|\psi\rangle = (\zeta \cdot \alpha_{-1} + \beta b_{-1} + \gamma c_{-1})|k, \downarrow\rangle, \quad (5.28)$$

which has 28 parameters: a 26-dimensional vector ζ_μ and two more constants β, γ . First we note that (5.24) yields the massless condition $k^2 = 0$. In addition, the BRST condition demands

$$0 = Q_B |\psi\rangle = 2((k \cdot \zeta)c_{-1} + \beta k \cdot \alpha_{-1})|k, \downarrow\rangle.$$

This only holds if $k \cdot \zeta = 0$ and $\beta = 0$. Therefore the BRST condition removes the unphysical anti-ghost excitations as well as all polarizations that are not orthogonal to the momentum, thereby eliminating two out of the 26 + 2 original states. Hence, there are only 26 parameters left.

Next, we have to make sure that this state is not Q_B -exact: a general state $|\chi\rangle$ is of the same form as (5.28), but with parameters ζ'^μ , β' and γ' . Thus, the most general Q_B -exact state at this level with $k^2 = 0$ will be

$$Q_B |\chi\rangle = 2(k \cdot \zeta' c_{-1} + \beta' k \cdot \alpha_{-1})|k, \downarrow\rangle.$$

This means that the c_{-1} part in (5.28) is BRST-exact and that the polarization has the equivalence relation $\zeta_\mu \sim \zeta_\mu + 2\beta' k_\mu$. This leaves us with the 24 physical degrees of freedom we expect for a massless vector particle in 26 dimensions. In sum, the physical state at level one is

$$\{|\zeta; k\rangle ; \quad k \cdot \zeta = 0\} / \zeta_\mu \sim \zeta_\mu + 2\beta' k_\mu .$$

The same procedure can be followed for the higher levels of open string states. It can be proved that $\mathcal{H}_{\text{light-cone}}$ is isomorphic to the Q_B -cohomology, and the inner product is positive-definite. In the case of the closed string, we have to use $Q_B + \bar{Q}_B$ for the BRST quantization.

6 Bosonic string amplitudes

Finally, we will discuss the string amplitude, which provides information about various features of string theory. The string amplitude can be expressed as follows:

$$A_n = \sum_g \int \frac{[\mathcal{D}h_{ab}]_{g,0}}{(\text{Diff} \times \text{Weyl})} \int \mathcal{D}X^\mu e^{-S_\sigma[X^\mu, h_{ab}]} \prod_{i=1}^n \int d^2\sigma \sqrt{h} V_i .$$

In this section, we will examine the degrees of freedom in the metric integration of the string amplitude after $(\text{Diff} \times \text{Weyl})$ gauge fixing. Specifically, for a closed oriented Riemann surface $\Sigma_{g,n}$, the dimension of the integration is

$$\dim \frac{[\mathcal{D}h_{ab}]_{g,n}}{(\text{Diff} \times \text{Weyl})} = 6g - 6 + 2n .$$

To understand the dimension of integration in the string amplitude, it is helpful to consider the term $6g$. Even after gauge fixing (e.g., fixing the metric locally), there is still freedom to change the "shape" of the world-sheet, known as metric moduli. If we increase the number of genus g , we can do so by attaching a cylinder to a Riemann surface Σ_g , as described in §4. To do this, we must specify the locations of the cylinder's endpoints and its shape. The endpoints are denoted by two complex points p_1 and p_2 , and around these points, we introduce complex coordinates z_1 and z_2 . We also impose the condition $z_1 z_2 = c$, where c is a constant whose phase specifies the twist and magnitude specifies the length (see Figure 9). In total, there are an additional six parameters that describe the shape of the world-sheet. The term $2n$ comes from the integration of the vertex operators, which is necessary for the vertex operator to be $\text{Diff} \times \text{Weyl}$ -invariant.

6.1 Teichmüller space and moduli space of Riemann surfaces

Let us briefly introduce Teichmüller space \mathcal{T}_g and moduli space \mathcal{M}_g of Riemann surfaces of genus g . For the sake of brevity, we focus on an oriented closed Riemann surface without punctures. It is known that every Riemann surface admits a Riemannian metric of constant curvature:

positive A two-sphere S^2 ($g = 0$) with a fixed radius in \mathbb{R}^3 has positive constant curvature with respect to the induced metric from the standard metric of \mathbb{R}^3

zero A two-torus T^2 ($g = 1$) constructed as a quotient space $T^2 = \mathbb{C}/\Gamma$ has zero curvature where Γ is a two-dimensional lattice.

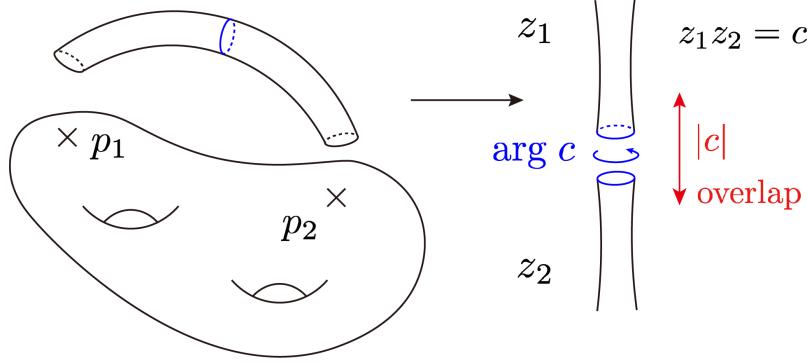


Figure 9: Attaching a pair of cylinders to Σ_g , and patching the pair of cylinders by $z_1z_2 = c$.

negative A Riemann surface of genus $g > 1$ can be constructed from a $4g$ -polygon with geodesic arcs and each angle $2\pi/4g$ in the Poincaré disc \mathbb{D} where the edges are identified via relation

$$a_1b_1a_1^{-1}b_1^{-1}a_2b_2a_2^{-1}b_2^{-1}\cdots a_gb_ga_g^{-1}b_g^{-1}=1.$$

(See Figure 10.) Hence, it admits a negative constant curvature.

A Riemann surface is a complex one-dimensional manifold. In short, the moduli space of Riemann surfaces classifies Riemann surfaces upto isometries while Teichmüller space classify Riemann surfaces upto isometries isotopic to the identity map. We will see its meaning below.

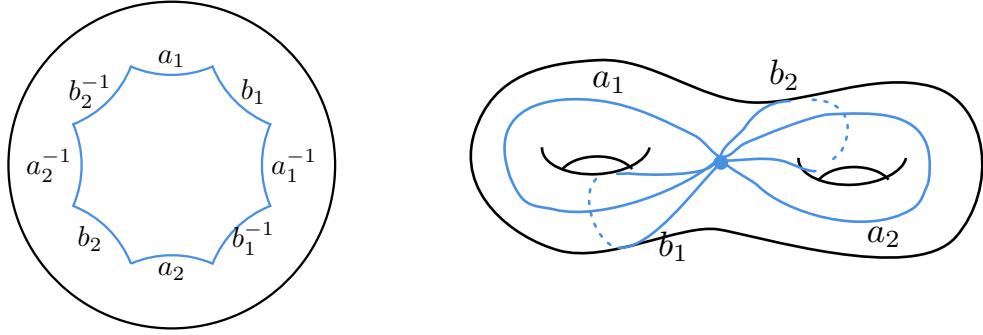


Figure 10: A $4g$ -polygon with geodesic arcs and each angle $2\pi/4g$ in the Poincaré disc \mathbb{D} ($g = 2$)

Let us first consider the case of genus one. As mentioned, a torus is constructed from a quotient space \mathbb{C}_w/Γ_τ . As far as “shape” is concerned, we can set one edge to $1 \in \mathbb{C}$ without loss of generality. Then, an identification is given by

$$w \cong w + 2\pi(m + n\tau), \quad (m, n) \in \mathbb{Z} \times \mathbb{Z}, \quad (6.1)$$

where $\tau = \tau_1 + i\tau_2$ with $\tau_2 > 0$. In this coordinate, the metric is flat $ds^2 = dwd\bar{w}$ so that w is a conformally-flat local coordinate. Another way to see “shape” of a torus is

$$(\sigma^1, \sigma^2) \cong (\sigma^1, \sigma^2) + 2\pi(m, n)$$

with a Riemannian metric

$$ds^2 = |d\sigma^1 + \tau d\sigma^2|^2. \quad (6.2)$$

In fact, the holomorphic coordinate w and (σ^1, σ^2) are related by

$$w = \sigma^1 + \tau \sigma^2, \quad (6.3)$$

so that τ encodes the information about holomorphicity. Thus, the parameter τ is called a complex structure that describes the “shape” of the torus, roughly speaking. Since (6.2) is invariant under $\tau_2 \rightarrow -\tau_2$, the space of complex structure τ is therefore the upper half place \mathbb{H} , which can be identified with the Teichmüller space $\mathcal{T}_{g=1}$ of a torus.

There are global transformations that leave “shape” as it is. The lattice Γ_τ is invariant under transformations

$$T : \tau \rightarrow \tau + 1, \quad S : \tau \rightarrow -\frac{1}{\tau}. \quad (6.4)$$

These transformations T, S generate the **modular transformation**

$$\tau \rightarrow \frac{a\tau + b}{c\tau + d}, \quad \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{PSL}(2, \mathbb{Z}), \quad (6.5)$$

where $\mathrm{PSL}(2, \mathbb{Z}) = \mathrm{SL}(2, \mathbb{Z})/\{\pm\}$. Hence, the modular transformations do not change the “shape” of a torus, but it acts on the Teichmüller space \mathbb{H} as in 11. As a result, the moduli space of complex structures of a torus is

$$\mathcal{M}_{g=1} = \mathcal{T}_{g=1}/\mathrm{PSL}(2, \mathbb{Z}), \quad (6.6)$$

which is the shaded region in Figure 11, called the **fundamental region**.

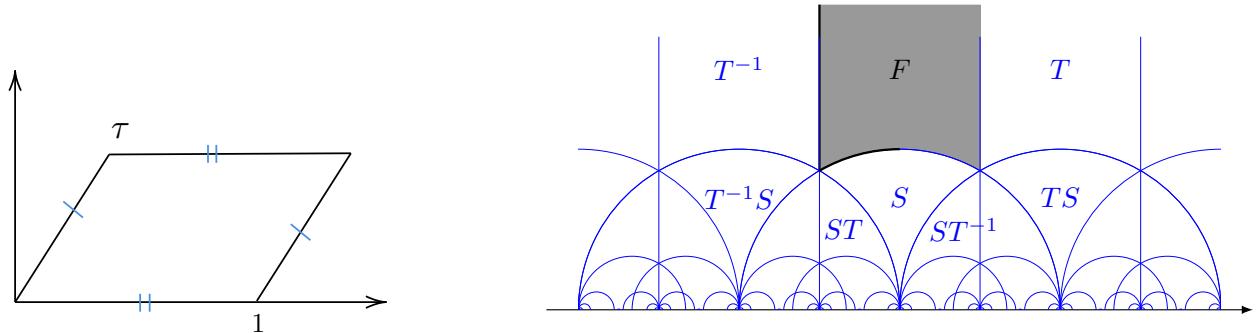


Figure 11: The Teichmüller space of a torus is the upper half-plane, and the mapping class group $\mathrm{PSL}(2, \mathbb{Z})$ acts on it. The moduli space is the fundamental region F (the shaded region).

Now let us consider a Riemann surface of genus $g > 1$. The Poincaré disk with the hyperbolic metric

$$(\mathbb{D}, ds^2 = \frac{4(dx^2 + dy^2)}{(1 - x^2 - y^2)^2})$$

has negative constant curvature, and geodesics are portions of circles that intersect the disk boundary at right angles. As briefly mentioned above, a Riemann surface of genus

$g > 1$ can be constructed by a $4g$ -polygon with geodesic arcs and each angle $2\pi/4g$ in the Poincare disk. However, not all Riemann surfaces of genus $g > 1$ with negative constant curvature (hyperbolic Riemann surfaces) are constructed in this way. The **moduli space** of Riemann surfaces indeed parametrizes hyperbolic Riemann surfaces that are not related by an isometry.

A hyperbolic Riemann surface of genus $g > 1$ admits a **pants decomposition** by cutting along $3g - 3$ simple closed curves into $2g - 2$ pants. In fact, one can take geodesic arcs for $3g - 3$ simple closed curves and measure their lengths with respect to the hyperbolic metric. The “shape” of each pants is uniquely determined by the lengths of three edges. Thus, when we construct a hyperbolic Riemann surface by gluing pants, the “shape” will be determined by twist angles at gluing along the simple closed curves. These data will determine how to construct a hyperbolic Riemann surface from $2g - 2$ pants, and the collection of the length coordinates l_i and the twist coordinates θ_i

$$(l_1, \dots, l_{3g-3}; \theta_1, \dots, \theta_{3g-3}) \quad (6.7)$$

are referred to as **Fenchel-Nielsen coordinates**. The Fenchel-Nielsen coordinates indeed parametrize the space of “complex structure” called **Teichmüller space** \mathcal{T}_g , and we have the bijection

$$\mathcal{T}_g \cong \mathbb{R}_+^{3g-3} \times \mathbb{R}^{3g-3}$$

In other words, \mathcal{T}_g can be identified with an open disk in \mathbb{R}^{6g-6} . Moreover, the Teichmüller space is endowed with a complex structure and Kähler form

$$\omega = \sum_{i=1}^{3g-3} dl_i \wedge d\theta_i, \quad (6.8)$$

referred to as the **Weil-Petersson form** so that it can be considered as a Kähler manifold.

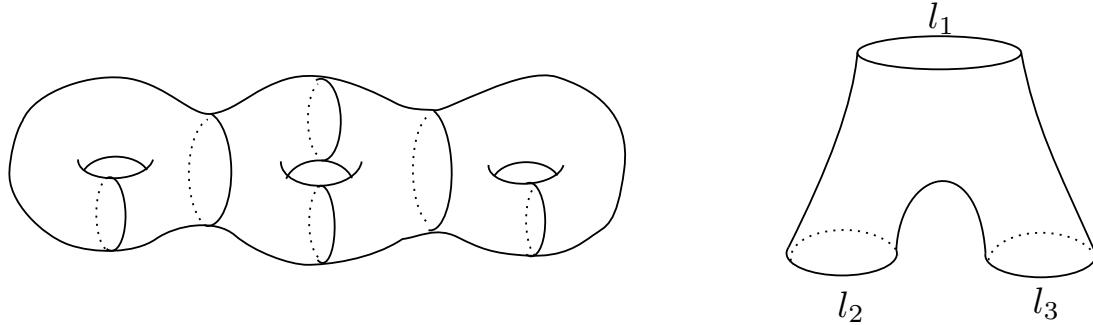


Figure 12: A pants decomposition of Riemann surface

The relation between the Teichmüller space \mathcal{T}_g and the moduli space \mathcal{M}_g is given by the mapping class group MCG_g . The Teichmüller space classifies hyperbolic Riemann surface upto isometries isotopic to the identity map. However, there are isometries that are not isotopic to the identity map. (In physics, they are sometimes called **large gauge transformations**.) The **mapping class group** appears as the quotient of orientation-preserving diffeomorphisms by diffeomorphisms isotopic to the identity map

$$\text{MCG}_g = \frac{\text{Diff}_+(\Sigma_g)}{\text{Diff}_0(\Sigma_g)}.$$

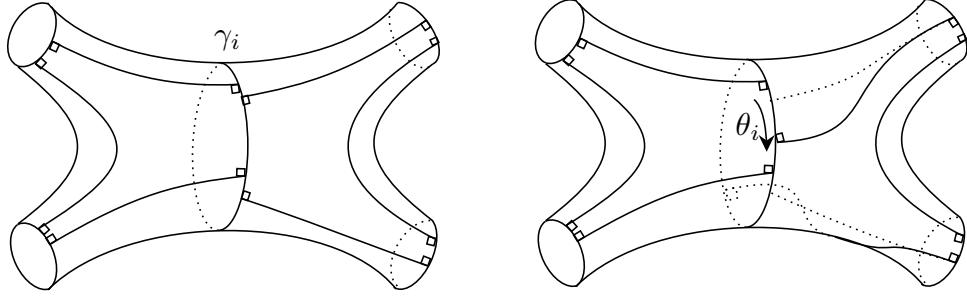


Figure 13: twist coordinates

Although this definition looks horrendous, the theorem of Lickorish states that the mapping class group MCG_g of Riemann surfaces of genus g is a discrete group generated by Dehn twists along $3g - 1$ simple closed curves shown in Figure 15. A Dehn twist is an isometry that generates a 2π -twist along a simple closed curve C of a hyperbolic Riemann surface Σ_g . Consequently, the moduli space of Riemann surfaces of genus g can be identified as the quotient space

$$\mathcal{M}_g = \mathcal{T}_g / \text{MCG}_g .$$

In the case of a torus, $\mathcal{T}_{g=1}$ is the upper half-plane and MCG_g is $\text{PSL}(2, \mathbb{Z})$, so that $\mathcal{M}_{g=1}$ is the fundamental region.

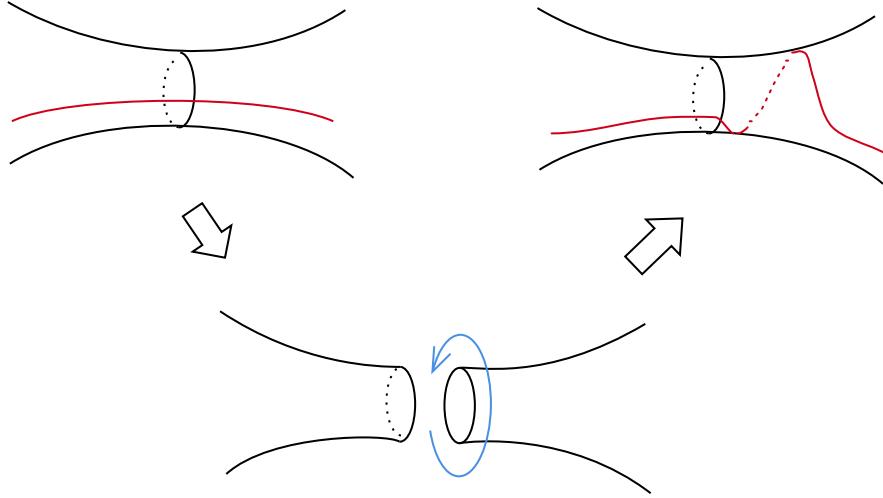


Figure 14: The Dehn twist along a simple closed curve

We can also discuss \mathcal{T}_g and \mathcal{M}_g from the viewpoint of complex analytics very briefly. More elaborate explanations can be found in [Nei87] for string theory. A Riemann surface Σ is defined by a local coordinate system $\{(U_j, z_j)\}_{j \in J}$ that are patched by biholomorphic mappings

$$z_k \circ z_j^{-1} : z_j(U_j \cap U_k) \rightarrow z_k(U_j \cap U_k) .$$

A local coordinate system defines a complex structure on Σ . Given Riemann surfaces R and S , a biholomorphic map $f : R \rightarrow S$ is a holomorphic map $f : R \rightarrow S$ which has the holomorphic inverse mapping $f^{-1} : S \rightarrow R$. The Teichmüller space is the space

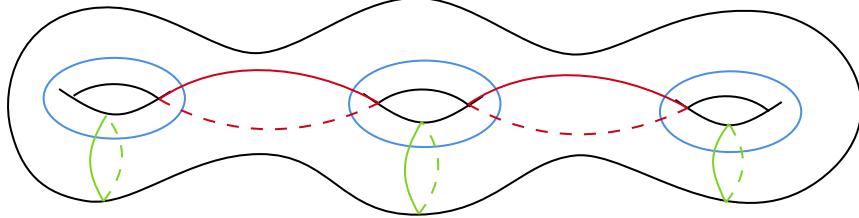


Figure 15: Choice of $3g - 1$ cycles for Dehn twists. ($g = 3$ here)

of complex structures modulo biholomorphic equivalence. One can reconstruct a complex structure from a real two-dimensional oriented closed manifold with a Riemannian metric, which can be expressed on a local coordinate (U, σ^a) as

$$ds^2 = h_{ab} d\sigma^a d\sigma^b.$$

Writing $z = \sigma^1 + i\sigma^2$, the metric takes the form

$$ds^2 = \lambda |dz + \mu d\bar{z}|^2. \quad (6.9)$$

where

$$\lambda = \frac{1}{4} \left(h_{11} + h_{22} + 2\sqrt{h_{11}h_{22} - h_{12}^2} \right), \quad \mu = \frac{h_{11} - h_{22} + 2ih_{12}}{h_{11} + h_{22} + 2\sqrt{h_{11}h_{22} - h_{12}^2}}.$$

Moreover, we can choose a conformally-flat coordinate w such that

$$\rho dw d\bar{w} = \rho |\partial w|^2 \left| dz + \frac{\bar{\partial} w}{\partial w} d\bar{z} \right|^2.$$

Comparing (6.9), the existence of a conformally-flat coordinate is indeed equivalent to the existence of a solution to

$$\mu = \frac{\bar{\partial} w}{\partial w}. \quad (6.10)$$

This is referred to as the **Beltrami equation**, and there is always a solution w . Moreover, $\{(U_j, w_j)\}_{j \in J}$ defines a complex structure, and we write the resulting Riemann surface Σ . For example, in the case of a torus, we define a complex structure (6.3) from a Riemannian metric (6.2). Even if we perform a Weyl transformation on the metric, the resulting Riemann surface Σ' is biholomorphically equivalent to the Riemann surface Σ constructed from the original metric. Therefore, the Teichmüller space can be identified with

$$\mathcal{T}_g = \frac{\text{Met}}{\text{Weyl} \times \text{Diff}_0}, \quad (6.11)$$

where Met is the space of metrics on real two-dimensional Riemannian manifolds of genus g .

In fact, $\mu = \mu_z^z$ in (6.10) can be considered as a $(-1, 1)$ -form, called **Beltrami differential**, which encodes the space of metrics up to Weyl transformations. However, there are gauge equivalent configurations due to diffeomorphisms. To distinguish non-trivial

variation in the space $L_{(-1,1)}^\infty(\Sigma_g)$ of Beltrami differential, Teichmüller has introduced a pairing (\cdot, \cdot) with a **holomorphic quadratic differential** (or $(2,0)$ -form) $\varphi_{zz}(z)dz \otimes dz$:

$$(\mu, \varphi) \equiv \int_{\Sigma_g} \mu \bar{z}^z \varphi_{zz} dz \wedge d\bar{z}. \quad (6.12)$$

Then, Teichmüller's Lemma states that the subspace of Beltrami differentials for a trivial variation is given by

$$\mathcal{N} = \{\mu \mid (\mu, \varphi) = 0 \text{ for } \forall \varphi \in H^0(\Sigma_g, K^{\otimes 2})\}.$$

Hence, the holomorphic tangent space of the Teichmüller space at Σ can be identified with

$$T_{\Sigma_g}^{(1,0)} \mathcal{T}_g \cong L_{(-1,1)}^\infty(\Sigma_g)/\mathcal{N}, \quad (6.13)$$

and the holomorphic cotangent space with the space of holomorphic quadratic differentials

$$T_{\Sigma_g}^{*(1,0)} \mathcal{T}_g \cong H^0(\Sigma_g, K^{\otimes 2}) \quad (6.14)$$

Note that K is the cotangent bundle on Σ_g , and it has degree $2g - 2$ so that the Riemann-Roch formula tells us

$$\dim H^0(\Sigma_g, K^{\otimes 2}) = \deg(K^{\otimes 2}) - g + 1 = 3g - 3.$$

The space of anti-holomorphic quadratic differentials has the same dimension.

6.2 Gauge fixing and string amplitudes

Conformal Killing vectors & Moduli space of metrics

The gauge transformation of a world-sheet metric under diffeomorphism and Weyl transformation is given in (5.4). Moreover, conformal Killing vectors ϵ

$$P \cdot \epsilon = 0. \quad (6.15)$$

does not change the metric infinitesimally. The existence of conformal killing vectors means that there are zero modes of the ghost c -field as pointed out in (5.7). Thus, we need extra care even after introducing the bc ghost. In addition, the physical inequivalent change of a metric is perpendicular to (5.4) as in Figure 8. Writing the change $\delta^\perp h_{ab}$, we have

$$\begin{aligned} 0 &= \int d^2\sigma \sqrt{h} \delta^\perp h_{ab} \left[(P \cdot \epsilon)^{ab} + 2\tilde{\omega} h^{ab} \right] \\ &= \int d^2\sigma \sqrt{h} \left[(P^T \cdot \delta^\perp h)_a^a \epsilon^a + 2\tilde{\omega} h^{ab} \delta^\perp h_{ab} \right]. \end{aligned}$$

To satisfy the orthogonality for arbitrary ϵ and ω , it is required that

$$h^{ab} \delta^\perp h_{ab} = 0, \quad (P^T \cdot \delta^\perp h)_a^a =: -\nabla^b \delta^\perp h_{ba} = 0.$$

Therefore, solutions to this equation are equivalent to zero modes of the ghost b -field (5.8). Since these zero modes are absent in the action, we need to insert appropriate zero

modes to derive a non-trivial amplitude (because $\int db \cdot 1 = 0 = \int dc \cdot 1$). In a conformal gauge, equations for conformal Killing vectors and metric moduli become

$$\begin{aligned}\partial\bar{\epsilon} &= \bar{\partial}\epsilon = 0, \\ \delta\delta^\perp h_{\bar{z}\bar{z}} &= \bar{\partial}\delta^\perp h_{zz} = 0.\end{aligned}$$

Hence, variations of the metric moduli correspond to holomorphic quadratic differentials, which can be identified with the holomorphic cotangent space (6.14) of the Teichmüller space at Σ .

Let us look at some examples. For a two-sphere, we have the stereographic projections, and the two local coordinates are related by $w = 1/z$, yielding

$$\begin{aligned}\delta w &= \frac{\partial w}{\partial z} \delta z = -z^{-2} \delta z \\ \delta h_{ww} &= \left(\frac{\partial w}{\partial z}\right)^{-2} \delta h_{zz} = z^4 \delta h_{zz}\end{aligned}$$

CKV δz , is holomorphic at $w = 0$ if it grows no more rapidly than z^2 as $z \rightarrow \infty$. On the other hand, there is no moduli for metric

$$\begin{aligned}\delta^\perp h_{zz} &= \delta^\perp h_{\bar{z}\bar{z}} = 0, \\ \epsilon &= a_0 + a_1 z + a_2 z^2, \\ \bar{\epsilon} &= a_0^* + a_1^* \bar{z} + a_2^* \bar{z}^2.\end{aligned}$$

Therefore, there are 6 CKVs and no modulus.

On the torus, the only holomorphic doubly periodic functions are the constants, so there are two real moduli and two real CKVs.

$$\delta^\perp h_{zz} = a, \quad \epsilon = b$$

where b represents translations along two directions.

As we have seen, the dimension of the metric moduli is $6g - 6$ for a Riemann surface of genus $g > 1$. However, there is no CKV. Thus, if we write the dimension of the conformal Killing group by μ and the dimension of the metric moduli by ν , we summarize the results in Table 1. Thus, the dimension of the Techmüller space is given by $\dim \mathcal{T}_g = \nu - \mu = 6g - 6$.

	$g = 0$	$g = 1$	$g \geq 2$
μ	6	2	0
ν	0	2	$6g - 6$

Table 1: The number of zero modes of b and c .

Ghost number anomaly

We can derive the conclusion that we need to insert the ghost fields in correlation functions to have non-trivial results by considering the ghost number anomaly. The ghost action in (5.11) is invariant under the transformation

$$\delta_g c = \epsilon c, \quad \delta_g b = -\epsilon b.$$

The corresponding Noether current is a ghost number current $j = cb$ under which we can define the ghost number $[c] = 1$ and $[b] = -1$. At a flat world-sheet, the current satisfies the conservation law $\bar{\partial}j = 0$. However, it has an anomaly in a curved world-sheet proportional to the scalar curvature like in (4.6)

$$\nabla^z j_z = \kappa \cdot R^{(2)}. \quad (6.16)$$

Let us determine the coefficient κ . The OPE of the energy-momentum tensor and the ghost number current leads to

$$\begin{aligned} T(z)j(w) &= -:(2b\partial c + \partial bc)(z)::cb(w): \\ &\sim \frac{-3}{(z-w)^3} + \frac{j(w)}{(z-w)^2} + \frac{\partial j(w)}{z-w}. \end{aligned}$$

Its infinitesimal version is

$$\delta j = \epsilon \partial j + \partial \epsilon j - \frac{3}{2} \partial^2 \epsilon. \quad (6.17)$$

On the other hand, in the conformal gauge (4.1), (6.16) can be expressed as

$$j_z = -4\kappa \partial \omega + j(z),$$

where $j(z)$ is the holomorphic current. Using this expression, we find a transformation of $j(z)$ as

$$j(z) = \frac{\partial \tilde{z}}{\partial z} \tilde{j}(\tilde{z}) + 2\kappa \frac{\partial}{\partial z} \ln \frac{\partial \tilde{z}}{\partial z},$$

which yields

$$\delta j(z) = \epsilon \partial j(z) + \partial \epsilon j(z) + 2\kappa \partial^2 \epsilon.$$

Comparing this with (6.17), we have $\kappa = -\frac{3}{4}$.

The zero modes of the b -field correspond to the dimension of the moduli space of Riemann surfaces, Σ_g , while the zero modes of the c -field match the number of conformal Killing vectors. Given that the zero modes of the bc -ghost field are fermionic, we must include an equivalent number of bc -ghost fields in the Grassmannian path integral to prevent it from vanishing.

The ghost number anomaly puts a constraint on the non-vanishing ghost correlation functions. Let us consider the variation of the following correlation function under the ghost number symmetry δ_g

$$\delta_g \left\langle \prod_{j=1}^{\nu/2} b(z_j) \prod_{i=1}^{\mu/2} c(z_i) \right\rangle = \frac{\mu - \nu}{2} \left\langle \prod_{j=1}^{\nu/2} b(z_j) \prod_{i=1}^{\mu/2} c(z_i) \right\rangle.$$

On the other hand, the Ward-Takahashi identity becomes

$$\begin{aligned} \delta_g \left\langle \prod_{j=1}^{\nu/2} b(z_j) \prod_{i=1}^{\mu/2} c(z_i) \right\rangle &= - \left\langle \left(\int \frac{d^2 z}{2\pi} \sqrt{h} \nabla^z j_z \right) \prod_{j=1}^{\nu/2} b(z_j) \prod_{i=1}^{\mu/2} c(z_i) \right\rangle \\ &= \left\langle \left(\int \frac{3d^2 z}{8\pi} \sqrt{h} R^{(2)} \right) \prod_{j=1}^{\nu/2} b(z_j) \prod_{i=1}^{\mu/2} c(z_i) \right\rangle \\ &= (3 - 3g) \left\langle \prod_{j=1}^{\nu/2} b(z_j) \prod_{i=1}^{\mu/2} c(z_i) \right\rangle. \end{aligned}$$

Therefore, we can conclude that $\#c - \#b = 3 - 3g$, and similarly we have $\#\bar{c} - \#\bar{b} = 3 - 3g$. This is consistent with $\mu - \nu = 6 - 6g$ as in Table 1.

General closed-string amplitude

Now let us heuristically derive the closed string amplitude. If the conformal Killing group is present, we insert the ghost zero modes $\#c + \#\bar{c} \equiv \mu$ at specified positions to fix the gauge. If enough numbers of vertex operators are present, the gauge fixing can be done by pairing $\mu/2$ vertex operators with the ghost zero modes at specific positions

$$c\bar{c}V(\hat{\sigma}_j) \quad (j = 1, \dots, \mu/2).$$

Moreover, we need to integrate over the physically inequivalent world-sheet metric, which is the moduli space of the world-sheet Riemann surface \mathcal{M}_g . This can be done by using the non-degenerate pairing with the cotangent and tangent bundle of the moduli space introduced in (6.12). Since the b and \bar{b} zero modes are sections of (anti-)holomorphic cotangent bundle of the moduli space, they are paired with (anti-)holomorphic tangent vectors as

$$(b, \partial_k \hat{h}) = \int \frac{d^2\sigma}{4\pi} \sqrt{\hat{h}} b_{ab} \frac{\partial}{\partial t_k} \hat{h}^{ab}(t), \quad (\bar{b}, \bar{\partial}_k \hat{h}) = \int \frac{d^2\sigma}{4\pi} \sqrt{\hat{h}} b_{ab} \frac{\partial}{\partial \bar{t}_k} \hat{h}^{ab}(t).$$

where t_k are coordinates of \mathcal{M}_g .

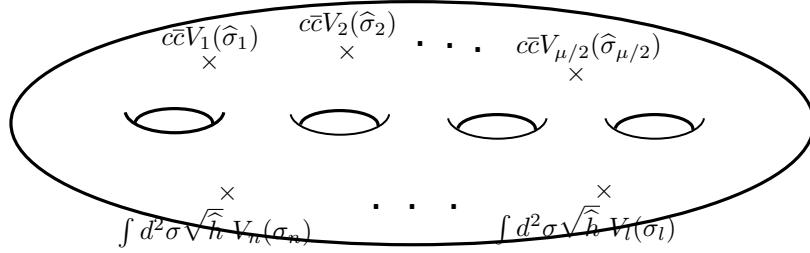


Figure 16: $\mu/2$ vertex operators are paired with $c\bar{c}$ at fixed positions and the others are integrated over the Riemann surface.

With these insertions, the general closed-string amplitude in the bosonic string theory is written as

$$A_{g,n} = \int_{\mathcal{M}_g} d^\nu t \int \mathcal{D}[c, \bar{c}, b, \bar{b}, X] e^{-S_\sigma[X, \hat{h}] - S_{\text{gh}}[b, c] - \Phi \chi(\Sigma_g)} \\ \prod_{k=1}^{\nu/2} (b, \partial_k \hat{h}) (\bar{b}, \bar{\partial}_k \hat{h}) \prod_{i=1}^{\mu/2} g_s c\bar{c} V_i(\hat{\sigma}_i) \prod_{l=\mu/2+1}^n \int d^2 \sigma_l \sqrt{\hat{h}} g_s V_l(\sigma_l), \quad (6.18)$$

where $\chi(\Sigma_g)$ is the Euler characteristics $(2 - 2g)$, and Φ is a vacuum expectation value of a dilaton that gives the string coupling $g_s = e^\Phi$. Let us now consider explicit examples and carry out computations.

6.3 Tree amplitude

The first example is the tree-level amplitude in string theory: a sphere with $n \geq 3$ punctures. Since we have $\mu = 6$ and $\nu = 0$ for a sphere, we must insert at least three

vertex operators to fix conformal Killing vectors. Considering a sphere $S^2 \cong \mathbb{C} \cup \{\infty\}$, we give the flat metric locally. Then, the amplitude formula becomes

$$\begin{aligned} A_{g=0,n} &= g_s^{-2} \int \mathcal{D}[c, \bar{c}, b, \bar{b}, X] e^{-S_\sigma[X, \hat{h}] - S_{\text{gh}}[b, c]} \prod_{i=1}^3 c\bar{c}(z_i, \bar{z}_i) g_s V_i(z_i, \bar{z}_i) \prod_{l=4}^n \int d^2 z_l g_s V_l(z_l, \bar{z}_l) \\ &= g_s^{n-2} \prod_{l=4}^n \int d^2 z_l \left\langle \prod_{j=1}^n V_j(z_j, \bar{z}_j) \right\rangle_X \left\langle c\bar{c}(z_1, \bar{z}_1) c\bar{c}(z_2, \bar{z}_2) c\bar{c}(z_3, \bar{z}_3) \right\rangle_{bc}, \end{aligned}$$

where we fix the positions of three operators $c\bar{c}(z_i, \bar{z}_i) g_s V_i(z_i, \bar{z}_i)$ for $i = 1, 2, 3$. As a result, the amplitude factorizes into the matter sector and the ghost sector.

Let us compute when all the vertex operators are Tachyons (3.36):

$$V_j(z_j, \bar{z}_j) = :e^{ik_j \cdot X(z_j, \bar{z}_j)}:$$

Then, the OPE of the Tachyon vertex operators can be computed from (3.7), which yields

$$\left\langle \prod_{j=1}^n V_j(z_j, \bar{z}_j) \right\rangle_X = C_X (2\pi)^D \delta^D(\sum k_i) \prod_{i < j}^n |z_{ij}|^{\alpha' k_i \cdot k_j}.$$

To compute the ghost sector, we perform an integral over the zero modes

$$\begin{aligned} c(z) &= c_0 + c_1 z + c_2 z^2, \\ \bar{c}(\bar{z}) &= \bar{c}_0 + \bar{c}_1 \bar{z} + \bar{c}_2 \bar{z}^2. \end{aligned}$$

Then, the integral is indeed very simple

$$\begin{aligned} \left\langle c\bar{c}(z_1, \bar{z}_1) c\bar{c}(z_2, \bar{z}_2) c\bar{c}(z_3, \bar{z}_3) \right\rangle_{bc} &= C_{bc} \int \prod_{i=0}^2 d\bar{c}_i d c_i c(z_1) \bar{c}(\bar{z}_1) c(z_2) \bar{c}(\bar{z}_2) c(z_3) \bar{c}(\bar{z}_3) \\ &= C_{bc} \det \begin{vmatrix} 1 & 1 & 1 \\ z_1 & z_2 & z_3 \\ z_1^2 & z_2^2 & z_3^2 \end{vmatrix} \det \begin{vmatrix} 1 & 1 & 1 \\ \bar{z}_4 & \bar{z}_5 & \bar{z}_6 \\ \bar{z}_4^2 & \bar{z}_5^2 & \bar{z}_6^2 \end{vmatrix} \\ &= C_{bc} |z_{12}|^2 |z_{23}|^2 |z_{31}|^2, \end{aligned}$$

where C_{bc} is the normalization constant independent of positions.

Shapiro-Virasoro amplitude

Let us explicitly perform the four-point amplitude

$$A_{0,4} = g_s^2 C_{4\text{pt}} (2\pi)^D \delta^D(\sum k_i) \int \prod_i d^2 z_i \prod_{i < j}^4 |z_{ij}|^{\alpha' k_i \cdot k_j} \prod_{i < j}^3 |z_{ij}|^2.$$

To fix the conformal Killing vectors, we can set (z_1, z_2, z_3) to $(0, 1, \infty)$. Then the expression is reduced to

$$A_{0,4} = g_s^2 C_{4\text{pt}} (2\pi)^D \delta^D(\sum k_i) B \left(-\frac{\alpha' s}{4} - 1, -\frac{\alpha' t}{4} - 1, -\frac{\alpha' u}{4} - 1 \right), \quad (6.19)$$

where

$$B(a, b, c) = \int d^2z |z|^{2a-2} |1-z|^{2b-2} = \pi \frac{\Gamma(a)\Gamma(b)\Gamma(c)}{\Gamma(a+b)\Gamma(b+c)\Gamma(c+a)} \quad (a+b+c=1). \quad (6.20)$$

Here we use the Mandelstam variables

$$\begin{aligned} s &= -k_{1+2}^2 = -k_{3+4}^2 = -2k_1 \cdot k_2 - \frac{8}{\alpha'}, \\ t &= -k_{1+3}^2 = -k_{2+4}^2 = -2k_1 \cdot k_3 - \frac{8}{\alpha'}, \\ u &= -k_{1+4}^2 = -k_{2+3}^2 = -2k_1 \cdot k_4 - \frac{8}{\alpha'}, \end{aligned}$$

with $k_{i+j}^2 = (k_i + k_j)^2$, satisfying the constraint

$$s + t + u = -\frac{16}{\alpha'}. \quad (6.21)$$

In (6.19), the Mandelstam variables s, t, u are symmetric under permutations. This is called the **Shapiro-Virasoro amplitude**. It is easy to see that Shapiro-Virasoro amplitude has a permutation symmetry among s, t, u . For the derivation of the B function, the reader can refer to [GSW87, §7(Vol.1 p.386 and p.373)].

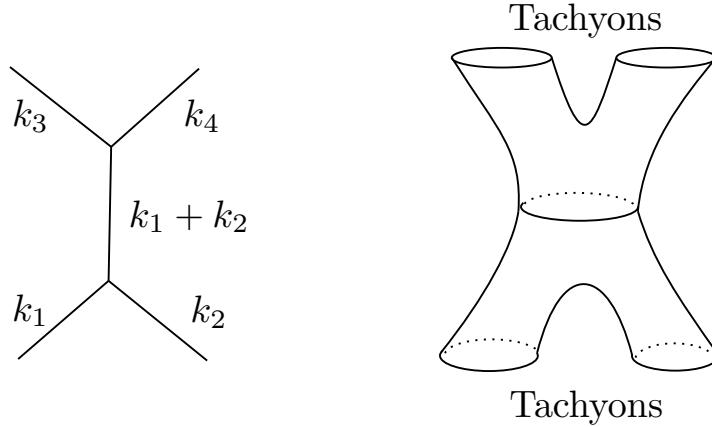


Figure 17: The s-channel of the Shapiro-Virasoro amplitude.

The Γ -function has a pole at non-positive integers. This shows that the amplitude in (6.19) has poles in the s -channel:

$$-\frac{\alpha' s}{4} - 1 \in \mathbb{Z}_{\leq 0}, \quad s = -k_{1+2}^2 = M^2 = \frac{4}{\alpha'}(n-1) \quad (n \in \mathbb{Z}_{\geq 0}),$$

where m is the mass of a propagating state as in Figure 17. Since the mass spectra are the same as in (2.18), this implies that the propagating states are tachyon, graviton, B -field, etc. The four-point amplitude can be divided into two three-point amplitudes as in Figure 17. At the s -channel tachyon pole, it becomes

$$\begin{aligned} A_{0,3} &= a_3 (2\pi)^D \delta^D(\sum k_i), \quad a_3 \simeq C g_s, \\ A_{0,4} &= a_4 (2\pi)^D \delta^D(\sum k_i), \quad a_4 \simeq -\frac{4\pi}{\alpha'} \frac{1}{s + \frac{4}{\alpha'}} C g_s^2. \end{aligned}$$

We can fix the overall constant from unitarity

$$a_4 = \frac{(a_3)^2}{s + \frac{4}{\alpha'}} , \quad \therefore \quad C = -\frac{4\pi}{\alpha'} .$$

Although we have considered a closed-string tree-level amplitude here, the result for open-string amplitude, so-called the **Veneziano amplitude** [Ven68], has historically proceeded. The discovery of the Veneziano amplitude is widely acknowledged as the starting point for the developments leading to string theory. The very first examples of string amplitudes by Veneziano [Ven68], Virasoro [Vir69], Shapiro [Sha70], as well as Koba and Nielsen [KN69a, KN69b] appeared long before the formulation of string theory. The early history of string theory before the superstring revolution in 1984 is recounted in [CCCDV12].

6.4 One-loop amplitude

Let us consider a torus amplitude which is analogous to one-loop amplitude so that it encodes quantum corrections. First, we will use the cylindrical holomorphic coordinate $w = t + i\sigma$ in Figure 5 with periodicity (6.1).

As seen in §6.2, there are two constant conformal Killing vectors and metric moduli. Note that an integral (with an appropriate normalization) over a torus automatically picks up zero modes

$$c(0) = \int \frac{d^2 w}{4\pi^2 \tau_2} c(w)$$

(The other zero modes $\bar{c}(0), b(0), \bar{b}(0)$ are similarly obtained.)

Let us consider an infinitesimal deformation of the flat metric

$$ds^2 = dw d\bar{w} \rightarrow d(w + \epsilon \bar{w}) d(\bar{w} + \bar{\epsilon} w) ,$$

where $\delta h_{ww} = \bar{\epsilon}$ and $h_{\bar{w}\bar{w}} = \epsilon$. One can set a new conformally-flat coordinate is $\tilde{w} = w + \epsilon \bar{w}$. This changes the period of a torus to

$$\tilde{w} = \tilde{w} + (2\pi(1 + \epsilon), 2\pi(\tau + \epsilon\bar{\tau})) .$$

Consequently, the complex modulus becomes

$$\tilde{\tau} = \frac{\tau + \epsilon\bar{\tau}}{1 + \epsilon} \simeq \tau - 2i\epsilon\tau_2$$

so that $\delta\tau = -2i\epsilon\tau_2$. Thus, a tangent vector of the metric moduli is

$$\partial_\tau h_{\bar{w}\bar{w}} = \frac{\delta h_{\bar{w}\bar{w}}}{\delta\tau} = \frac{i}{2\tau_2} ,$$

so that we have

$$(b, \partial_\tau h) = \int \frac{d^2 w}{2\pi} \sqrt{h} b_{ww}(w) h^{w\bar{w}} h^{w\bar{w}} \frac{\partial}{\partial\tau} h_{\bar{w}\bar{w}}(\tau) = 2\pi i b_{ww}(0) .$$

The torus amplitude is expressed as an integral over the fundamental region F in Figure 11 with bc zero modes insertion

$$A_{1,n} = g_s^n \frac{1}{2} \int_F d^2\tau \left\langle (b, \partial_\tau h)(\bar{b}, \partial_{\bar{\tau}} h) c(0) \bar{c}(0) V_1(0) \prod_{j=2}^n \int d^2 w_j \sqrt{h} V_j(w_j, \bar{w}_j) \right\rangle .$$

where $\frac{1}{2}$ is due to the symmetry $w \rightarrow -w$. This \mathbb{Z}_2 symmetry is a discrete subgroup of CKG. Since the zero mode on a torus can be obtained by integrating an operator out, the expression can be manipulated into

$$A_{1,n} = g_s^n \frac{1}{2} \int_F \frac{d^2\tau}{\tau_2} \langle b(0)\bar{b}(0)c(0)\bar{c}(0) \rangle_{bc} \left\langle \prod_{j=1}^n \int d^2w_j \sqrt{h} V_j(w_j, \bar{w}_j) \right\rangle_X.$$

Torus partition function for bosonic string

Let us explicitly compute the amplitude without vertex operators ($n = 0$):

$$A_{1,0} = \frac{1}{2} \int_F \frac{d^2\tau}{\tau_2} \langle b(0)\bar{b}(0)c(0)\bar{c}(0) \rangle_{bc} \langle 1 \rangle_X.$$

The Feynman path integral over a torus with complex structure τ can be written in terms of the Hamiltonian formalism

$$\langle 1 \rangle_X = \text{Tr} \exp [2\pi i (\tau_1 P + i\tau_2 H)] = \text{Tr} \left[q^{L_0^X - \frac{c^X}{24}} \bar{q}^{\bar{L}_0^X - \frac{c^X}{24}} \right],$$

where $q = e^{2\pi i \tau}$. Note that the space and time translations are generated by the zero modes of the Virasoro generator (2.15)

$$P = L_0^X - \bar{L}_0^X, \quad H = L_0^X + \bar{L}_0^X - \frac{c^X}{12}.$$

The trace is taken over the Hilbert space (2.14). Writing the holomorphic Fock space with a fixed target coordinate as

$$|n_1, n_2, n_3, \dots; k\rangle = \alpha_{-1}^{n_1} \alpha_{-2}^{n_2} \alpha_{-3}^{n_3} \dots |0; k\rangle, \quad \text{with } n_i \in \mathbb{Z}_{\geq 0}, \quad (6.22)$$

its generating function can be written as

$$\begin{aligned} \text{Tr}_{\mathcal{H}^X}(q^{L_0^X - \frac{c^X}{24}}) &= q^{-\frac{1}{24}} \sum_{n_1=0}^{\infty} \sum_{n_2=0}^{\infty} \dots \langle k; n_1, n_2, \dots | q^{L_0^X} | n_1, n_2, \dots; k \rangle \\ &= q^{\frac{\alpha'}{4}k^2 - \frac{1}{24}} \prod_{n=1}^{\infty} \frac{1}{1-q^n} = q^{\frac{\alpha'}{4}k^2} \frac{1}{\eta(\tau)}, \end{aligned}$$

where $\eta(\tau)$ is the **Dedekind eta-function**. Hence, we have the bosonic torus partition function

$$\begin{aligned} \langle 1 \rangle_X &= \int \frac{d^D x d^D k}{(2\pi)^D} \exp \left(-4\pi\tau_2 \frac{\alpha'}{4} k^2 \right) \left| \frac{q^{-\frac{1}{24}}}{\prod_{n \geq 1} (1-q^n)} \right|^{2D} \\ &= i \frac{V_D}{(2\pi\ell_s)^D} (\tau_2)^{-D/2} |\eta(\tau)|^{-2D}, \end{aligned}$$

where i in the second line is from Wick-rotation of spacetime momentum $k^0 \rightarrow ik_E^0$.

Similarly, we can write the ghost partition function over a torus with complex structure τ as

$$\langle b(0)\bar{b}(0)c(0)\bar{c}(0) \rangle_{bc} = \text{Tr} \left[(-1)^F b_0 \bar{b}_0 c_0 \bar{c}_0 q^{L_0^g - \frac{c^g}{24}} \bar{q}^{\bar{L}_0^g - \frac{c^g}{24}} \right]$$

with $c^g = -26$, and the zero mode of the Virasoro generator (5.15) is

$$L_0^g = - \sum_{n \in \mathbb{Z}} n : b_n c_{-n} : -1 . \quad (6.23)$$

Here F is the fermion number operator; $F = 1$ for fermion and $F = 0$ for boson. The insertion of $(-1)^F$ means that the periodic boundary condition is imposed on the ghost fields along the time circle. The trace is taken over the Hilbert space spanned by physical states that are constructed creation operators b_{-n}, c_{-m} for $n, m \in \mathbb{Z}_{>0}$ (as well as anti-holomorphic operators) on the bc ghost vacuum $|\downarrow\downarrow\rangle$

$$\text{Tr}_{\mathcal{H}^g}[(-1)^F b_0 c_0 q^{L_0^g - \frac{c^g}{24}}] = q^{-1 - \frac{c^g}{24}} \prod_{n \geq 1} (1 - q^n)^2$$

Thus, we have

$$\langle b(0)\bar{b}(0)c(0)\bar{c}(0) \rangle_{bc} = q^{-1 - \frac{c^g}{24}} \bar{q}^{-1 - \frac{c^g}{24}} \prod_{n \geq 1} (1 - q^n)^2 (1 - \bar{q}^n)^2 = |\eta(\tau)|^4 .$$

In conclusion, the torus amplitude is

$$A_{1,0} = \frac{iV_D}{(2\pi\ell_s)^D} \int_F \frac{d^2\tau}{2\tau_2} (\tau_2)^{-D/2} |\eta(\tau)|^{-2(D-2)} . \quad (6.24)$$

At the critical dimension $D = 26$, the expansion of the Dedekind eta-function

$$|\eta(\tau)|^{-48} \simeq \left| q^{-1} + 24 + \mathcal{O}(q) \right|^2 ,$$

gives the spectrum of the bosonic string theory. Furthermore, it is invariant under $\text{PSL}(2, \mathbb{Z})$. In fact, the measure $\frac{d^2\tau}{\tau_2^2}$ is invariant under $\text{PSL}(2, \mathbb{Z})$, and $\tau_2^{\frac{1}{2}} |\eta(q)|^2$ is also invariant under $\text{PSL}(2, \mathbb{Z})$ (see §7.4).

Note that the limit $\tau_2 \rightarrow 0$ describes the ultraviolet (UV) regime since the Euclidean time becomes small. The integral over the fundamental domain F avoids this UV region. Consequently, there is no UV divergence in the one-loop amplitude of the bosonic string theory thanks to the modular invariance of a torus.

7 Superstring theories

We have seen so far that bosonic strings suffer from two major problems:

- Their spectrum always contains a tachyon. In that respect, their vacuum is unstable.
- They do not contain spacetime fermions. This lack of fermionic states is in contrast to observations and makes the bosonic string unrealistic.

Both of these challenges are remedied in superstring theory. Supersymmetry is a symmetry that exchanges bosons and fermions. The world-sheet superstring theory consists of a bosonic and a fermionic sector. The bosonic sector is identical to the world-sheet theory of the bosonic string. Therefore, we can view our efforts up to now as a preliminary study of one-half of the superstring theory. In fact, we will see that the presence of fermions resolves the problem of Tachyon. Moreover, we will learn that the critical dimension of superstring theory is $D = 10$.

There are five superstring theories as follows, and we will study them in this order.

Type IIA & IIB

Oriented string theories that can incorporate open strings if there are D-branes.
 IIA: Ramond ground states with opposite chirality, and Dp -branes (p even). IIB:
 Ramond ground states with the same chirality, and Dp -branes (p odd).

Type I

An open and closed unoriented string theory, including Yang-Mills degrees of freedom with $SO(32)$ gauge group, that can incorporate $D1$, $D5$, $D9$ -branes.

Heterotic $SO(32)$ & $E_8 \times E_8$

Type II right-movers & bosonic left-movers, including Yang-Mills degrees of freedom with either $SO(32)$ or $E_8 \times E_8$ gauge group.

There exist two major formulations of superstring theory. Both formulations enjoy supersymmetry on the world-sheet and in spacetime, but they differ in the following respect:

- In the **Ramond-Neveu-Schwarz (RNS) formulation**, supersymmetry is manifest on the world-sheet, but not in spacetime.
- In the **Green-Schwarz (GS) formulation** [GSW87, §5] [BBS06, §5], supersymmetry is manifest in spacetime, but not on the world-sheet.

More recently, the pure-spinor formalism [Ber04] has been developed as yet another approach to superstring. In this section, let us discuss the RNS formalism of superstring theory.

7.1 RNS formulation

With the complex coordinate convention, the action becomes

$$S^m = \frac{1}{4\pi} \int d^2z \left(\frac{2}{\alpha'} \partial X^\mu \bar{\partial} X_\mu + \psi^\mu \bar{\partial} \psi_\mu + \bar{\psi}^\mu \partial \bar{\psi}_\mu \right) \quad (7.1)$$

where the equations of motion tell us $\psi^\mu(z)$ (resp. $\bar{\psi}^\mu(\bar{z})$) is chiral (resp. anti-chiral). The action is invariant under **supersymmetric transformation** (Exercise)

$$\delta X^\mu = -\sqrt{\frac{\alpha'}{2}} (\epsilon \psi^\mu + \bar{\epsilon} \bar{\psi}^\mu), \quad \delta \psi^\mu = \sqrt{\frac{2}{\alpha'}} \epsilon \partial X^\mu, \quad \delta \bar{\psi}^\mu = \sqrt{\frac{2}{\alpha'}} \bar{\epsilon} \bar{\partial} X^\mu. \quad (7.2)$$

The Noether theorem implies that there are currents for the supersymmetry

$$T_F(z) = i \sqrt{\frac{2}{\alpha'}} \psi^\mu \partial X_\mu, \quad \bar{T}_F(z) = i \sqrt{\frac{2}{\alpha'}} \bar{\psi}^\mu \bar{\partial} X_\mu. \quad (7.3)$$

which is called **supercurrents**. Indeed, ψ^μ and $\bar{\psi}^\mu$ are primary fields of conformal dimension, $(\frac{1}{2}, 0)$ $(0, \frac{1}{2})$, respectively, and therefore their OPEs are

$$X^\mu(z, \bar{z}) X^\nu(0, 0) \sim -\sqrt{\frac{\alpha'}{2}} \eta^{\mu\nu} \ln |z|^2, \quad \psi^\mu(z) \psi^\nu(0) \sim \frac{\eta^{\mu\nu}}{z}, \quad \bar{\psi}^\mu(\bar{z}) \bar{\psi}^\nu(0) \sim \frac{\eta^{\mu\nu}}{\bar{z}}. \quad (7.4)$$

Using the OPEs, one can show the supersymmetric transformation (7.2) (exercise).

The energy-momentum tensor of the action (7.1) is

$$T_B(z) = -\frac{1}{\alpha'} \partial X^\mu \partial X_\mu - \frac{1}{2} \psi^\mu \partial \psi_\mu \quad (7.5)$$

along with their complex conjugates \bar{T}_B, \bar{T}_F . Their OPEs can be computed by using (7.4)

$$\begin{aligned} T_B(z)T_B(w) &\sim \frac{3D}{4(z-w)^4} + \frac{2T_B(w)}{(z-w)^2} + \frac{\partial_w T_B(w)}{z-w} \\ T_B(z)T_F(w) &\sim \frac{3T_F(w)}{2(z-w)^2} + \frac{\partial_w T_F(w)}{z-w} \\ T_F(z)T_F(w) &\sim \frac{D}{(z-w)^3} + \frac{2T_B(w)}{z-w}. \end{aligned} \quad (7.6)$$

and similarly for the anti-chiral part. The central charge of the theory is

$$c^m = \frac{3}{2}D \quad (7.7)$$

where each scalar and fermion contributes 1 and 1/2, respectively.

Ramond vs Neveu-Schwarz

In superstring theory, the fermionic fields on the closed string may be either periodic or anti-periodic on the circle around the string, corresponding to two different spinor bundles. It is conventional to denote these spin structures by **Ramond (R)** and **Neveu-Schwarz (NS)**, defined as follows.

$$\begin{aligned} \psi^\mu(t, \sigma + 2\pi) &= +\psi^\mu(t, \sigma) && \text{R: periodic on cylinder} \\ \psi^\mu(t, \sigma + 2\pi) &= -\psi^\mu(t, \sigma) && \text{NS: anti-periodic on cylinder} \end{aligned} \quad (7.8)$$

As in Figure 5, the mapping $z = e^{-iw}$ from the cylinder $w = \sigma + it$ to the 2-plane $z \in \mathbb{C}$ is a conformal map. Under the conformal map, the primary field ψ^μ with conformal dimension $(\frac{1}{2}, 0)$ is transformed as

$$\psi^\mu(z) = \left(\frac{dz}{dw}\right)^{-\frac{1}{2}} \psi^\mu(w) = \text{const} \times e^{-\frac{iw}{2}} \psi^\mu(w).$$

Hence the (anti-)periodicity assignments are reversed between the cylinders and the plane:

$$\begin{aligned} \psi^\mu(e^{2\pi i} z) &= -\psi^\mu(z) && \text{R: anti-periodic on plane} \\ \psi^\mu(e^{2\pi i} z) &= +\psi^\mu(z) && \text{NS: periodic on plane} \end{aligned} \quad (7.9)$$

The boundary conditions for anti-chiral fields $\bar{\psi}^\mu$ are defined in a similar fashion.

As in the bosonic string, one can decompose ψ^μ and $\bar{\psi}^\mu$ in modes

$$\psi^\mu(z) = \sum_{n \in \mathbb{Z} + \nu} \frac{\psi_n^\mu}{z^{n+1/2}}, \quad \bar{\psi}^\mu(\bar{z}) = \sum_{n \in \mathbb{Z} + \nu} \frac{\bar{\psi}_n^\mu}{\bar{z}^{n+1/2}}$$

where ν takes the values 0 (R) and $\frac{1}{2}$ (NS). The canonical quantization leads to the algebra

$$\{\psi_m^\mu, \psi_n^\nu\} = \eta^{\mu\nu} \delta_{m+n,0} \quad \{\bar{\psi}_m^\mu, \bar{\psi}_n^\nu\} = \eta^{\mu\nu} \delta_{m+n,0}. \quad (7.10)$$

The mode expansion must be carried out with care here since we must distinguish between Ramond and Neveu-Schwarz sectors.

$$T_B(z) = \sum_{n \in \mathbb{Z}} \frac{L_n^m}{z^{n+2}}, \quad T_F(z) = \sum_{r \in \mathbb{Z}+\nu} \frac{G_r^m}{z^{r+3/2}}, \quad (7.11)$$

where the generators can be written in terms of the modes (exercise)

$$\begin{aligned} L_m^m &= \frac{1}{2} \sum_{n \in \mathbb{Z}} \eta_{\mu\nu} : \alpha_{m-n}^\mu \alpha_n^\nu : + \frac{1}{4} \sum_{r \in \mathbb{Z}+\nu} (2r-m) \eta_{\mu\nu} : \psi_{m-r}^\mu \psi_r^\nu : + a^m \delta_{m,0} \\ G_r^m &= \sum_{n \in \mathbb{Z}} \eta_{\mu\nu} \alpha_n^\mu \psi_{r-n}^\nu. \end{aligned} \quad (7.12)$$

The normal ordering constant a^m can be determined like in the bosonic string theory. Each periodic boson contributes $-\frac{1}{24}$. The fermionic contributions are

$$\begin{aligned} -\frac{1}{2} \sum_{r=0}^{\infty} r &= \frac{1}{24} \quad \text{R-sector}, \\ -\frac{1}{2} \sum_{r=0}^{\infty} (r + \frac{1}{2}) &= -\frac{1}{48} \quad \text{NS-sector}. \end{aligned} \quad (7.13)$$

Including the shift $\frac{1}{24}c^m = \frac{1}{16}D$ gives

$$\begin{aligned} a^m &= \frac{1}{24}c^m + \left(-\frac{1}{24} + \frac{1}{24} \right) D = \frac{1}{16}D \quad \text{R-sector}, \\ a^m &= \frac{1}{24}c^m + \left(-\frac{1}{24} - \frac{1}{48} \right) D = 0 \quad \text{NS-sector}. \end{aligned} \quad (7.14)$$

In fact, the generators L_m^m and G_r^m form the algebra called the **$\mathcal{N} = 1$ superconformal algebra** with central charge (7.7) (exercise):

$$\begin{aligned} [L_m, L_n] &= (m-n)L_{m+n} + \frac{c}{12} (m^3 - m) \delta_{m,-n}, \\ \{G_r, G_s\} &= 2L_{r+s} + \frac{c}{12} (4r^2 - 1) \delta_{r,-s}, \\ [L_m, G_r] &= \frac{m-2r}{2} G_{m+r}. \end{aligned} \quad (7.15)$$

where we omit the upper script m .

Ghost CFT

In the bosonic string theory, we study the BRST quantization with the Faddeev-Popov bc ghost. In superstring theory, ghost fields also appear with their supersymmetric partners:

$$S^{\text{gh}} = \frac{1}{2\pi} \int d^2z (b\bar{\partial}c + \beta\bar{\partial}\gamma)$$

where b, c are fermionic and β, γ are bosonic fields. Hence, the standard method tells us the $\beta\gamma$ OPEs

$$\gamma(z)\beta(w) = -\beta(z)\gamma(w) \sim \frac{1}{z-w}$$

We have seen that the conformal dimensions of X and ψ differ by $\frac{1}{2}$. This is the same for the ghost fields. Since the b and c ghosts have conformal dimensions 2 and -1 respectively, β and γ are primary fields of conformal dimensions $(\frac{3}{2}, 0)$ and $(-\frac{1}{2}, 0)$ respectively. Hence, the form of the energy-momentum tensor and the supercurrent are

$$\begin{aligned} T_B^{\text{gh}}(z) &= :(\partial b)c: - 2\partial :bc: + :(\partial\beta)\gamma: - \frac{3}{2}\partial:\beta\gamma: \\ T_G^{\text{gh}}(z) &= (\partial\beta)c + \frac{3}{2}\beta\partial c - 2b\gamma. \end{aligned} \quad (7.16)$$

Then, the TT OPE determines the central charge of the ghost SCFT. The bc system contributes -26 to the central charge as we know, while the $\beta\gamma$ system contributes $+11$. Hence, the total central charge

$$c^{\text{tot}} = c^m + c^{\text{gh}} = \frac{3}{2}D - 26 + 11.$$

Then, we happily obtain the critical dimension $D = 10$ of superstring theory if we impose the Weyl-anomaly-free condition $c^{\text{tot}} = 0$. In the following, we assume $D = 10$.

Now, let us write the Virasoro generators of the ghost SCFT. The $\beta\gamma$ ghosts have the same boundary condition as the fermionic fields $\psi^\mu\bar{\psi}^\mu$ so that we have the mode expansions

$$\beta(z) = \sum_{r \in \mathbb{Z}+\nu} \frac{\beta_r}{z^{r+\frac{3}{2}}}, \quad \gamma(z) = \sum_{r \in \mathbb{Z}+\nu} \frac{\gamma_r}{z^{r-\frac{1}{2}}},$$

which satisfy the commutation relation

$$[\gamma_s, \beta_r] = \delta_{r,-s}.$$

Using these modes, we express the Virasoro generator of the ghost SCFT as

$$\begin{aligned} L_m^{\text{gh}} &= \sum_{n \in \mathbb{Z}} (2m-n) :b_n c_{m-n}: + \frac{1}{2} \sum_{r \in \mathbb{Z}+\nu} (m+2r) :\beta_{m-r}\gamma_r: + a^{\text{gh}}\delta_{m,0} \\ G_r^{\text{gh}} &= - \sum_{n \in \mathbb{Z}} \left[\frac{1}{2}(n+2r)\beta_{r-n}c_n + 2b_n\gamma_{r-n} \right] \end{aligned} \quad (7.17)$$

They also satisfy the $\mathcal{N} = 1$ superconformal algebra (7.15). Again, using the commutation relations of the ghost modes, one can determine the normal ordering constant

$$\begin{aligned} a^{\text{gh}} &= \frac{-15}{24} + \left(\frac{1}{12} - \frac{1}{12} \right) = -\frac{5}{8} && \text{R-sector,} \\ a^{\text{gh}} &= \frac{-15}{24} + \left(\frac{1}{12} + \frac{1}{24} \right) = -\frac{1}{2} && \text{NS-sector.} \end{aligned}$$

where two fermions and bosons contribute $\frac{1}{12}$ and $-\frac{1}{12}$, respectively, in the R sector while their contributions are $\frac{1}{12}$ and $\frac{1}{24}$, respectively, in the NS sector. Combining them with (7.14) at $D = 10$, we have the total vacuum energy

$$\begin{aligned} a^{\text{tot}} &= 0 && \text{R-sector,} \\ a^{\text{tot}} &= -\frac{1}{2} && \text{NS-sector.} \end{aligned}$$

In the R sector, the vacuum energy is zero so that the Tachyon is absent. On the other hand, Tachyon is still present in the NS sector. This will be projected out by the GSO projection, as we will see below.

7.2 Physical spectrum and the GSO Projection

Before discussing the GSO projection, let us study the fermionic spectrum generated by fermionic modes ψ_r^μ . We first consider the NS spectrum since it's simpler. Since r takes half integers, we can define the ground state of the NS sector as

$$\psi_r^\mu |0;k\rangle_{\text{NS}} = 0 \quad \text{for } r > 0.$$

When we include the ghost part of the vertex operator, it contributes to the total fermion number F , so that the matter plus ghost ground state has the odd fermion number

$$(-1)^F |0;k\rangle_{\text{NS}} = -|0;k\rangle_{\text{NS}}. \quad (7.18)$$

Because there exist the zero modes ψ_0^μ , the R sector is more subtle. In fact, the canonical commutation relation (7.10) of the zero modes satisfies the **Clifford algebra**

$$\{\sqrt{2}\psi_0^\mu, \sqrt{2}\psi_0^\nu\} = 2\eta^{\mu\nu}.$$

Therefore, we can identify them with Gamma matrices $\Gamma^\mu = \sqrt{2}\psi_0^\mu$, and the ground state of the R sector becomes the spin representation of $\text{SO}(1, D - 1)$. For mathematics of Clifford algebra, spin group, and spin representations, we refer to [Pol98, Appendix B]. However, we can heuristically understand the spin representation as follows. The following basis for this representation is often convenient:

$$\begin{aligned} \Gamma_i^\pm &= \frac{1}{\sqrt{2}} (\psi_0^{2i} \pm i\psi_0^{2i+1}) & i = 1, \dots, 4 \\ \Gamma_0^\pm &= \frac{1}{\sqrt{2}} (\psi_0^1 \pm \psi_0^0) \end{aligned} \quad (7.19)$$

In this basis, the Clifford algebra takes the form

$$\{\Gamma_i^+, \Gamma_j^-\} = \delta_{ij}, \quad \{\Gamma_i^+, \Gamma_j^+\} = 0 = \{\Gamma_i^-, \Gamma_j^-\}. \quad (7.20)$$

The Γ_i^\pm , $i = 0, \dots, 4$ act as raising and lowering operators, generating the $2^5 = 32$ Ramond ground states:

$$|s_0, s_1, s_2, s_3, s_4; k\rangle = |\mathbf{s}; k\rangle \quad (7.21)$$

where each of the s_i is $\pm \frac{1}{2}$, and where

$$\Gamma_i^- |-\frac{1}{2}, -\frac{1}{2}, -\frac{1}{2}, -\frac{1}{2}, -\frac{1}{2}; k\rangle = 0 \quad (7.22)$$

while Γ_i^+ raises s_i from $-\frac{1}{2}$ to $\frac{1}{2}$. One can further define the chirality operator

$$\Gamma_{11} = (2)^5 \psi_0^0 \psi_0^1 \psi_0^2 \cdots \psi_0^9, \quad (7.23)$$

which acts on $|\mathbf{s}\rangle$ as

$$\Gamma_{11} |\mathbf{s}; k\rangle = (-1)^F |\mathbf{s}; k\rangle = \begin{cases} +|\mathbf{s}; k\rangle & \text{even \# of } -\frac{1}{2} \\ -|\mathbf{s}; k\rangle & \text{odd \# of } -\frac{1}{2} \end{cases}.$$

Hence, the Dirac representation **32** decomposes into a **16_s** with an even number of $-\frac{1}{2}$'s and **16_c** with an odd number.

$$\mathbf{32} = \mathbf{16}_s \oplus \mathbf{16}_c. \quad (7.24)$$

Now, let us study the physical spectrum of superstring theories, using the GSO projection. We start with the open string. As in §5.2, we use the BRST quantization for the constraints on physical states. We define the BRST current and charge as

$$Q_B = \frac{1}{2\pi i} \oint dz j_B .$$

where

$$j_B = c T_B^m + \gamma T_F^m + \frac{1}{2} (c T_B^{gh} + \gamma T_F^{gh}) .$$

Using the OPE relation, it is straightforward, though tedious, to show that the L_n and G_r can be obtained by

$$\{Q_B, b_n\} = L_n, \quad [Q_B, \beta_r] = G_r.$$

In the BRST quantization, the space of physical states are isomorphic to the Q_B -cohomology (5.16). Therefore, a physical state $|\psi\rangle$ is annihilated by the positive modes

$$L_n |\psi\rangle = 0 \quad (n > 0), \quad G_r |\psi\rangle = 0 \quad (r > 0) . \quad (7.25)$$

The physical states are defined modulo

$$L_n |\chi\rangle \cong 0, \quad G_r |\chi\rangle \cong 0, \quad \text{for } n, r < 0 . \quad (7.26)$$

As in (5.22) and (5.23), we also impose the condition that the zero modes annihilate the physical state, which gives the mass-shell condition

$$L_0 |\psi\rangle = (\alpha' k^2 + N + a^{\text{tot}}) |\psi\rangle = 0. \quad (7.27)$$

In addition, in the R sector we impose

$$\beta_0 |\psi\rangle = G_0 |\psi\rangle = 0,$$

As at the end of §5.2, by taking Q_B -cohomology, the action of all the ghost modes $b, c\beta, \gamma$ are null, namely not included in the space of physical states.

Furthermore, in the RNS theory, we need to impose the **GSO (Gliozzi-Scherk-Olive) projection** in order to have an equal number of bosonic and fermionic states at each mass level.

In the NS sector, the GSO projection is just to remove states with odd fermion numbers so that the GSO projection operator on the NS sector is expressed as

$$P_{\text{GSO}}^{\text{NS}} = \frac{1 + (-1)^{F_{\text{NS}}}}{2} \quad \text{NS sector} . \quad (7.28)$$

At level 0, from the L_0 condition, the lowest state $|0; k\rangle_{\text{NS}}$ is tachyonic

$$m^2 = -k^2 = -\frac{1}{2\alpha'} \quad (7.29)$$

Fortunately, the GSO projection removes this state because it has an odd fermion number as in (7.18). The NS ground state is at level $\frac{1}{2}$, where we have a massless state with vector polarization

$$|\zeta; k\rangle_{\text{NS}} = \zeta \cdot \psi_{-\frac{1}{2}} |0; k\rangle_{\text{NS}},$$

with even fermion number so we need to keep it in the spectrum. The physical state conditions (7.25), (7.27) are

$$\begin{aligned} 0 &= L_0 |\zeta; k\rangle_{\text{NS}} = \alpha' k^2 |\zeta; k\rangle_{\text{NS}}, \\ 0 &= G_{\frac{1}{2}} |\zeta; k\rangle_{\text{NS}} = \sqrt{2\alpha'} k \cdot \zeta |0; k\rangle_{\text{NS}}, \end{aligned} \quad (7.30)$$

while the null state at this level leads

$$G_{-\frac{1}{2}} |0; k\rangle_{\text{NS}} = \sqrt{2\alpha'} k \cdot \psi_{-\frac{1}{2}} |0; k\rangle_{\text{NS}}.$$

Therefore, we have

$$m^2 = -k^2 = 0, \quad k \cdot \zeta = 0, \quad \zeta^\mu \cong \zeta^\mu + k^\mu.$$

Thus, the NS ground state is massless with degrees of freedom for 8 spacelike polarizations, forming the vector representation $\mathbf{8}_v$ of $\text{SO}(8)$.

A Ramond ground state can be expressed with spinor polarization

$$|u; k\rangle_R = u_s |\mathbf{s}; k\rangle_R.$$

The mass-shell condition (7.27) implies that the Ramond ground state is massless

$$0 = L_0 |u; k\rangle_R = \alpha' k^2 |u; k\rangle_R,$$

while the physical condition (7.25)

$$0 = G_0 |u; k\rangle_R = \sqrt{\alpha'} u_s k \cdot \Gamma_{ss'} |\mathbf{s}'; k\rangle_R$$

leads to the Dirac equation

$$u k \cdot \Gamma = 0.$$

By choosing the momentum vector $k^\mu = (k, k, 0, \dots, 0)$, this amounts to

$$\Gamma_0^+ u = 0 \quad \longrightarrow \quad s_0 = +\frac{1}{2}, \quad (7.31)$$

giving 16 degeneracies $|+, s_1, s_2, s_3, s_4\rangle$ for the physical Ramond vacuum. This is a representation $\mathbf{16}$ of $\text{SO}(8)$ which again decomposes into $\mathbf{8}_s$ with an even number of $-\frac{1}{2}$'s and $\mathbf{8}_c$ with an odd number:

$$\mathbf{16} = \mathbf{8}_s \oplus \mathbf{8}_c. \quad (7.32)$$

In the R sector, the GSO projections will pick one of these two irreducible representations, and therefore the GSO projection operators can be written as

$$P_{\text{GSO}}^{R\pm} = \frac{1 \pm (-1)^{F_R}}{2} \quad \text{R sector.} \quad (7.33)$$

where $F_R = \sum_{i=0}^4 (\frac{1}{2} + s_i)$. Indeed, the two choices $\mathbf{8}_s$ and $\mathbf{8}_c$ differ by the spacetime parity redefinition.

7.3 Torus partition functions for superstring theory

Let us see the role of the GSO projections in the torus (one-loop) partition function of superstring theory. Although the computation is very similar to that in §6.4, fermionic torus partition functions are more interesting due to boundary conditions along circles.

Since the critical dimension of superstring theory is $D = 10$, the bosonic contribution can be read off from (6.24). Consequently, the torus partition function for superstring theory can be written as follows

$$Z = \frac{iV_{10}}{(2\pi\ell_s)^{10}} \int \frac{d^2\tau}{2\tau_2^2} \frac{1}{\tau_2^4 |\eta(\tau)|^{16}} Z_F(\tau) \bar{Z}_F(\bar{\tau}),$$

where Z_F is ψ contribution and \bar{Z}_F is $\bar{\psi}$ contribution.

As usual, we will focus on the holomorphic sector Z_F , and we will sum up all the possible boundary conditions for fermions. Since the NS & R boundary conditions give rise to disconnected Fock spaces, we can introduce a relative phase $e^{i\theta}$ between them. Hence, the partition function in the Hamilton formalism is written as follows.

$$\begin{aligned} Z_F(\tau) &= Z_{NS}^+ + Z_{NS}^- + e^{i\theta} Z_R^+ + e^{i\theta} Z_R^- , \\ Z_S^\pm &= \text{Tr} [(\pm)^F e^{2\pi i \tau H_S}] , \end{aligned}$$

where S is either NS or R, and the superscript is an option for periodicity of world-sheet time t direction ($-$ is periodic). In the light-cone gauge, the “Hamiltonian”s for fermions are

$$\begin{aligned} H_{NS} &= L_0 = \sum_{r=\frac{1}{2}}^{\infty} r \psi_{-r} \cdot \psi_r - \frac{D-2}{48} , \\ H_R &= L_0 = \sum_{r=1}^{\infty} r \psi_{-r} \cdot \psi_r + \frac{D-2}{24} , \end{aligned} \tag{7.34}$$

where D should be 10 for superstring, and the light-cone gauge means that the spacetime indices i run from 2 to 9, namely

$$\psi \cdot \psi = \sum_{i=2}^9 \psi^i \psi^i.$$

Although the ground state of the NS sector is unique, the R sector has vacuum degeneracies as in (7.21) before imposing any physical condition. However, the physical state condition requires $s_0 = +\frac{1}{2}$ (7.31) so that there are $2^4 = 16$ degeneracies on which the operator $(-1)^{F_R}$ acts as

$$(-1)^{F_R} |\mathbf{s}; k\rangle = (-1)^{\sum_{i=0}^4 (\frac{1}{2} + s_i)} |\mathbf{s}; k\rangle$$

Therefore, we have

$$\sum_{\mathbf{s}} \langle \mathbf{s}; k | (-1)^{F_R} | \mathbf{s}; k \rangle = \begin{cases} 16 & \text{for + sign} \\ 0 & \text{for - sign,} \end{cases}$$

where sum over s_i ($i = 1, 2, 3, 4$) is understood. In the presence of the fermion zero modes, the index $\text{Tr}(-1)^{F_R}$ vanishes because $s_i = \pm\frac{1}{2}$ gives the opposite sign. (This is the generic feature of the index.)

Counting the eigenvalues of (7.34) in the fermionic Fock spaces with a given boundary condition, the torus partition functions are therefore written as

$$\begin{aligned} Z_{\text{NS}}^+ &= \text{Tr} \left[e^{2\pi i \tau H_{\text{NS}}} \right] = q^{-\frac{1}{6}} \prod_{n=1}^{\infty} (1 + q^{n-\frac{1}{2}})^8 = \left(\frac{\vartheta_3(\tau)}{\eta(\tau)} \right)^4, \\ Z_{\text{NS}}^- &= \text{Tr} \left[(-1)^F e^{2\pi i \tau H_{\text{NS}}} \right] = -q^{-\frac{1}{6}} \prod_{n=1}^{\infty} (1 - q^{n-\frac{1}{2}})^8 = - \left(\frac{\vartheta_4(\tau)}{\eta(\tau)} \right)^4, \\ Z_{\text{R}}^+ &= \text{Tr} \left[e^{2\pi i \tau H_{\text{R}}} \right] = 16q^{\frac{1}{3}} \prod_{n=1}^{\infty} (1 + q^n)^8 = \left(\frac{\vartheta_2(\tau)}{\eta(\tau)} \right)^4, \\ Z_{\text{R}}^- &= \text{Tr} \left[(-1)^F e^{2\pi i \tau H_{\text{R}}} \right] = 0q^{\frac{1}{3}} \prod_{n=1}^{\infty} (1 - q^n)^8 = 0 \left(\frac{\vartheta_1(\tau)}{\eta(\tau)} \right)^4 = 0, \end{aligned}$$

where $q = e^{2\pi i \tau}$, and modular functions are summarized in §7.4. Remarkably, we obtain the Jacobi theta functions $\vartheta_i(\tau)$ ($i = 1, \dots, 4$) depending on boundary conditions. Moreover, the $\text{SL}(2, \mathbb{Z})$ action on the boundary condition over a torus easily tells us their modular properties as in Figure 18. Note that the minus sign in Z_{NS}^- comes from the fact (7.18) that the NS vacuum is fermionic.

Therefore, the torus partition function for fermions is given by

$$Z_{\text{F}}(\tau) = \frac{1}{\eta^4(\tau)} \left\{ \vartheta_3(\tau))^4 - (\vartheta_4(\tau))^4 + e^{i\theta} (\vartheta_2(\tau))^4 \right\}$$

In order for $Z_{\text{F}}(\tau)$ to be modular invariant, the relative phase factor is uniquely determined as $e^{i\theta} = -1$ so that

$$Z_{\text{F}}(\tau) = \frac{1}{\eta^4(\tau)} \left\{ \vartheta_3(\tau))^4 - (\vartheta_4(\tau))^4 - (\vartheta_2(\tau))^4 \right\} = 0.$$

In fact, the partition function is zero due to supersymmetry and this is called the Jacobi-Riemann identity (7.37). The partition function is zero so that it is trivially modular invariant.

The combination derived above is indeed expressed by using the GSO projection operators (7.28) (7.33):

$$\begin{aligned} Z_{\text{F}}(\tau) &= 2 \text{Tr} \left[\frac{1 + (-1)^F}{2} e^{2\pi i \tau H_{\text{NS}}} \right] - 2 \text{Tr} \left[\frac{1 \pm (-1)^F}{2} e^{2\pi i \tau H_{\text{R}}} \right] \\ &= 2 \text{Tr} \left[P_{\text{GSO}}^{\text{NS}} e^{2\pi i \tau H_{\text{NS}}} \right] - 2 \text{Tr} \left[P_{\text{GSO}}^{\text{R}, \pm} e^{2\pi i \tau H_{\text{R}}} \right]. \end{aligned}$$

and the minus sign in the R sector indeed comes from spacetime spin-statistics. Note that a choice of the GSO projections \pm in the R sector does not affect the result due to the fermion zero modes. In other words, the GSO projection is compatible with the modular-invariance of the torus partition function.

7.4 Modular functions

It is quite amusing to see that the modular functions such as the Dedekind eta function and Jacobi theta functions defined in the 19th century naturally arise in string theory.

Here we summarize the definition and the basic properties of the modular functions describing torus partition functions. The reader may also refer to [Pol98, §7.2] and [BP09, §4.2].

The infinite product form of them are

$$\begin{aligned}\eta(\tau) &= q^{\frac{1}{24}} \prod_{n=1}^{\infty} (1 - q^n) , \\ \vartheta_2(\tau) &= 2q^{\frac{1}{8}} \prod_{n=1}^{\infty} (1 - q^n)(1 + q^n)^2 , \\ \vartheta_3(\tau) &= \prod_{n=1}^{\infty} (1 - q^n)(1 + q^{n-\frac{1}{2}})^2 , \\ \vartheta_4(\tau) &= \prod_{n=1}^{\infty} (1 - q^n)(1 - q^{n-\frac{1}{2}})^2 ,\end{aligned}\tag{7.35}$$

where $q = e^{2\pi i\tau}$. The first one is called the **Dedekind eta function**, and the others are the **Jacobi theta functions**.

Their T - and S -transformations are read off

$$\begin{aligned}\eta(\tau + 1) &= e^{i\pi/12} \eta(\tau) , & \eta(-1/\tau) &= \sqrt{-i\tau} \eta(\tau) , \\ \vartheta_2(\tau + 1) &= e^{i\pi/4} \vartheta_2(\tau) , & \vartheta_2(-1/\tau) &= \sqrt{-i\tau} \vartheta_4(\tau) \\ \vartheta_3(\tau + 1) &= \vartheta_4(\tau) , & \vartheta_3(-1/\tau) &= \sqrt{-i\tau} \vartheta_3(\tau) \\ \vartheta_4(\tau + 1) &= \vartheta_3(\tau) , & \vartheta_4(-1/\tau) &= \sqrt{-i\tau} \vartheta_2(\tau) .\end{aligned}\tag{7.36}$$

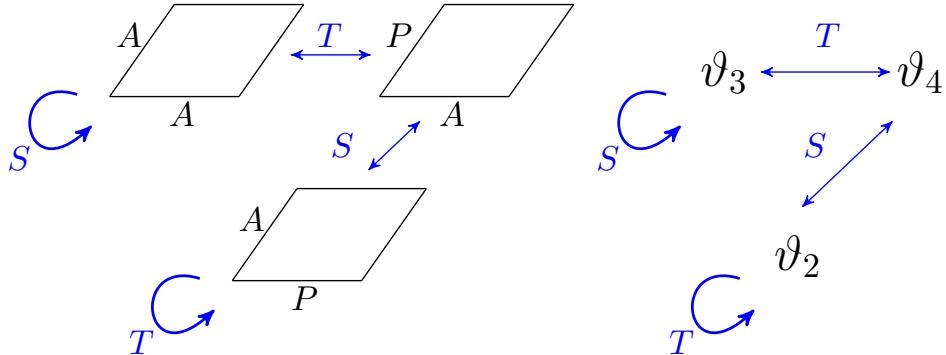


Figure 18: the T and S transformations of ϑ -functions. A and P represent anti-periodic and periodic boundary conditions of fermions.

There are two important identities. One is called the **Jacobi-Riemann identity**:

$$(\vartheta_3(\tau))^4 = (\vartheta_2(\tau))^4 + (\vartheta_4(\tau))^4 .\tag{7.37}$$

The other is **Jacobi triple product identity**:

$$\vartheta_2(\tau)\vartheta_3(\tau)\vartheta_4(\tau) = 2\eta^3(\tau) ,\tag{7.38}$$

from which the modular property of the Dedekind eta function follows.

8 Type II superstring theories

We have seen that the triality $\mathbf{8}_v, \mathbf{8}_s, \mathbf{8}_c$ of the eight-dimensional irreducible representations of $SO(8)$ appear as the massless spectrum of the RNS superstring theory after the GSO projections. One way to explain the critical dimension $D = 10$ of superstring theory is that the little group $SO(8)$ of the Lorentz group $SO(1,9)$ of the spacetime has these special eight-dimensional irreducible representations. As explained, the GSO projections (7.33) in the R sector pick one of the irreducible spinor representations $\mathbf{8}_s, \mathbf{8}_c$ of $SO(8)$. In this section, we will obtain superstring theories of two different types, depending on the sign in the GSO projection operators. The analysis of massless fields in Type II theories predicts extended objects, called **D-branes** [Pol95].

8.1 Type II superstrings

Like bosonic closed string theory, massless spectrum even in closed superstring can be determined by taking tensor products of massless states in the left- and right-moving sector. However, we now have a choice between $\mathbf{8}_s$ and $\mathbf{8}_c$ in the R sector. Hence, there are two inequivalent ways to construct superstring theory: the same (IIB) or opposite (IIA) choices on the right- and left-moving spectrum. These lead to the massless sectors

$$\begin{aligned} \text{Type IIA: } & (\mathbf{8}_v \oplus \mathbf{8}_s) \otimes (\mathbf{8}_v \oplus \mathbf{8}_c) \\ \text{Type IIB: } & (\mathbf{8}_v \oplus \mathbf{8}_s) \otimes (\mathbf{8}_v \oplus \mathbf{8}_s) \end{aligned} \quad (8.1)$$

of $SO(8)$. Although one can also choose

$$\begin{aligned} \text{Type IIA': } & (\mathbf{8}_v \oplus \mathbf{8}_c) \otimes (\mathbf{8}_v \oplus \mathbf{8}_s) \\ \text{Type IIB': } & (\mathbf{8}_v \oplus \mathbf{8}_c) \otimes (\mathbf{8}_v \oplus \mathbf{8}_c) \end{aligned} \quad (8.2)$$

they are equivalent after the spacetime parity redefinition. By the construction, Type IIB theory is chiral whereas IIA theory is not chiral. Also, the tensor products of the NS-NS and R-R sectors provide massless bosonic fields whereas those of the NS-R and R-NS sectors give massless fermionic fields.

IIA	$\mathbf{8}_v$	$\mathbf{8}_c$	IIB	$\mathbf{8}_v$	$\mathbf{8}_s$
$\mathbf{8}_v$	$\mathbf{1} \oplus \mathbf{28} \oplus \mathbf{35}$ $\phi \ B_{\mu\nu} \ G_{\mu\nu}$	$\mathbf{8}_s \oplus \mathbf{56}_c$ $\lambda^- \ \psi_m^+$	$\mathbf{8}_v$	$\mathbf{1} \oplus \mathbf{28} \oplus \mathbf{35}$ $\phi \ B_{\mu\nu} \ G_{\mu\nu}$	$\mathbf{8}_c \oplus \mathbf{56}_s$ $\lambda^+ \ \psi_m^-$
$\mathbf{8}_s$	$\mathbf{8}_c \oplus \mathbf{56}_s$ $\lambda^+ \ \psi_m^-$	$\mathbf{8}_v \oplus \mathbf{56}_t$ $C_n \ C_{nmp}$	$\mathbf{8}_s$	$\mathbf{8}_c \oplus \mathbf{56}_s$ $\lambda^+ \ \psi_m^-$	$\mathbf{1} \oplus \mathbf{28} \oplus \mathbf{35}_+$ $C \ C_{mn} \ C_{mnpq}$

Table 2: Massless fields in Type IIA and IIB theory.

Let us first look at the massless bosonic sector. In the NS-NS sector, this is the same as bosonic string theory

$$\mathbf{8}_v \otimes \mathbf{8}_v = \phi \oplus B_{\mu\nu} \oplus G_{\mu\nu} = \mathbf{1} \oplus \mathbf{28} \oplus \mathbf{35}. \quad (8.3)$$

Thus, the new ingredients come from the R-R sector, and the IIA and IIB spectra are respectively

$$\begin{aligned} \mathbf{8}_s \otimes \mathbf{8}_c &= [\mathbf{1}] \oplus [\mathbf{3}] = \mathbf{8}_v \oplus \mathbf{56}_t \\ \mathbf{8}_s \otimes \mathbf{8}_s &= [\mathbf{0}] \oplus [\mathbf{2}] \oplus [\mathbf{4}]_+ = \mathbf{1} \oplus \mathbf{28} \oplus \mathbf{35}_+. \end{aligned} \quad (8.4)$$

Here $[n]$ denotes the n -the antisymmetric representation of $\text{SO}(8)$, and we associate R-R n -form $C_{(n)}$ to it. Also, here $[4]_+$ means its R-R field strength $G_{(5)} = dC_{(4)}$ is self-dual

$$*G_{(5)} = G_{(5)} . \quad (8.5)$$

Note that the representations $[n]$ and $[8 - n]$ are related by the Hodge dual so that they are related by contraction with the 8-dimensional ϵ -tensor. As we will see next, these R-R fields are associated to D-branes in Type II theories, and $[n]$ and $[8 - n]$ are related by the electro-magnetic duality.

Let us first look at the massless fermionic sector. The tensor products of the NS-R and R-NS sectors are given by

$$\begin{aligned} \mathbf{8}_v \otimes \mathbf{8}_c &= \mathbf{8}_s \oplus \mathbf{56}_c \\ \mathbf{8}_v \otimes \mathbf{8}_s &= \mathbf{8}_c \oplus \mathbf{56}_s . \end{aligned} \quad (8.6)$$

The $\mathbf{56}_{s,c}$ correspond to gravitinos $\psi_{m,\alpha}^\pm$ that are superpartners of gravitons, and they have vector and one spinor indices by construction. As we will see in §12, they will give spacetime supersymmetry. The $\mathbf{8}_{s,c}$ are dilatino λ_α^\pm which are superpartners of the dilaton field.

8.2 Introduction to D-branes

The massless spectrum from closed strings analyzed above does not incorporate gauge fields because gauge fields arise from an open string. To incorporate open strings in Type II theories, we need to introduce D-branes. Note that closed superstring theories are consistent in themselves as we see in Heterotic string theory §11. However, only open string theory is inconsistent, and it requires closed strings as well (see Figure 19).

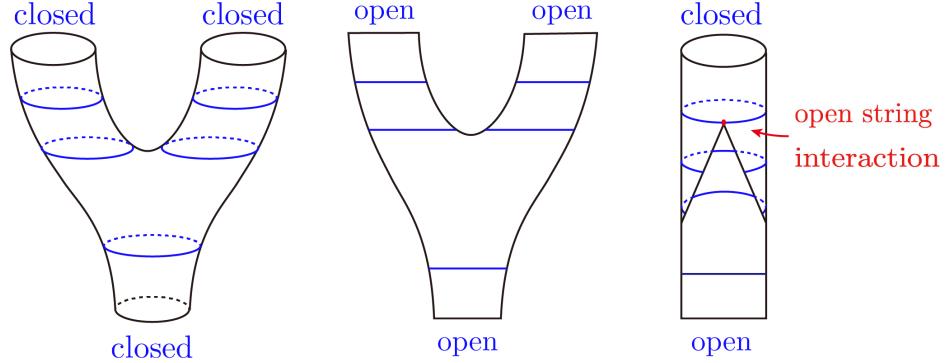


Figure 19: Open string interaction induces closed string

Boundary conditions & D-brane

As seen in (2.26), we have imposed the two types of boundary conditions for open strings. In fact, the Dirichlet boundary condition (Figure 20) does not conserve the momenta of open strings (exercise). It implies that there must be an object into which the momentum goes. We call this object a D-brane where the D stands for Dirichlet and brane

comes from the membrane. Indeed, we can impose the Neumann boundary condition to some coordinates and the Dirichlet to the other coordinates:

$$\begin{aligned} \text{Neumann condition on } X^a & \quad (a = 0, 1, \dots, p) \\ \text{Dirichlet condition on } X^I & \quad (I = p + 1, \dots, D - 1) \end{aligned}$$

The corresponding D-brane is called a **D p -brane**, which extends to a p -dimensional subspace or a $(p + 1)$ -dimensional spacetime. Now we can visualize a configuration of D-brane and string as in Figure 20. A D-brane is a dynamical object as it should receive the momentum so that it has action and interactions with string. (See §15.) On the other hand, in order to give the Dirichlet boundary condition $X^I = c^I$, a D-brane must be infinitely heavier than a string.

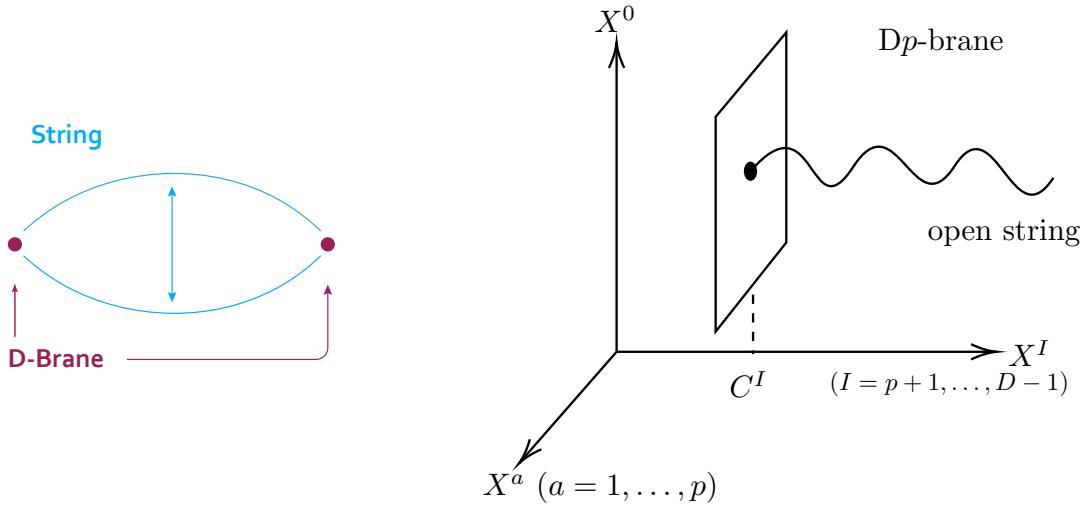


Figure 20: D-brane must exist at the ends of the open string so that momentum can escape from the string.

Chan-Paton factor

We can consider not only a single D-brane but also a stack of D-branes, and a string now has a choice of D-branes to which a string ends. Let us label this option i ($i = 1, \dots, n$), which is called **Chan-Paton factor**. As an open string has two endpoints, Chan-Paton degree of freedom is specified by

$$|N; k\rangle \rightarrow |N; k; ij\rangle . \quad (8.7)$$

Now we have n^2 massless vector states in both bosonic string and superstring theory. As usual, we use $n \times n$ Hermitian matrices T^a normalized to

$$\text{Tr}\left(T^a T^b\right) = \delta^{ab} ,$$

which consist of a complete set for data of the open string endpoints:

$$|N = 1; k; a\rangle = T_{ij}^a |N = 1; k; ij\rangle . \quad (8.8)$$

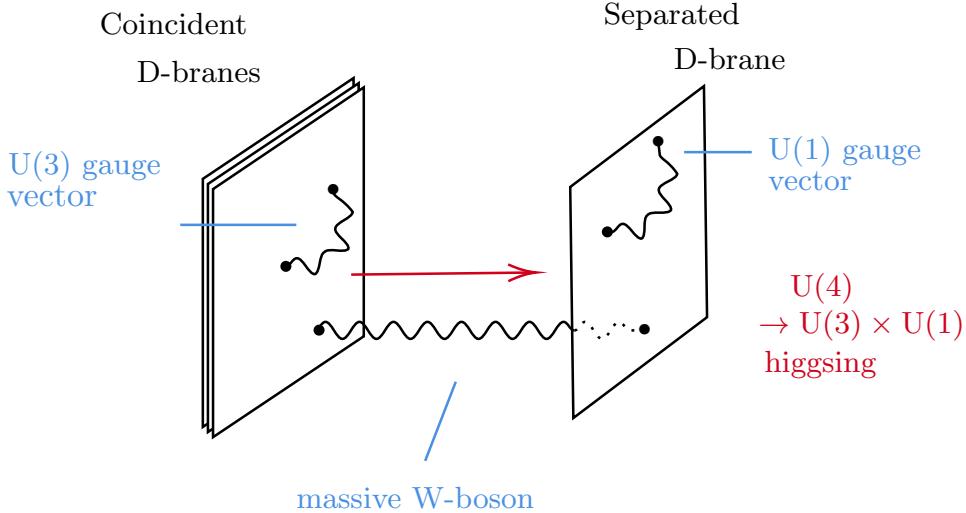


Figure 21: Many D-branes and Chan-Paton factors. $U(n)$ gauge group and Higgs mechanism.

These states correspond to a $U(n)$ gauge field. (You can check it by a three-point amplitude.)

Note that in order to realize $U(n)$ gauge group, the D-branes must coincide at a point. Otherwise, the gauge group is broken by Higgs mechanism (see Figure 21).

We will see in §10.1 that other types of gauge groups (SO or Sp) can be realized in string theory with orientifold planes.

D-branes in IIA/IIB superstring theory

D-branes arise as boundary conditions of open strings, and they carry gauge fields. Since D-branes are intrinsic to string theory, it reveals many interesting facets so that we will further investigate their properties. In §8.1, we saw that Type II superstring theories are endowed with R-R fields, which are anti-symmetric tensors analogous to gauge fields. Like an electron/monopole is coupled to the gauge fields, D-branes are coupled to the R-R fields in Type II theories.

To use the analogy of electromagnetism, let us quickly review an electron/monopole interacting with the electromagnetic fields in the $(3 + 1)$ -dim spacetime. An electron is expressed by a source

$$J_e^\mu = (\rho, \mathbf{j}) = (q_e \delta^3(\mathbf{r} - \mathbf{r}(t)), \partial_t \mathbf{r}(t) \rho),$$

and its coupling to the gauge field A is given by

$$S_J = \int A_\mu J^\mu d^4x = q_e \int_\gamma A_\mu dx^\mu = q_e \int_\gamma A,$$

where q_e is an electric charge, and γ is the world-line of the electron. Since the Maxwell equation is

$$*d * F_{(2)} = J_e, \quad (8.9)$$

we can obtain the electric charge by

$$q_e = \int_{S^2} *F_{(2)}.$$

On the other hand, the electromagnetic duality

$$dF_{(2)} = J_m \quad (8.10)$$

tells us that a magnetic monopole is a source $J_m = q_m \delta^3(r)$ of a magnetic flux. Therefore, the magnetic charge can be measured as

$$q_m = \int_{S^2} F_{(2)} = \int_B dF_{(2)},$$

where a surface S^2 around the monopole. It is well-known that the charges of the electron and monopole obey the Dirac quantization condition (for instance, see [Pol98, §13.3]):

$$q_e q_m \in 2\pi\mathbb{Z}. \quad (8.11)$$

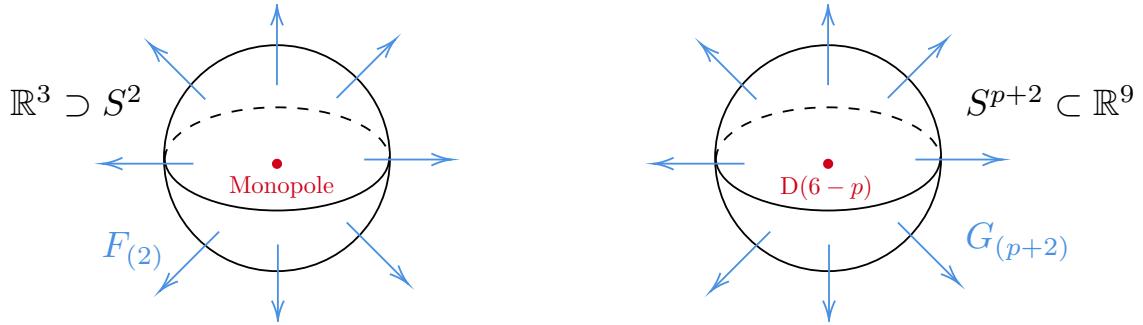


Figure 22: Higher dimensional analog of monopole: magnetic flux of $D(6 - p)$ -brane.

In a similar fashion, a Dp -brane ($p \leq 3$) supported on a world-membrane M_{p+1} is electrically coupled to the R-R ($p + 1$)-form $C_{(p+1)}$ as

$$S_p = \mu_p \int_{M_{p+1}} C_{(p+1)}. \quad (8.12)$$

where μ_p is the charge of the Dp -brane. An exterior derivative of the R-R potential gives its R-R field strength $G_{(p+2)} = dC_{(p+1)}$, and its electromagnetic dual is given by its Hodge dual

$$\tilde{G}_{(8-p)} = *G_{(p+2)}$$

From Gauss's law, the charge is given by the flux of $\tilde{G}_{(8-p)}$ over a sphere S^{8-p} around the Dp -brane

$$\mu_p = 2\kappa_{10}^2 \int_{S^{8-p}} \tilde{G}_{(8-p)},$$

where $2\kappa_{10}^2 = (2\pi)^7 \alpha'^4$. The magnetic dual of the Dp -brane is a $D(6 - p)$ -brane, and its magnetic charge is accordingly given by

$$\mu_{6-p} = 2\kappa_{10}^2 \int_{S^{p+2}} G_{(p+2)}.$$

They must obey the Dirac quantization condition

$$2\kappa_{10}^2 \mu_p \mu_{6-p} \in 2\pi\mathbb{Z}. \quad (8.13)$$

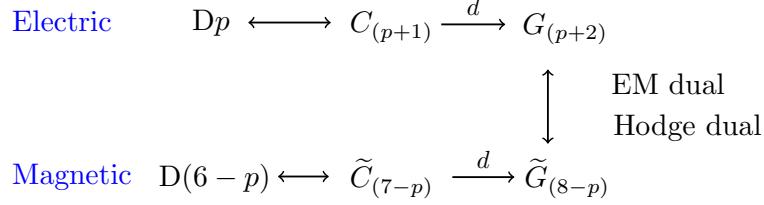


Figure 23: Electro-magnetic duality and D-brane.

Writing the flux $\tilde{G}_{(8-p)} = d\tilde{C}_{(7-p)}$, the $D(6-p)$ -brane is magnetically coupled to $\tilde{C}_{(7-p)}$. Now, reading off R-R fields from Table 2, we see that there are Dp -branes for even p in Type IIA and for odd p in Type IIB theory.

Even from the NS-NS sector, we can predict the presence of an extended object in string theory. A string or a fundamental string, denoted as F1, is electrically coupled to the B -field. On the other hand, there is an object magnetically coupled to the B -field, which is called **NS5-brane**.

In conclusion, we summarize extended objects in Type II superstring theory in Table 8.2.

IIA	$B_{(2)}$	$C_{(1)}$	$C_{(3)}$	IIB	$B_{(2)}$	$C_{(0)}$	$C_{(2)}$	$C_{(4)}^+$
Electric	F1	D0	D2	Electric	F1	$D(-1)$	D1	D3
Magnetic	NS5	D6	D4	Magnetic	NS5	D7	D5	D3

Table 3: Extended objects in Type II superstring theory

Note that a $D(-1)$ -brane is a timely localized object (called an instanton). There are D8-branes in Type IIA theory, which are non-dynamical so that there is no corresponding R-R anti-symmetric tensor. Moreover, these branes can be understood as a decay of the space-filling D9-brane and anti-D9-brane pair $D9-\bar{D}9$ [Sen98b], which admits a beautiful mathematical interpretation [Wit98c] by K-theory.

9 T-duality

We have seen that the critical dimensions D of bosonic and supersymmetric string theory are $D = 26$ and $D = 10$, respectively. To obtain an effective theory in lower dimensions, we can make use of **Kaluza-Klein compactifications** where the true spacetime takes the form of a direct product $M_d \times K_{D-d}$, where M_d is the d -dimensional Minkowski spacetime, and K_{D-d} is a very tiny compact manifold. As we will see, an effective theory in M_d still sees interesting “stringy” effects in this Kaluza-Klein scheme.

First, let us focus on the simplest form of compactification, known as *toroidal compactifications*, denoted by $K = T^{D-d}$. A torus is essentially a product of circles (S^1) and is characterized by its flat geometry. Consequently, the nonlinear sigma model for a torus can be described by a free 2d CFT. Intriguingly, such simple compactifications introduce the concept of **T-duality**. Furthermore, **Heterotic string theories**, which we will explore in §11, have been developed using these toroidal compactifications. To gain an understanding of the fundamental properties, we will first examine the toroidal compactifications in bosonic string theory. The toroidal compactifications of superstring theories are equally fascinating and merit detailed exploration in relation to string dualities later on.

The concept of D-branes was introduced as boundary conditions of open strings. T-duality gets particularly rich when we include D-branes so that we will study their properties more in detail. The explanation of T-duality in Polchinski's lecture notes [Pol96b] is so elegant that we just simply follow it in this section.

9.1 S^1 compactification in closed bosonic string

To begin with, let us first study the simplest case of the spacetime $\mathbb{R}^{1,24} \times S^1$ where we compactify 25-th direction on a circle S^1 of radius R . For closed strings, we have the familiar mode expansion (3.3). Now let us take a close look at the zero modes which can be written as

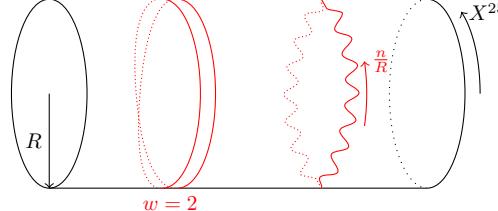
$$X^\mu(z, \bar{z}) = x^\mu - i\sqrt{\frac{\alpha'}{2}} (\alpha_0^\mu + \bar{\alpha}_0^\mu) t + \sqrt{\frac{\alpha'}{2}} (\bar{\alpha}_0^\mu - \alpha_0^\mu) \sigma + \text{oscillators.}$$

where the spacetime momentum of the string is

$$p^\mu = \frac{1}{\sqrt{2\alpha'}} (\alpha_0^\mu + \bar{\alpha}_0^\mu).$$

Under $\sigma \rightarrow \sigma + 2\pi$, the oscillator term is periodic and $X^\mu(z, \bar{z})$ changes by $2\pi\sqrt{(\alpha'/2)}(\bar{\alpha}_0^\mu - \alpha_0^\mu)$. For a non-compact spatial direction $\mathbb{R}^{1,24}$, X^μ is single-valued $X^\mu(t, \sigma) = X^\mu(t, \sigma + 2\pi)$, which requires

$$\alpha_0^\mu = \bar{\alpha}_0^\mu = \sqrt{\frac{\alpha'}{2}} p^\mu, \quad \mu = 0, 1, \dots, 24.$$



On the other hand, since the 25-th direction is put on the circle S^1 of radius R , it has a period $X^{25} \sim X^{25} + 2\pi R$. Hence, the momentum p^{25} can take the values n/R for $n \in \mathbb{Z}$ where n is called **Kaluza-Klein momentum**. Also, under $\sigma \sim \sigma + 2\pi$, $X^{25}(z, \bar{z})$ can change by $2\pi w R$ where w is called the **winding number**. Thus, we have

$$\alpha_0^{25} + \bar{\alpha}_0^{25} = \frac{2n}{R} \sqrt{\frac{\alpha'}{2}}, \quad \bar{\alpha}_0^{25} - \alpha_0^{25} = \sqrt{\frac{2}{\alpha'}} w R$$

implying

$$\alpha_0^{25} = \left(\frac{n}{R} - \frac{wR}{\alpha'} \right) \sqrt{\frac{\alpha'}{2}}, \quad \bar{\alpha}_0^{25} = \left(\frac{n}{R} + \frac{wR}{\alpha'} \right) \sqrt{\frac{\alpha'}{2}}. \quad (9.1)$$

Now, let us study their mass spectrum. The mass formula for the string with one dimension compactified on a circle can be interpreted from a 25-dimensional viewpoint

in which one regards each of the Kaluza-Klein momenta, which are given by n , as distinct particles. Thus, the mass formula is given by

$$\begin{aligned} M^2 &= - \sum_{\mu=0}^{24} p^\mu p_\mu = \frac{2}{\alpha'} (\alpha_0^{25})^2 + \frac{4}{\alpha'} (N - 1) \\ &= \frac{2}{\alpha'} (\bar{\alpha}_0^{25})^2 + \frac{4}{\alpha'} (\bar{N} - 1). \end{aligned} \quad (9.2)$$

where μ runs only over the non-compact dimensions. Here N and \bar{N} are the right and left number operators (2.19). Using (9.1), the difference between the two expressions above yield

$$N - \bar{N} = nw, \quad (9.3)$$

so that the level-matching condition is modified due to the S^1 compactification. In a similar fashion, the mass can be expressed in terms of the Kaluza-Klein momentum and the winding number

$$M^2 = \frac{n^2}{R^2} + \frac{w^2 R^2}{\alpha'^2} + \frac{2}{\alpha'} (N + \bar{N} - 2) \quad (9.4)$$

The mass spectra (9.4) of the theories at radius R and α'/R are identical when the winding and Kaluza-Klein modes are interchanged $n \leftrightarrow w$. This symmetry of the bosonic string theory is called **T-duality**. From the viewpoint of strings, a circle of radius R is equivalent to that of radius α'/R , and this is the main reason why we can avoid UV divergence in string theory. This shows that strings see geometry in an unprecedented and intriguing way, and this is one of the remarkable phenomena called “stringy geometry”.

Here's a revised version of the passage for improved clarity and flow:

From Equation (9.1), it is evident that this interchange results in:

$$\alpha_0^{25} \rightarrow -\alpha_0^{25}, \quad \bar{\alpha}_0^{25} \rightarrow \bar{\alpha}_0^{25}. \quad (9.5)$$

In fact, under the T-duality transformation, not only the zero mode but also the entire right-moving segment of the compact coordinate reverses its sign:

$$X'^{25}(z, \bar{z}) = -X^{25}(z) + \bar{X}^{25}(\bar{z}). \quad (9.6)$$

Remarkably, the energy-momentum tensor, the operator product expansions (OPEs), and all correlation functions remain invariant under this transformation. In other words, T-duality, which connects the theories at radii R and α'/R , represents an exact symmetry in perturbative closed string theory.

Due to T-duality, a theory with a compactification radius of R is equivalent to one with a radius of α'/R . This equivalence suggests the existence of a “minimal radius,” $R = \sqrt{\alpha'}$, known as the self-dual radius in string theory. At the self-dual radius, the duality $R \rightarrow \alpha'/R$ maps R back to itself, potentially leading to intriguing phenomena. We will explore the physics at the self-dual radius in the following.

Self-dual radius: $R = \sqrt{\alpha'}$

As we know, the massless spectra of bosonic string theory include gravitons. Hence, let us see the effect of the S^1 compactification on the gravitons. In the Kaluza-Klein

mechanism $M_{25} \times S^1$, the metric is decomposed into compact and non-compact space-time direction

$$ds^2 = G_{MN}dx^M dx^N = \tilde{G}_{\mu\nu}dx^\mu dx^\nu + G_{25,25}(dx^{25} + A_\mu dx^\mu)^2. \quad (9.7)$$

$$\tilde{G}_{\mu\nu} = G_{\mu\nu} - G_{25,25}A_\mu A_\nu$$

where the fields $G_{\mu\nu}$, $G_{25,25}$, and A_μ are allowed to depend only on the non-compact coordinates x^μ ($\mu = 0, 1, \dots, 24$). Under a coordinate transformation

$$x'^{25} = x^{25} + \lambda(x^\mu)$$

the part $G_{\mu,25} = G_{25,\mu}$ of the metric transforms as

$$A'_\mu = A_\mu - \partial_\mu \lambda.$$

Thus, it behaves as U(1) gauge field, and gauge transformations arise as part of the higher-dimensional coordinate transformation. On the other hand, the part $G_{25,25}$ of the metric behaves as a scalar field. Indeed, writing $G_{25,25} = e^{2\sigma}$, the Ricci scalar for the metric (9.7) can be written as

$$R_{26} = R_{25} - 2e^{-\sigma}\nabla^2 e^\sigma - \frac{1}{4}e^{2\sigma}F_{\mu\nu}F^{\mu\nu}.$$

Actually, it is straightforward to see the corresponding vertex operators at generic radius R :

$$\begin{aligned} \partial X^\mu \bar{\partial} X^\nu e^{ik \cdot X} &\longleftrightarrow G_{\mu\nu}, B_{\mu\nu}, \phi \\ \partial X^\mu \bar{\partial} X^{25} e^{ik \cdot X}, \partial X^{25} \bar{\partial} X^\mu e^{ik \cdot X} &\longleftrightarrow A^\mu, B_{\mu,25} \\ \partial X^{25} \bar{\partial} X^{25} e^{ik \cdot X} &\longleftrightarrow e^\sigma \end{aligned} \quad (9.8)$$

where $\nu = 0, \dots, 24$ runs the coordinate indices for M_{25} . In fact, the middle line indicates that the theory has $U(1)_\ell \times U(1)_r$ gauge symmetry at generic radius R .

However, at the self-dual radius $R = \sqrt{\alpha'}$, the mass formula (9.4) becomes

$$M^2 = \frac{1}{\alpha'}(n^2 + w^2 + 2(N + \bar{N} - 2)),$$

so that the massless spectra actually get enlarged. In addition to the generic solution $n = w = 0$, $N = \bar{N} = 1$, there are also massless modes as in the following table:

	n	w	\bar{N}	N
A	± 1	± 1	0	1
B	± 1	∓ 1	1	0
C	± 2	0	0	0
D	0	± 2	0	0

Hence, the states corresponding to A and B contain four new gauge bosons with vertex operators

$$\bar{\partial} X^\mu e^{\pm 2iX^{25}(z)/\sqrt{\alpha'}} e^{ik \cdot X}, \quad \partial X^\mu e^{\pm 2i\bar{X}^{25}(\bar{z})/\sqrt{\alpha'}} e^{ik \cdot X}.$$

Thus, $U(1)_\ell \times U(1)_r$ gets enhanced to $SU(2)_\ell \times SU(2)_r$. Consequently, the T-duality can be understood as the Weyl group \mathbb{Z}_2 of $SU(2)_r$. Indeed, C and D also give rise to new

gauge bosons (exercise). It is expected that the new gauge bosons must combine with the old into a non-Abelian theory. In fact, if one can define the current

$$j^\pm(z) = j^1(z) \pm i j^2(z) := e^{\pm 2iX^{25}(z)/\sqrt{\alpha'}} \quad j^3(z) := i \partial X^{25}(z) / \sqrt{\alpha'} ,$$

they satisfy the OPEs (Exercise)

$$j^a(z) j^b(0) \sim \frac{k \delta^{ab}}{2z^2} + \frac{i \epsilon^{abc} j^c(0)}{z} .$$

with $k = 1$. Here ϵ^{abc} is the structure constant of $SU(2)$. This is precisely the definition of $SU(2)$ affine Lie algebra with level $k = 1$. The same story is repeated for the left movers. Hence we see that we have an enhancement of gauge symmetry from $U(1)_\ell \times U(1)_r$ to $SU(2)_\ell \times SU(2)_r$ at $R = \sqrt{\alpha'}$.

In fact, when the theory moves away from the self-dual radius $R = \sqrt{\alpha'}$, the $SU(2)_\ell \times SU(2)_r$ gauge symmetry is Higgsed. The world-sheet action is deformed by turning on the marginal operator

$$V_{a\bar{a}} := j_a \bar{j}_{\bar{a}} e^{ik \cdot X} ,$$

which is equivalent to giving the VEV of the $(3,3)$ -component of the Higgs field. As a result, when the theory is away from the self-dual radius, the $SU(2)_\ell \times SU(2)_r$ gauge symmetry is spontaneously broken down to a $U(1)_\ell \times U(1)_r$.

9.2 T-duality of open strings

Now let us consider T-duality on the open string spectrum in the S^1 compactification. At the end of §2.2, we briefly study open string spectra. There, we learned that Neumann boundary condition on the X^μ -direction is achieved by imposing $\alpha_n^\mu = \bar{\alpha}_n^\mu$ while Dirichlet boundary condition is by $\alpha_n^\mu = -\bar{\alpha}_n^\mu$. Since T-duality along the 25-th direction transforms the modes as in (9.5), it therefore exchanges Neumann to Dirichlet condition [DLP89, Lei89]:

$$\partial_\sigma X^{25}|_{\sigma=0,\pi} = 0 \quad \rightarrow \quad \partial_t X'^{25}|_{\sigma=0,\pi} = 0 .$$

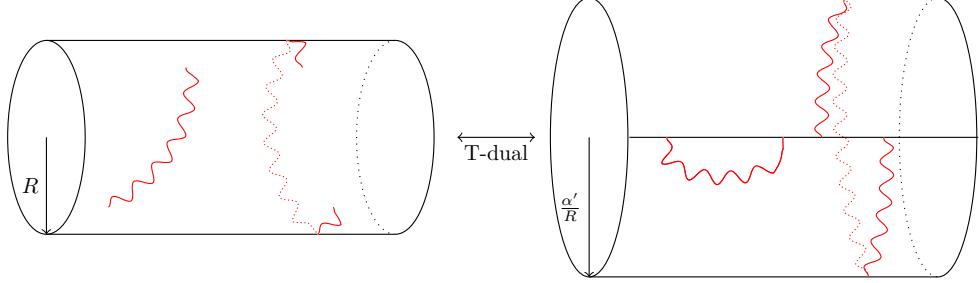
Suppose that an open string with Neumann boundary condition has KK momentum n along the 25-th direction, and we perform T-duality on that circle. Then, a simple calculation

$$X'^{25}(\pi) - X'^{25}(0) = \int_0^\pi d\sigma \partial_\sigma X'^{25} = i \int_0^\pi d\sigma \partial_t X^{25} \tag{9.9}$$

$$= 2\pi\alpha' p^{25} = \frac{2\pi\alpha' n}{R} = 2\pi n R'. \tag{9.10}$$

tells us that the X^{25} coordinate of the open string endpoints is fixed after T-duality. T-duality transforms the KK momentum n to winding number n as in the figure below. This can be interpreted as follows. Open strings can freely move on a space-filling D25-brane. After T-duality, the space-filling D25-brane becomes a D24-brane with its X^{25} coordinate fixed. As a result, all the endpoints of open strings are constrained to the fixed X^{25} -direction whereas they are free to move in the other 24 spatial dimensions.

More generally, the boundary condition of open strings can be imposed at any dimension. Since T-duality interchanges Neumann and Dirichlet boundary conditions, T-duality tangent to a Dp -brane brings it to a $D(p-1)$ -brane. On the other hand, T-duality orthogonal to a Dp -brane turns it into a $D(p+1)$ -brane.



We can ask a question which position of the X^{25} coordinate a D24-brane is located in the figure above. This question is related to Chan-Paton factors we encountered in §8.2. Suppose that there are K space-filling D25-branes, which give rise to $U(K)$ spacetime gauge fields. In the current setting, X^{25} direction is compactified on the circle S^1 so that we can include a Wilson loop on S^1 as

$$A_{25} = \text{diag}\{\theta_1, \theta_2, \dots, \theta_K\}/2\pi R \quad (9.11)$$

The insertion of the Wilson loop breaks the gauge group as $U(K) \rightarrow U(1)^K$, and the broken gauge group is abelian so that we can write it as

$$A_{25} = -i\Lambda^{-1}\partial_{25}\Lambda, \quad \Lambda = \text{diag}\{e^{iX^{25}\theta_1/2\pi R}, e^{iX^{25}\theta_2/2\pi R}, \dots, e^{iX^{25}\theta_N/2\pi R}\}.$$

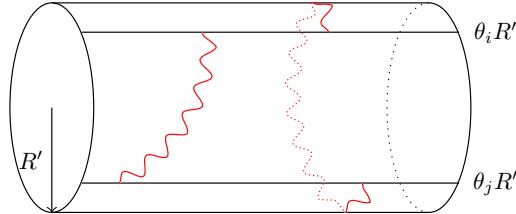
Then, under the translation $X^{25} \rightarrow X^{25} + 2\pi R$, an open string state $|N=1; n; ij\rangle$ is shifted by a phase

$$e^{i(\theta_j - \theta_i)}, \quad (9.12)$$

which means its momentum is shifted by $(\theta_j - \theta_i)/2\pi$ from an integer n . Under T-duality, the momentum is transformed into the winding number. Therefore, θ_i can be understood as the X^{25} coordinate of the i -th D24-brane after T-duality. Namely, the open string state $|N=1; n; ij\rangle$ is mapped to an open string of length

$$X'^{25}(\pi) - X'^{25}(0) = (2\pi n + \theta_j - \theta_i) R'.$$

There are in general K D24-branes at different positions as schematically depicted in the following figure.



Then, the $(D-1) = 24$ -dimensional mass of this open string is

$$\begin{aligned} M^2 &= (p^{25})^2 + \frac{1}{\alpha'}(N-1) \\ &= \left(\frac{[2\pi n + (\theta_j - \theta_i)] R'}{\pi\alpha'} \right)^2 + \frac{1}{\alpha'}(N-1). \end{aligned} \quad (9.13)$$

Hence, massless gauge bosons arise only from open strings $|N = 1; n = 0; ii\rangle$ whose endpoints are on the same D-brane.

Due to the string tension, open strings that stretch between different D-branes become massive (gauge field). Therefore, this can be understood as the Higgs mechanism in which X^{25} expectation values (coordinates) of K D24-branes take different values, and the gauge group is broken as $U(K) \rightarrow U(1)^K$. This situation is called **Coulomb phase**, where the Goldstone bosons are “eaten” by gauge fields $|N = 1; n = 0; ij\rangle$ which become massive.

Let us make an important remark. So far, we have treated D-branes just as rigid boundary conditions. However, D-branes are dynamical so that they can fluctuate in shape and position. As we will see in §15, §16, §17, the gravitational and gauge dynamics on a stack of D-branes makes string theory very intriguing.

9.3 T-duality of Type II superstrings

Let’s refine the phrasing for clarity and flow:

Let us examine how T-duality operates on Type II superstrings compactified on $M_9 \times S^1$. We observe that T-duality performs a parity transformation on the right-moving sector of the worldsheet, transforming

$$X^9(z) \longleftrightarrow -X'^9(z)$$

The superconformal invariance requires

$$\psi^9 \leftrightarrow -\psi'^9.$$

This alteration reverses the chirality of the right-moving R-sector ground state by swapping the raising and lowering operators $\psi^8 \pm i\psi'^9$ (refer to (7.19)). Essentially, T-duality acts as a spacetime parity operation on one side of the worldsheet, reversing the chiralities of the right- and left-moving ground states. Consequently, the Type IIA theory with a compactification radius R is T-dual to the Type IIB theory with radius α'/R .

Chiral spinors are defined through the relation:

$$\Gamma_{11}\psi_{\pm} = \pm\psi_{\pm}$$

with $\bar{\psi}_{\pm} = \psi_{\pm}^\dagger \Gamma^0$. The vertex operators for the R-R field strengths $G_{\mu_1 \dots \mu_{p+2}}$ are represented as spinor bilinears:

$$\text{IIA : } \bar{\psi}_-^L \Gamma^{\mu_1 \dots \mu_{p+2}} \psi_+^R, \quad \text{IIB : } \bar{\psi}_+^L \Gamma^{\mu_1 \dots \mu_{p+2}} \psi_+^R,$$

where ψ^R and ψ^L are derived from the right and left movers, respectively. The antisymmetric product of $(p+2)$ gamma matrices is denoted by $\Gamma^{\mu_1 \dots \mu_{p+2}}$.

Given that IIA and IIB theories have distinct R-R fields, T-duality in the 9th direction transforms one into the other, impacting the spin fields as follows:

$$\psi_{\alpha}^R(z) \rightarrow \psi_{\alpha}^R(z), \quad \bar{\psi}_{\alpha}^L(\bar{z}) \rightarrow \bar{\psi}_{\alpha}^L(\bar{z})\beta_9,$$

where $\beta_9 = \Gamma^{11}\Gamma^9$ represents the 9-reflection on spinors. The Γ_{11} just gives sign ± 1 when it acts on spinor fields because the R ground states have definite chirality. The Γ^9 adds a “9” to the set $\mu_1 \dots \mu_p$ if none is present, or removes one if it is present via $(\Gamma^9)^2 = 1$. This is how T-duality acts on the RR field strengths, adding or subtracting the index for

the dualized dimensions. Thus, if we start from the IIA string we get the IIB RR fields and vice versa. Consequently, the R-R fields are transformed as:

$$\begin{aligned} C_9 &\rightarrow C \\ C_\mu, C_{\mu\nu 9} &\rightarrow C_{\mu 9}, C_{\mu\nu} \\ C_{\mu\nu\lambda} &\rightarrow C_{\mu\nu\lambda 9}. \end{aligned} \tag{9.14}$$

Figure 24 illustrates the behavior of D-branes under T-duality, and Table 4 lists the corresponding D-brane configurations in Type II theories that are dual to each other.

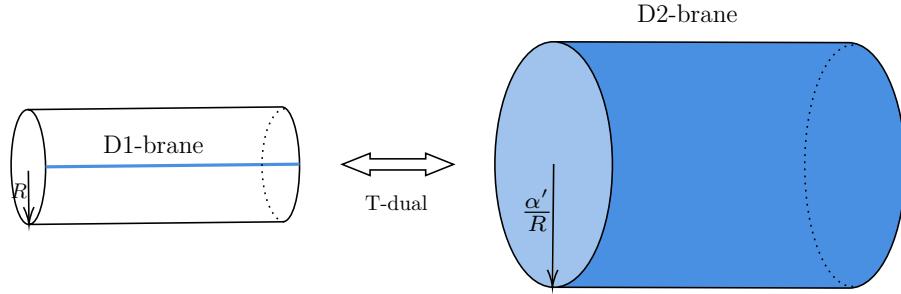


Figure 24: D1-brane and its T-dual D2-brane on $\mathbb{R} \times S^1$.

10 Unoriented strings and Type I theory

So far we have considered oriented world-sheet Riemann surfaces. However, it is quite natural to consider processes in Figure 25, which yields an unoriented world-sheet. Therefore, we will consider the so-called **unoriented string**, which leads to Type I su-

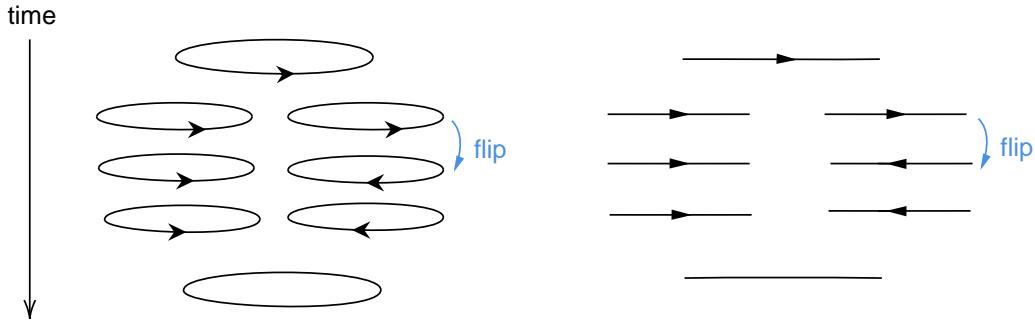


Figure 25: Unoriented processes. The left is closed string one, and the right is open string one.

perstring theory. As we will see below, an unoriented string theory needs to incorporate (unoriented) open strings.

Open strings can end on D-branes in Type II theories. However, without D-branes, Type II theories cannot incorporate open strings due to supersymmetry as follows. In fact, $D = 10$ $\mathcal{N} = 2$ supersymmetry requires 8_v to have two superpartners as for a

	0	1	2	3	4	5	6	7	8	9
IIB	NS5	x	x	x	x	x	x			
	D1	x	x							
	D5	x	x				x	x	x	x
IIA						T-dual on X^6				
	NS5	x	x	x	x	x				
	D2	x	x				x			
IIB	D4	x	x				x	x	x	
						T-dual on X^2				
	NS5	x	x	x	x	x				
IIB	D3	x	x	x			x			
	D5	x	x	x			x	x	x	
						S-dual				
	0	1	2	3	4	5	6	7	8	9
	D5	x	x	x	x	x				
	D3	x	x	x			x			
	NS5	x	x				x	x	x	

Table 4: D-branes and dualities in Type II theories. The last two configurations will appear in §15.3.

graviton to have two gravitinos in Table 2. On the other hand, after the GSO projection, open strings can have one of the following massless states:

$$\begin{aligned} P_{\text{GSO}} : \quad & \text{NS, R+} = \mathbf{8}_v + \mathbf{8}_s , \\ \tilde{P}_{\text{GSO}} : \quad & \text{NS, R-} = \mathbf{8}_v + \mathbf{8}_c , \end{aligned}$$

where $\mathbf{8}_v$ represents a gauge field and $\mathbf{8}_{s,c}$ are gauginos. Hence, the gauge field has only one superpartner so that it cannot exist in Type II theories without D-branes.

In fact, the existence of (parallel) D-branes breaks a half of supersymmetries so that open strings can end on D-branes. To consider unoriented string theory, we will introduce an **orientifold plane** or **O-plane** that breaks a half of supersymmetry. Although the argument below mainly focuses on the bosonic string for simplicity, it is straightforward to apply for superstrings.

10.1 Unoriented string theory

As in Figure 25, a world-sheet parity symmetry exchanges left- and right-movers for closed strings and a reversal of the spatial direction for open strings.

Let us first take a look at a world-sheet parity flip operator Ω on a bosonic closed string. It exchanges the left and right-moving modes which ends up with a reversal of

the spatial direction up to the overall sign:

$$\Omega : \begin{cases} X(t, \sigma) & \leftrightarrow X(t, -\sigma) \\ b(t, \sigma) & \leftrightarrow \bar{b}(t, -\sigma) \\ c(t, \sigma) & \leftrightarrow -\bar{c}(t, -\sigma) \end{cases}, \quad \Omega : \begin{cases} \alpha_n^\mu & \leftrightarrow \bar{\alpha}_n^\mu \\ b_n & \leftrightarrow \bar{b}_n \\ c_n & \leftrightarrow \bar{c}_n \end{cases}, \quad (10.1)$$

where t is the Euclidean world-sheet time, not τ . Only Ω invariant states such as the tachyon vacuum, gravitons and dilaton are survived under this projection, while the antisymmetric B -field is projected out. Since a closed string is electrically coupled to the B -field, this implies that an unoriented closed string is not stable and it should decay. We define the space-time after the projection as a space-time-filling $O(25)$ - plane. This is a special case of orientifold that will be introduced soon.

Now, let us introduce a space-filling D25-brane and consider an open string with the Neumann boundary condition. As in (2.26), the Neumann boundary condition identifies the left and right-moving modes. Then, the world-sheet parity operator acts as

$$\Omega : \begin{cases} X(t, \sigma) & \leftrightarrow X(t, \pi - \sigma) \\ b(t, \sigma) & \leftrightarrow \bar{b}(t, \pi - \sigma) \\ c(t, \sigma) & \leftrightarrow -\bar{c}(t, \pi - \sigma) \end{cases}, \quad \Omega : \begin{cases} \alpha_n^\mu & \leftrightarrow (-1)^n \alpha_n^\mu \\ b_n & \leftrightarrow (-1)^n b_n \\ c_n & \leftrightarrow (-1)^n c_n \end{cases}, \quad (10.2)$$

Now recall that we can incorporate Chan-Paton degrees of freedom (8.7) for open string states. The orientation flip Ω exchanges the Chan-Paton indices $i \leftrightarrow j$. Hence, writing an open string state as in (8.8)

$$|N, k; T\rangle \equiv |N, k; ij\rangle T_{ij},$$

Ω acts on the state as

$$\Omega |N, k; T\rangle = (-1)^N |N, k; T^t\rangle. \quad (10.3)$$

Therefore, it imposes the condition

$$\begin{aligned} T^t &= T && \text{for } N \text{ even} \\ T^t &= -T && \text{for } N \text{ odd}. \end{aligned} \quad (10.4)$$

In particular, a gauge boson arises at level one (2.29) so that we need to impose the anti-symmetric condition $T^t = -T$ on the Chan-Paton factor. Consequently, it gives rise to a $SO(n)$ gauge field.

In general, when the world-sheet parity is flipped, we can also transform the Chan-Paton index by a matrix P at the same time.

$$\Omega |N, k; ij\rangle = (-1)^N P_{jj'} |N, k; j'i'\rangle P_{i'i}^{-1}$$

The projection operator is subject to $\Omega^2 = 1$, which gives the constraint

$$\Omega^2 |N, k; ij\rangle = [(P^T)^{-1} P]_{ii'} |N, k; i'i'\rangle (P^{-1} P^T)_{j'j}, \Rightarrow P^t P^{-1} = \pm 1$$

Furthermore, there is a $U(n)$ gauge equivalence relation $T \sim \tilde{T} = U T U^{-1}$ ($U \in U(n)$), which amounts to $\tilde{P} \sim U P U^t$. Taking the gauge equivalence into account, $P^t P^{-1} = 1$

leads to $P = 1$, in which the gauge group becomes $\mathrm{SO}(n)$ as above. On the other hand, for $P^t P^{-1} = -1$, the rank needs to be even $n = 2k$, and we can take a basis such that

$$P = i \begin{pmatrix} 0 & -\mathbf{1}_{k \times k} \\ \mathbf{1}_{k \times k} & 0 \end{pmatrix}$$

In this situation, the gauge Lie algebra becomes $\mathrm{Sp}(k)$ so that $T^t = -PTP$. The choice of the gauge group is related to the sign of charge and tension for the orientifold plane. We then denote orientifold plane O^+ or O^- for gauge group $\mathrm{Sp}(k)$ and $\mathrm{SO}(n)$ respectively. We will discuss it in detail at the end of this section.

Unoriented superstring spectrum

Now, let us consider the unoriented closed string in superstring theories. Since Ω exchanges the left- and right-moving modes, both must have the same spectra. Hence, it can only act on Type IIB theory but not on Type IIA theory. (See (8.1).) In Table 2, the flip Ω projects out the B -field [2] in the NS-NS sector, as well as a half of the NS-R R-NS sector $\mathbf{8}_c + \mathbf{56}_s$ (only the diagonal part survives). Supersymmetry requires that the number of bosons and fermions are the same, which implies that $[0]$ and $[4]_+$ are eliminated and only the second rank anti-symmetric field [2] survives in the R-R sector. This implies that Ω projects out D(-1)-, D3-, D7-branes. D1- and D5-branes survive, which are electrically and magnetically coupled to the R-R two-form $C_{(2)}$, respectively. The remaining states are

$$\begin{aligned} [0] \oplus [2] \oplus (2) \oplus \mathbf{8}_c \oplus \mathbf{56} = & \mathbf{1} \oplus \mathbf{28} \oplus \mathbf{35} \oplus \mathbf{8}_c \oplus \mathbf{56}_s \\ & \Phi \quad C_{\mu\nu} \quad G_{\mu\nu} \quad \lambda^+ \quad \psi_\mu^- \end{aligned} \tag{10.5}$$

This gives us 10d $\mathcal{N} = 1$ supergravity multiplet. However, this theory is anomalous. We thus need to incorporate unoriented open strings to cancel this anomaly [GS84, GS85]. In conclusion, Type I is an unoriented, open plus closed superstring theory in which the massless spectrum is

$$[0] \oplus [2] \oplus (2) \oplus \mathbf{8}_c \oplus \mathbf{56} \oplus (\mathbf{8}_v \oplus \mathbf{8}_s)_{\mathrm{SO}(n) \text{ or } \mathrm{Sp}(n)} .$$

It can be shown that $\mathcal{N} = 1$ supersymmetric Yang-Mills theory with gauge group $\mathrm{SO}(32)$ or $E_8 \times E_8$ can be the additional sector to cancel this anomaly. However, $E_8 \times E_8$ gauge group can not be constructed from unoriented open string by the Chan-Paton factors. The gauge group $\mathrm{SO}(32)$ is the only option for type I string theory. From the world-sheet perspective, the anomaly cancellation corresponds to the **tadpole cancellation**. In what follows, we will investigate the tadpole cancellation in bosonic string theory, and will see that the gauge group needs to be $\mathrm{SO}(2^{13})$. A similar analysis tells us that the gauge group must be $\mathrm{SO}(32)$ in Type I superstring theory.

10.2 Amplitude in unoriented string theory

As above, unoriented open strings give rise to either $\mathrm{SO}(n)$ or $\mathrm{Sp}(n)$ massless gauge fields. However, unless the tadpole cancellation is satisfied, the theory is still anomalous. We will see it in bosonic string theory.

Cylinder

First, let us evaluate an “oriented” open string one-loop amplitude where a worldsheet is a cylinder (or annulus) as in the left of Figure 26. We consider that there are n D25-branes (or D9-branes for superstring) so that all of the boundary conditions are of Neumann type. The evaluation is parallel to §6. It is clear from Figure 26 that we notice

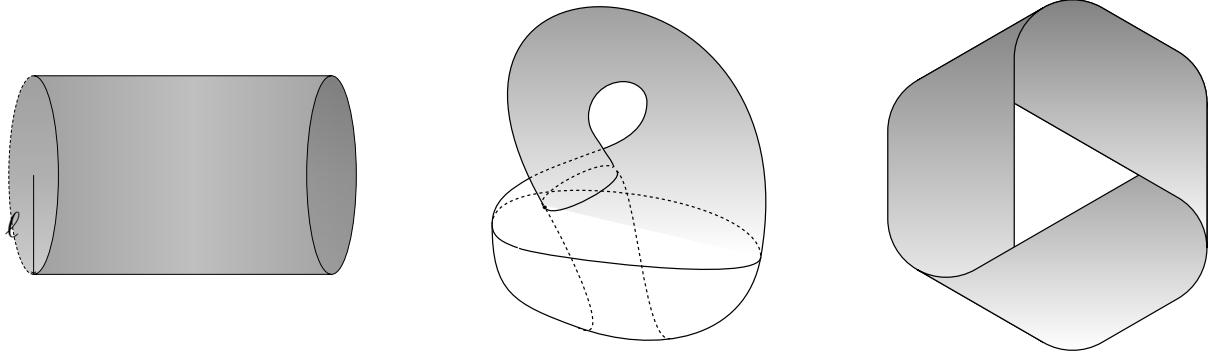


Figure 26: Cylinder, Klein bottle and Möbius strip.

the following facts:

- The range is $0 \leq \text{Re } w \leq \pi$, the period is $w \sim w + 2\pi i l$.
- There is a real modulus l so that the amplitude needs one b zero mode insertion.
- There is a time translational isometry, a shift of $\text{Im } w$ so that the amplitude needs one c zero mode insertion.

Consequently, the cylinder partition function is written as

$$A_{0,C} = \int \frac{dl}{2l} \langle b_0 c_0 \rangle_{\text{gh}} \langle 1 \rangle_{\text{mat}},$$

where the ghost zero modes in open strings are read off

$$\begin{aligned} b_0 &= \frac{1}{2\pi} \int_0^\pi [dw b(w) + d\bar{w} \bar{b}(\bar{w})] , \\ c_0 &= \frac{i}{2\pi} \int_0^\pi [dw c(w) - d\bar{w} \bar{c}(\bar{w})] . \end{aligned}$$

Using operator formalism, we can derive each contribution as

$$\begin{aligned} \langle 1 \rangle_{\text{mat}} &= n^2 \text{ Tr} \left[q^{L_0^X - \frac{c}{24}} \right] = n^2 \cdot \frac{iV_{26}}{(2\pi)^{26}} (2l\alpha')^{-13} \cdot \eta(il)^{-26} , \\ \langle b_0 c_0 \rangle_{\text{gh}} &= \text{Tr} \left[(-1)^F b_0 c_0 q^{L_0^S - \frac{c}{24}} \right] = \eta(il)^2 , \end{aligned}$$

where $q = e^{-2\pi l}$ and the Virasoro generators are in (5.27) for open strings. Note that n^2 comes from the trace over the Chan-Paton degrees of freedom. Therefore, the cylinder

partition function is

$$A_{0,C} = n^2 \cdot \frac{iV_{26}}{(2\pi\ell_s)^{26}} \int_0^\infty \frac{dl}{2l} \frac{1}{(2l)^{13} \eta(il)^{24}}, \quad (10.6)$$

where $\ell_s = \sqrt{\alpha'}$ is the string length.



Figure 27: Pictorial description of the equivalence between open string one-loop and closed string propagation.

Let us look into its physical implication. When $l \rightarrow 0$, the cylinder becomes a thin tube, and it can be regarded as a propagation of a closed string as in Figure 27. This is called the **open-closed duality**. This is justified by rewriting the amplitude. Using the modular property (7.36) of the Dedekind η -function, we can rewrite it in terms of $l = s^{-1}$ as

$$A_{0,C} = \frac{n^2}{2^{13}} \cdot \frac{iV_{26}}{(2\pi\ell_s)^{26}} \int_0^\infty \frac{ds}{2} \eta(is)^{-24}, \quad (10.7)$$

where

$$\eta(is)^{-24} = q^{-1} + 24 + \dots \equiv \sum_{N=0}^{\infty} \mathcal{N}_N q^{N-1} \quad (q = e^{-2\pi s}).$$

The leading term is divergent due to the closed string tachyon. While the second term gives a divergence when $s \rightarrow \infty$. The UV divergence $l \rightarrow 0$ in an open string is now replaced by the IR divergence $s \rightarrow \infty$ of a closed string propagation, which can be understood as particle propagations (sum of lines) as follows (see also Figure 28).

$$\int_0^\infty ds \sum_{N=0}^{\infty} \mathcal{N}_N e^{-2\pi s(N-1)} \sim \sum_i \int_0^\infty ds e^{-s(k^2 + m_i^2)} = \sum_i \frac{1}{k^2 + m_i^2} \Big|_{k=0}.$$

We can see that the IR divergence is from massless particle propagation (graviton etc.), which is absorbed or emitted from D25-branes. This point will be revisited at the end of §15.1. In conclusion, the divergence is due to the existence of the D25-branes, which have definite tension (this is why they emit graviton/dilaton). To cure this divergence, we will consider unoriented string amplitudes when world-sheets are Klein bottle and Möbius strip as in Figure 26. The reason will become clear later.

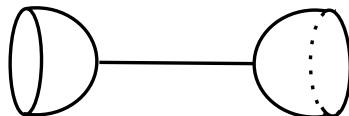


Figure 28: Intermediate propagation is replaced by particles (lines).

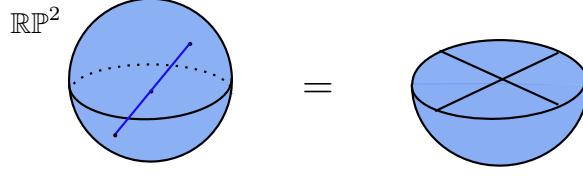


Figure 29: $\mathbb{R}\mathbb{P}^2$ is formed by identifying antipodal points on S^2 , akin to connecting the southern hemisphere to the equator.

Klein bottle amplitude

We proceed to evaluate the one-loop amplitude of an unoriented closed string propagating on a Klein bottle world-sheet, as depicted in Figure 30. The Klein bottle arises from identifying the top and bottom edges of the string world-sheet via the world-sheet parity operator.

The amplitude takes the form:

$$A_{0,K} = \int \frac{dl}{2l} \text{Tr} \left[\Omega(-1)^F \frac{1}{2} (b_0 + \bar{b}_0) \frac{1}{2} (c_0 + \bar{c}_0) q^{L_0 + \bar{L}_0 - \frac{c}{12}} \right] \quad (q = e^{-2\pi l}).$$

Using (10.1), we can express it as:

$$A_{0,K} = \int \frac{dl}{2l} \text{Tr} \left[(-1)^F b_0 c_0 q^{2L_0^{\text{tot}} - \frac{c^{\text{tot}}}{12}} \right],$$

where L_0 and c denote the Virasoro generators and the central charge for closed strings, respectively. Hence, the resulting expression becomes:

$$\begin{aligned} A_{0,K} &= \frac{iV_{26}}{(2\pi\ell_s)^{26}} \int \frac{dl}{2l} \frac{1}{l^{13} \eta(2il)^{24}} \\ &= 2^{13} \frac{iV_{26}}{(2\pi\ell_s)^{26}} \int \frac{ds}{2} \eta(is)^{-24}. \end{aligned} \quad (10.8)$$

where we apply the modular S -transformation ($2l = s^{-1}$) from the first to second line. Similarly, we encounter a tadpole divergence originating from the Klein bottle.

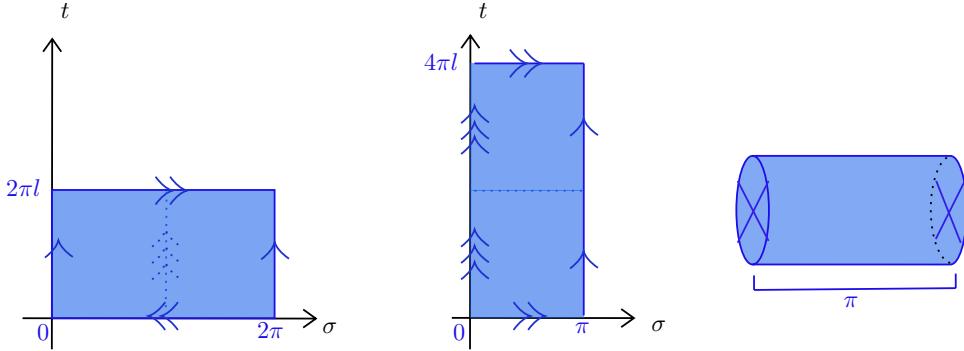


Figure 30: Representation of the Klein bottle as a cylinder with cross-cap boundaries at both ends.

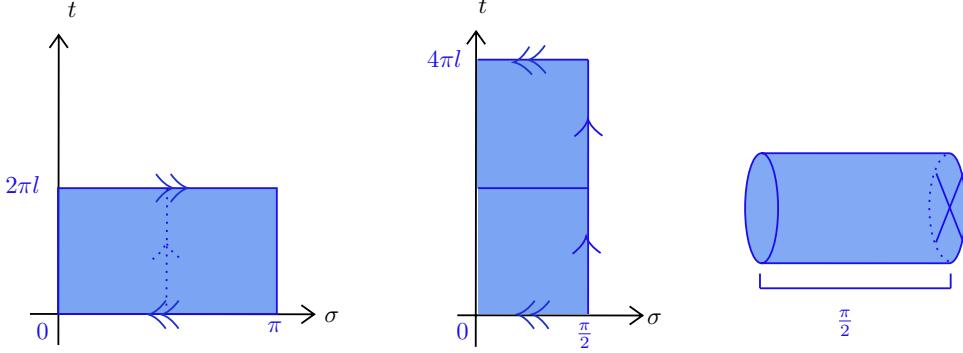


Figure 31: Möbius strip.

Möbius strip amplitude

Finally, let us evaluate the one-loop amplitude of an unoriented open string on a Möbius strip world-sheet, as depicted in Figure 31. The Möbius strip exhibits several key characteristics:

- The spatial range is $0 \leq \sigma \leq \pi$, with the time period is $2\pi l$ together with the orientation flip: $(t, \sigma) \sim (t + 2\pi l, \pi - \sigma)$.
- There is a real modulus l ; the amplitude needs one b zero mode insertion.
- There is a real isometry, which is the time translation; the amplitude needs one c zero mode insertion.

Thus, the Möbius strip amplitude is given by

$$A_{0,M} = \int \frac{dl}{2l} \text{Tr} \left[\Omega(-1)^F b_0 c_0 q^{L_0 - \frac{c}{12}} \right] \quad (q = e^{-2\pi l}).$$

The trace is over Hilbert space of the matter and ghost sectors on the strip as well as the Chan-Paton factors, which give in total n^2 degeneracy for each state. We can divide the effect of Ω into two parts as follows.

$$\Omega|\Phi; T\rangle = |\Omega\Phi; PT^t P\rangle \equiv \Omega_\Phi \cdot \Omega_T |\Phi; T\rangle.$$

Let us see the Ω_T , which is defined as

$$\Omega_T = \frac{PT^t P}{T}.$$

For $\text{SO}(n)$, where $P = 1$, we obtain:

$$\Omega_{T,\text{SO}} = \frac{T^t}{T} = \begin{cases} +1 & (\text{for symmetric } T) \\ -1 & (\text{for anti-symmetric } T) \end{cases},$$

resulting in:

$$\text{Tr}_{T,\text{SO}} [\Omega] = \frac{n(n+1)}{2} - \frac{n(n-1)}{2} = n.$$

In the case of $\text{Sp}(n)$ (exercise), we have

$$\text{Tr}_{T,\text{Sp}} [\Omega] = -n.$$

Since the action of Ω to a state of level N introduces the sign $(-1)^N$ (10.3), the contribution from the matter and ghost part can be evaluated as

$$\begin{aligned}\text{Tr}_\Phi \left[\Omega(-1)^F b_0 c_0 q^{L_0 - \frac{c}{12}} \right] &= \frac{iV_{26}}{(2\pi\ell_s)^{26}} \frac{1}{(2l)^{13}} \cdot q^{-1} \prod_{N=1}^{\infty} \frac{(1 - (-q)^N)^2}{(1 - (-q)^N)^{26}} \\ &= \frac{iV_{26}}{(2\pi\ell_s)^{26}} \frac{1}{(2l)^{13}} \cdot \frac{-1}{\eta(il + \frac{1}{2})^{24}}.\end{aligned}$$

Therefore, the Möbius string amplitude is

$$\begin{aligned}A_{0,M} &= \pm n \cdot \frac{iV_{26}}{(2\pi\ell_s)^{26}} \int \frac{dl}{2l} \frac{1}{(2l)^{13}} \cdot \frac{-1}{\eta(il + \frac{1}{2})^{24}}, \\ &= \mp 2n \cdot \frac{iV_{26}}{(2\pi\ell_s)^{26}} \int \frac{ds}{2} \eta(is + \frac{1}{2})^{-24},\end{aligned}$$

where, by setting $2l = s^{-1}$, we apply the modular property (7.36) to the following expression:

$$q^{-1} \prod_{N=1}^{\infty} (1 - (-q)^N)^{-24} = -\eta(il + \frac{1}{2})^{-24} = \vartheta_3(2il)^{-12} \eta(2il)^{-12}. \quad (10.9)$$

To sum up, three amplitudes are (introduced additional $\frac{1}{2}$ factor for Cylinder as an unoriented amplitude)

$$\begin{aligned}A_{0,C} &= \frac{n^2}{2^{13}} \cdot \frac{iV_{26}}{(2\pi\ell_s)^{26}} \int_0^\infty \frac{ds}{2} \eta(is)^{-24}, \\ A_{0,K} &= 2^{13} \cdot \frac{iV_{26}}{(2\pi\ell_s)^{26}} \int \frac{ds}{2} \eta(is)^{-24}, \\ A_{0,M} &= \mp 2n \cdot \frac{iV_{26}}{(2\pi\ell_s)^{26}} \int \frac{ds}{2} \eta(is + \frac{1}{2})^{-24},\end{aligned}$$

where

$$\begin{aligned}\eta(is)^{-24} &= q^{-1} + 24 + \mathcal{O}(q) \quad (q = e^{-2\pi s}), \\ \eta(is + \frac{1}{2})^{-24} &= -q^{-1} + 24 + \mathcal{O}(q).\end{aligned}$$

As we saw in the oriented string case, the massless states lead to IR singularity. On the other hand, in the unoriented open string case, we have

$$\frac{1}{2^{13}} \left[n \mp 2^{13} \right]^2 \cdot \frac{iV_{26}}{(2\pi\ell_s)^{26}} \int_0^\infty \frac{ds}{2} \cdot (24 + \dots),$$

where the upper (minus) sign corresponds to $\text{SO}(n)$ and the lower (plus) sign to $\text{Sp}(n)$. Therefore, the unoriented IR divergence vanishes only for $\text{SO}(2^{13}) = \text{SO}(8192)$. This cancellation can be illustrated in Figure 32. The cross cap shows another object (other than D-brane) that absorbs and emits gravitons etc., which is called O-plane. In this situation, it should be space-filling. Hence, it is $O25^-$ -plane, and a single $O25^-$ -plane cancels the tension of 2^{13} D25-branes.

$$[\text{circle} + \text{circle with dot}]^2 = \text{circle connected to circle with dot} + 2 \text{circle connected to circle with dot} + \text{circle connected to circle with dot}$$

Figure 32: Pictorial expression for the unoriented open string amplitude.

Although our discussion was in the bosonic string, parallel argument perfectly works for superstring. Here we give the final results for the cylinder, Klein bottle, and Möbius strip amplitude in R-R sectors (exercise):

$$\begin{aligned} A_{0,C} &= n^2 \cdot \frac{iV_{10}}{(2\pi\ell_s)^{10}} \int_0^\infty \frac{dl}{8l} \frac{\vartheta_2(il)^4}{(2l)^5 \eta(il)^{12}}, \\ A_{0,K} &= 2^5 \cdot \frac{iV_{10}}{(2\pi\ell_s)^{10}} \int_0^\infty \frac{dl}{8l} \frac{\vartheta_2(2il)^4}{(2l)^5 \eta(2il)^{12}}, \\ A_{0,M} &= \mp n \cdot \frac{iV_{10}}{(2\pi\ell_s)^{10}} \int_0^\infty \frac{dl}{8l} \frac{\vartheta_2(il + \frac{1}{2})^4}{(2l)^5 \eta(il + \frac{1}{2})^{12}}, \end{aligned} \quad (10.10)$$

respectively. Consequently, we have the unoriented superstring amplitude

$$\frac{1}{2^5} [n \mp 2^5]^2 \cdot \frac{iV_{10}}{(2\pi\ell_s)^{10}} \int_0^\infty \frac{ds}{8} \cdot (16 + \dots) \quad (10.11)$$

where the IR divergence vanishes for $\mathrm{SO}(2^5) = \mathrm{SO}(32)$ [PC88]. This is another way to show the Green-Schwarz anomaly cancellation [GS84, GS85]. This means that

$$\text{Type I} = \text{Type IIB} + 32 \text{ D9-branes} + \text{O9}^- \text{-plane}.$$

Note that in the superstring case, D-branes and O-plane have RR-charge in addition to tension, which has relations

$$T_{\text{O9}^\pm} = \pm 32 \cdot T_{\text{D9}} \quad (\text{tension}), \quad \mu_{\text{O9}^\pm} = \pm 32 \cdot \mu_{\text{D9}} \quad (\text{RR-charge}).$$

10.3 T-duality of Type I theory

Let us recall T-duality. Consider X^i is S^1 compactified and T-duality acts as follows:

$$T_i : \quad X^i(z, \bar{z}) = X^i(z) + \bar{X}^i(\bar{z}) \quad \rightarrow \quad X'^i(z, \bar{z}) = -X^i(z) + \bar{X}^i(\bar{z}).$$

On the other hand, the orientation flip acts as follows:

$$\Omega : \quad X^i(z, \bar{z}) = X^i(z) + \bar{X}^i(\bar{z}) \quad \rightarrow \quad X^i(\bar{z}, z) = \bar{X}^i(z) + X^i(\bar{z}).$$

Therefore, in the T-dual coordinate X' , the orientation flip acts as

$$\Omega : \quad X'^i(z, \bar{z}) \quad \rightarrow \quad -X'^i(\bar{z}, z).$$

This is understood as spacetime **orbifold** as well as world-sheet orientation flip (see Figure 33), which is called **orientifold**. Therefore, the dual space is not S^1 but S^1/\mathbb{Z}_2 with radius $\tilde{R} = \frac{\alpha'}{R}$. Note that there are two fixed points where O-planes sit and induce the spacetime reversal and the orientation flip.

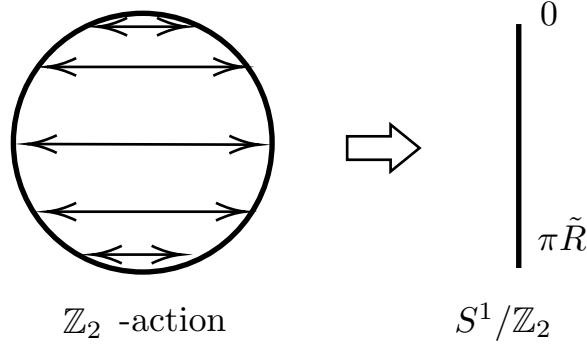


Figure 33: \mathbb{Z}_2 orbifold of S^1 .

Let us consider Type I superstring theory with X^9 compactified on S^1 and take T-duality along the S^1 . With a proper Wilson line

$$A_9 = i \begin{pmatrix} & -a_1 & & \\ a_1 & & & \\ & & -a_2 & \\ & a_2 & & \ddots \end{pmatrix}$$

D8-branes sit at different points in \mathbb{Z}_2 symmetric way (see Figure 34). Note that an $O9^-$ -

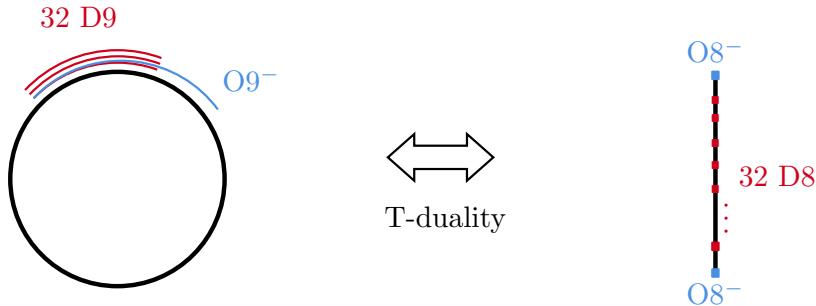


Figure 34: T-dual of Type I superstring theory.

plane splits into two $O8^-$ -planes. Accordingly, tension and RR-charge reduce by 2. In the end, the T-dual of Type I on S^1 is

$$\text{Type I}' = \text{Type IIA on } S^1/\mathbb{Z}_2 \text{ with } 2 \text{ } O8^- \text{-plane} + (16 + 16) \text{ D8-branes}. \quad (10.12)$$

10.4 Op-plane

Of course, one can consider further T-dualities along other directions. In particular, an orientifold p -plane (Op) in Type II string theory is defined [DP96, Dab97] as the \mathbb{Z}_2 involution⁵

$$Op : \mathbb{R}^{1,p} \times \mathbb{R}^{9-p}/I_{9-p}\Omega \cdot \begin{cases} 1 & p = 0, 1 \bmod 4 \\ (-1)^{F_L} & p = 2, 3 \bmod 4, \end{cases} \quad (10.13)$$

⁵Note that we have already seen the special case $p = 9$, where the space-time is fixed under the involution.

where I_{9-p} is the involution of all coordinates in the transverse space \mathbb{R}^{9-p} and F_L is the left-moving spacetime fermion number operator. The presence of $(-1)^{F_L}$ for $p = 2, 3 \bmod 4$ can be understood from the requirement that the generator square to one on fermions. Due to the spacetime involution I_{9-p} transverse to the Op -plane, the number of Dp -branes parallel to the Op -plane is effectively doubled. Therefore, a stack of n Dp -branes along the Op^- -plane leads to an $SO(2n)$ gauge theory (Dynkin type D_n) in Type II theory. To realize an $SO(2n+1)$ gauge group (Dynkin type B_{n+1}), we need a combination of an Op^- -plane and $\frac{1}{2}$ Dp -brane, which is denoted by $\widetilde{O}p^-$. Consequently, n Dp -branes with $\widetilde{O}p^-$ -plane give rise to an $SO(2n+1)$ gauge theory as a world-volume theory. As we will see in §13.1, it is natural to incorporate an $\widetilde{O}p^+$ -plane in Type II theory, which also gives rise to an $Sp(k)$ gauge theory like an Op^+ -plane. They behave differently under the S-duality of Type IIB theory.

Each T-duality doubles the number of O-planes, and hence, reduces the tension and the R-R charges. Namely, we have the following relations:

$$T_{Op^\pm} = \pm 2^{p-5} \cdot T_{Dp} \quad (\text{tension}), \quad \mu_{Op^\pm} = \pm 2^{p-5} \cdot \mu_{Dp} \quad (\text{R-R charge}).$$

Since an $\widetilde{O}p^-$ -plane is a combination of an Op^- -plane and $\frac{1}{2}$ Dp -brane, there is $\frac{1}{2}$ -shift in R-R charge as summarized in Table 10.4. Op -planes bring richness to string theory such as quiver gauge theories, the topology of \mathbb{RP}^p , real K-theory, Dirac quantization conditions, etc, which are beyond the scope of this lecture. We refer to [Wit98b, HK00, HZ99, TY19] and references therein.

O-plane	gauge group	RR-charge
Op^-	$SO(2n)$	-2^{p-5}
Op^+	$Sp(2n)$	$+2^{p-5}$
$\widetilde{O}p^-$	$SO(2n+1)$	$-2^{p-5} + \frac{1}{2}$
$\widetilde{O}p^+$	$Sp(2n)$	$+2^{p-5}$

11 Heterotic string theories

We now embark on the study of **Heterotic string theories** [GHMR85b, GHMR85a, GHMR86]. Heterotic string is a hybrid construction of the left-moving sector of the 26-dimensional bosonic string and the right-moving sector of 10-dimensional superstring. In the Heterotic formulation, the 16 extra bosonic degrees of freedom from the left-moving sector are compactified on specific 16-dimensional tori, giving rise to two possible gauge groups: $SO(32)$ or $E_8 \times E_8$. The choice of 16-dimensional tori is of crucial importance, as these tori possess exceptional properties. Alternatively, we can describe the left-moving sector in terms of 32 free fermions, whose current algebra corresponds to either $SO(32)$ or $E_8 \times E_8$ at level $k = 1$. This combination of two distinct types of modes has been aptly named **heterosis**, highlighting the unique hybridization inherent in the Heterotic string theory.

11.1 Bosonic construction

Toroidal compactifications

In §9, we explored the compactification of bosonic string theory on a circle S^1 . Now, we will extend our analysis to the compactification of bosonic string theory on a d -dimensional torus T^d , resulting in an effective theory with $(26 - d)$ dimensions. The torus is defined by identifying points in the d -dimensional internal space according to

$$X^I \sim X^I + 2\pi e_i^I w^i = X^I + 2\pi W^I, \quad \text{for } w^i \in \mathbb{Z}. \quad (11.1)$$

where capital letters denote the compact dimensions. The vectors $\mathbf{e}_i = \{e_i^I\}$ ($i = 1 \dots d$) are d linearly independent vectors known as **vielbein**, which generate a d -dimensional lattice Λ . Moreover, the vielbein brings the metric into the standard Euclidean form

$$G_{ij} = \mathbf{e}_i \cdot \mathbf{e}_j = e_i^I e_j^J \delta_{IJ}, \quad X^I \equiv e_i^I X^i \quad (11.2)$$

The torus T^d on which we perform the compactification is obtained by dividing \mathbb{R}^d by Λ

$$T^d = \frac{\mathbb{R}^d}{2\pi\Lambda}.$$

The momentum p^I conjugate to the coordinates X^I on the torus is quantized as $\mathbf{p} \cdot \mathbf{W} \in \mathbb{Z}$. Hence, the momentum p takes values on the dual lattice Λ^* (Figure 35)

$$\Lambda^* \equiv \{e^{*I} n_i; \quad n_i \in \mathbb{Z}\}, \quad G^{ij} = \mathbf{e}^{*i} \cdot \mathbf{e}^{*j} = e_I^{*i} e_J^{*j} \delta^{IJ}.$$

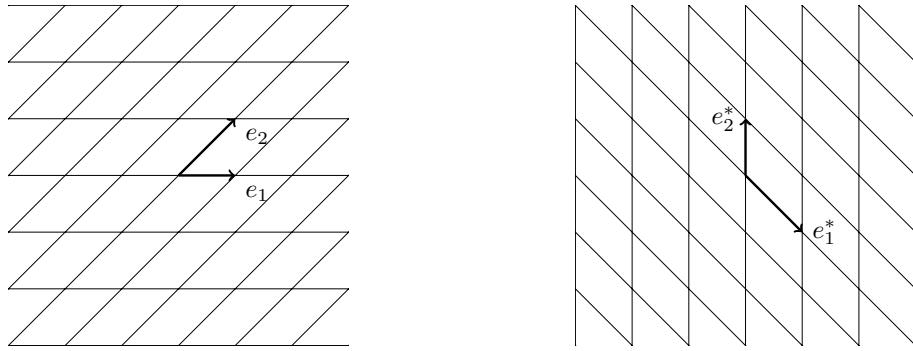


Figure 35: Lattice and its dual lattice

The closed string in the compact directions $X^I(z, \bar{z}) = X_R^I(z) + \bar{X}_L^I(\bar{z})$ obeys the condition

$$X^I(\sigma + 2\pi, \tau) = X^I(\sigma, \tau) + 2\pi W^I$$

where W^I serves as an analogue of winding numbers. The mode expansion for the compact directions is given by

$$\begin{aligned} X_R^I(z) &= x^I - i\sqrt{\frac{\alpha'}{2}} p_R^I \ln z + i\sqrt{\frac{\alpha'}{2}} \sum_{m \neq 0} \frac{\alpha_m^I}{m z^m}, \\ \bar{X}_L^I(\bar{z}) &= \bar{x}^I - i\sqrt{\frac{\alpha'}{2}} p_L^I \ln \bar{z} + i\sqrt{\frac{\alpha'}{2}} \sum_{m \neq 0} \frac{\bar{\alpha}_m^I}{m \bar{z}^m}. \end{aligned}$$

where the zero modes are

$$\begin{aligned}\mathbf{p}_L := p_L^I &= \frac{1}{\sqrt{2}} \left[\sqrt{\alpha'} p^I + \frac{W^I}{\sqrt{\alpha'}} \right] = \frac{1}{\sqrt{2}} \left[\sqrt{\alpha'} e^{*Ii} n_i + \frac{e_i^I}{\sqrt{\alpha'}} w^i \right], \\ \mathbf{p}_R := p_R^I &= \frac{1}{\sqrt{2}} \left[\sqrt{\alpha'} p^I - \frac{W^I}{\sqrt{\alpha'}} \right] = \frac{1}{\sqrt{2}} \left[\sqrt{\alpha'} e^{*Ii} n_i - \frac{e_i^I}{\sqrt{\alpha'}} w^i \right].\end{aligned}\quad (11.3)$$

The mass formula and the level-matching condition are now

$$\begin{aligned}\alpha' M^2 &= 2(N + \bar{N} - 2) + (\alpha' p_I p^I + \frac{1}{\alpha'} W_I W^I) \\ &= 2(N + \bar{N} - 2) + (\alpha' n_i n_j G^{ij} + \frac{1}{\alpha'} w^i w^j G_{ij}) \\ N - \bar{N} &= p_I W^I = n_i w^i\end{aligned}\quad (11.4)$$

As previously observed, the expressions for p_L and p_R suggest **T-duality** between the winding number W^I and the momentum p^I . In fact, **T-duality** corresponds to an equivalence between a pair of compactification lattices \mathbf{e}_i and \mathbf{e}'_i related as $\sqrt{\alpha'} \mathbf{e}'_i = \frac{\mathbf{e}^{*i}}{\sqrt{\alpha'}}$. These two compactifications yield the same spectrum since their allowed values of momenta are related as

$$\mathbf{p}_L \leftrightarrow \mathbf{p}'_L; \quad \mathbf{p}_R \leftrightarrow -\mathbf{p}'_R \quad (11.5)$$

by interchanging the labels n_i and w^i .

Now let us combine the zero modes into the $(d+d)$ -dimensional vectors $\mathbf{P} = (\mathbf{p}_L, \mathbf{p}_R)$. This construction treats Λ and Λ^* on equal footing as

$$\mathbf{P} = \mathbf{E}^{*i} n_i + \mathbf{E}_j w^j,$$

where

$$\mathbf{E}_j = \frac{1}{\sqrt{\alpha'}} (\mathbf{e}_j, -\mathbf{e}_j), \quad \mathbf{E}^{*i} = \sqrt{\alpha'} (\mathbf{e}^{*i}, \mathbf{e}^{*i}).$$

Note that the length of the lattice is normalized by the string length, given as $\sqrt{\alpha'} = \ell_s$. Therefore, \mathbf{P} takes values in a $(d+d)$ -dimensional lattice $\Gamma_{d,d}$ spanned by $\{\mathbf{E}^{*i}\}$ and $\{\mathbf{E}_j\}$. This lattice satisfies the following properties:

- **Lorentzian** if the signature of the metric G is $((+1)^d, (-1)^d)$,
- **integral** if $v \cdot w \in \mathbb{Z}$ for all $v, w \in \Gamma_{d,d}$,
- **even** if $\Gamma_{d,d}$ is integral and v^2 is even for all $v \in \Gamma_{d,d}$,
- **self-dual** if $\Gamma_{d,d} = (\Gamma_{d,d})^*$,
- **unimodular** if $\text{Vol}(\Gamma_{d,d}) = |\det G| = 1$.

In fact, the metric of this lattice is defined by

$$\mathbf{P} \cdot \mathbf{P}' = (\mathbf{p}_L \cdot \mathbf{p}'_L - \mathbf{p}_R \cdot \mathbf{p}'_R) = n_i w'^i + n'_i w^i$$

which confirms that it is Lorentzian. Since $\mathbf{P} \cdot \mathbf{P} \in 2\mathbb{Z}$, it is also even. The derivation of the self-duality property will be left as an exercise. The unimodular property $\text{Vol}(\Gamma_{d,d}) = \text{Vol}(\Gamma_{d,d}) = 1$ immediately follows from the self-dual property. The lattice $\Gamma_{d,d}$ in the torus compactification of the string is called **Narain lattice**.

The partition function of the bosonic string compactified on a torus T^d can be expressed as

$$Z_{\Gamma_{d,d}}^{\text{bos}} = \frac{1}{\tau_2^{(24-d)/2} |\eta(q)|^{48}} \sum_{(\mathbf{p}_R, \mathbf{p}_L) \in \Gamma_{d,d}} q^{\frac{1}{2} \mathbf{p}_R^2} \bar{q}^{\frac{1}{2} \mathbf{p}_L^2}$$

where $|\eta(q)|^{48}$ represents the contribution from the bosonic oscillators and $\tau_2^{(24-d)/2}$ arises from the integration over the non-compact momenta. This expression can be readily extended to the Type II string compactified on T^d

$$Z_{\Gamma_{d,d}}^{\text{Type II}} = \frac{1}{\tau_2^{(8-d)/2} |\eta(q)|^{24}} \frac{1}{4} \left| -\vartheta_2^4(\tau) + \vartheta_3^4(\tau) - \vartheta_4^4(\tau) \right|^2 \sum_{(\mathbf{p}_R, \mathbf{p}_L) \in \Gamma_{d,d}} q^{\frac{1}{2}} \mathbf{p}_R^2 \bar{q}^{\frac{1}{2}} \mathbf{p}_L^2$$

which vanishes by virtue of the Jacobi-Riemann identity.

Heterotic strings

After studying toroidal compactifications, we are now ready to introduce the $D = 10$ Heterotic string. As mentioned at the beginning, the Heterotic string combines the left-moving sector of the 26-dimensional bosonic string with the right-moving sector of the 10-dimensional superstring. The left-moving bosonic string is compactified on a 16-dimensional torus, where the momenta of the additional chiral bosons $X^I(\bar{z})$ take values on a 16-dimensional lattice Γ_{16} . Thus, the partition function of the Heterotic string can be expressed as

$$Z^{\text{het}}(\tau) = \frac{1}{\tau_2^4 \eta(q)^{12} \eta(\bar{q})^{24}} \left(-\vartheta_2^4(\tau) + \vartheta_3^4(\tau) - \vartheta_4^4(\tau) \right) \sum_{\mathbf{p}_L \in \Gamma_{16}} \bar{q}^{\frac{1}{2}} \mathbf{p}_L^2 \quad (11.6)$$

Here, $\eta(q)^8 \eta(\bar{q})^{24}$ represents the contribution from the bosonic oscillators, the τ_2^4 factor arises from the zero modes of the noncompact transverse coordinates, and $\vartheta_i^4 / \eta(q)^4$ comes from the world-sheet fermions. The most interesting part of this partition function is the lattice sum

$$P(\bar{\tau}) := \sum_{\mathbf{p}_L \in \Gamma_{16}} \bar{q}^{\frac{1}{2}} \mathbf{p}_L^2$$

In order for the partition function (11.6) to be invariant under the modular transformation $\text{SL}(2, \mathbb{Z})$, the modular transformations of η and ϑ_i reveal the following properties:

$$T : P(\bar{\tau} + 1) = P(\bar{\tau}), \quad S : P(-1/\bar{\tau}) = \bar{\tau}^8 P(\bar{\tau}).$$

The invariance under the T -transformation implies that $\mathbf{p}_L^2 \in 2\mathbb{Z}$, requiring the lattice Γ_{16} to be **even**. For the S -transformation, we employ the Poisson resummation formula

$$\sum_{\mathbf{p} \in \Lambda} e^{-\pi\alpha(\mathbf{p}+\mathbf{x})^2 + 2\pi i \mathbf{y} \cdot (\mathbf{p}+\mathbf{x})} = \frac{1}{\text{Vol}(\Lambda)\alpha^{\dim \Lambda/2}} \sum_{\mathbf{q} \in \Lambda^*} e^{-2\pi i \mathbf{q} \cdot \mathbf{x} - \frac{\pi}{\alpha}(\mathbf{y}+\mathbf{q})^2}$$

which amounts to

$$P(-1/\bar{\tau}) = \frac{\bar{\tau}^8}{\text{Vol}(\Gamma_{16})} \sum_{\mathbf{p}_L \in (\Gamma_{16})^*} \bar{q}^{\frac{1}{2}} \mathbf{p}_L^2.$$

This requires that the lattice Γ_{16} is **self-dual**, *i.e.* $(\Gamma_{16})^* = \Gamma_{16}$ so that $\text{Vol}(\Lambda) = 1$.

It turns out that there are only two even self-dual Euclidean lattices in 16 dimensions

- the root lattice of $E_8 \times E_8$
- the weight lattice of $\text{Spin}(32)/\mathbb{Z}_2$

The metric G_{ij} of the root lattice of E_8 is the Cartan matrix of E_8 ⁶:

$$\begin{pmatrix} 2 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ -1 & 2 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 2 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 2 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 2 & -1 & 0 & -1 \\ 0 & 0 & 0 & 0 & -1 & 2 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 & 2 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 & 0 & 2 \end{pmatrix}, \quad \text{Diagram: } \circ-\circ-\circ-\circ-\circ-\circ-\circ-\circ$$

Let us examine the massless fields in Heterotic string theory in more detail. As usual, there is the tachyonic vacuum of the bosonic string. At the massless level, we have oscillator excitations $\bar{\alpha}_{-1}^\mu |0\rangle$, $\bar{\alpha}_{-1}^I |0\rangle$ in the left-moving sector. The former transform like spacetime vectors while the internal oscillator excitations correspond to the left-moving part of the Abelian $U(1)^{16}$ gauge boson. They form the **Cartan subalgebra** of $E_8 \times E_8$ or $SO(32)$. Since the mass formula is given by

$$\alpha' M^2 = 2(N + \bar{N} - 2) + \mathbf{p}_L^2, \quad (11.7)$$

additional massless modes come from $\mathbf{p}_L^2 = 2$ for $N = 1, \bar{N} = 0$. Since Both the root lattice of $E_8 \times E_8$ and the weight lattice of $Spin(32)/\mathbb{Z}_2$ contain 480 vectors of (length)² = 2 and generate the 496-dimensional non-Abelian gauge bosons of these groups. It is remarkable that, despite Heterotic strings being closed strings, gauge fields emerge due to the presence of the extra 16-dimensional torus. This phenomenon can be understood as a unique **stringy effect**, and the gauge groups are restricted to either $E_8 \times E_8$ or $SO(32)$ for the consistency of the theory.

Moreover, we have observed that the $SO(32)$ gauge group arises in Type I string theory. This connection is not coincidental, as we will explore in §13. It turns out that Type I and Heterotic $SO(32)$ strings are related through **S-duality**, further highlighting the profound interplay between different string theories.

As a result, the massless spectra of Heterotic string are as follows

- Gravitons, B -fields, dilaton in $D = 10$

$$\psi_{-\frac{1}{2}}^\mu |0\rangle_{NS} \otimes \bar{\alpha}_{-1}^\nu |0\rangle$$

- their supersymmetric partners, gravitino and dilatino

$$|\mathbf{s}\rangle_R \otimes \bar{\alpha}_{-1}^\nu |0\rangle$$

- 496 gauge bosons of $E_8 \times E_8$ or $SO(32)$

$$\psi_{-\frac{1}{2}}^\mu |0\rangle_{NS} \otimes \bar{\alpha}_{-1}^I |0\rangle, \quad \psi_{-\frac{1}{2}}^\mu |0\rangle_{NS} \otimes |\mathbf{p}_L^2 = 2\rangle$$

- 496 supersymmetric partners, gaugini

$$|\mathbf{s}\rangle_R \otimes \bar{\alpha}_{-1}^I |0\rangle, \quad |\mathbf{s}\rangle_R \otimes |\mathbf{p}_L^2 = 2\rangle$$

Indeed, Heterotic string theory is $D = 10 \mathcal{N} = 1$ supergravity coupled to $D = 10 \mathcal{N} = 1$ $E_8 \times E_8$ or $SO(32)$ super-Yang-Mills theory so that it has 16 real supersymmetric charges.

⁶Due to the limited scope of our discussion, we are unable to delve into exceptional Lie algebras or the classification of semi-simple Lie algebras in detail [Kir08]. However, if you would like to develop some intuition about weight and root lattices, I recommend referring to Figure 7.3, Figure 8.1, and Figure 8.2 in [Kir08] for the example of A_2 . For a deeper understanding of the structure of E_8 in the context of string theory, I suggest exploring Section 6 of [GSW87].

11.2 Fermionic construction

The 16 bosonic fields compactified on the self-dual lattice can be described by fermionic fields, which is called **fermionization**. Therefore, let us delve into the fermionic construction of Heterotic string theory.

The world-sheet action of Heterotic string theory is given by the sum of the matter and ghost actions:

$$\begin{aligned} S^m &= \frac{1}{4\pi} \int d^2z \left(\frac{2}{\alpha'} \partial X^\mu \bar{\partial} X_\mu + \psi^\mu \bar{\partial} \psi_\mu + \bar{\lambda}^A \partial \bar{\lambda}_A \right) \\ S^g &= \frac{1}{2\pi} \int d^2z (b \bar{\partial} c + \bar{b} \partial \bar{c} + \beta \bar{\partial} \gamma) \end{aligned} \quad (11.8)$$

where μ represents the 10-dimensional indices and the right-moving sector is supersymmetric. For the theory to be Weyl-anomaly free, the total central charge c^{tot} must vanish:

$$c^{\text{tot}} = c^X + c^\psi + c^{bc} + c^{\beta\gamma} + c^\lambda = 10 + \frac{5}{2} - 26 + \frac{11}{2} + c^\lambda = c^\lambda - 8$$

Since each Majorana-Weyl anti-chiral fermion $\bar{\lambda}^A$ contributes $\frac{1}{4}$ to the central charge, we need 32 left-moving fermions $\bar{\lambda}^A$ in the action.

Interestingly, there are two possible boundary conditions for the left-moving fermions $\bar{\lambda}_A$ that yield consistent string theories. If we impose the same boundary condition on all fermions, we obtain the Heterotic SO(32) gauge group. On the other hand, if we impose one boundary condition on half of the fermions and the opposite boundary condition on the other half, we obtain the Heterotic $E_8 \times E_8$ gauge group.

Heterotic SO(32) (HO)

In the case of Heterotic SO(32) (HO), the same boundary condition is imposed on all left-moving fermions

$$\begin{aligned} \bar{\lambda}^A(t, \sigma + 2\pi) &= +\bar{\lambda}^A(t, \sigma) && \text{R: periodic on cylinder} \\ \bar{\lambda}^A(t, \sigma + 2\pi) &= -\bar{\lambda}^A(t, \sigma) && \text{NS: anti-periodic on cylinder} \end{aligned} \quad (11.9)$$

This leads to a global SO(32) symmetry that rotates the $\bar{\lambda}^A$ fermions ($A = 1, \dots, 32$). In order for the theory to be consistent, we need to impose the GSO projection on the left-moving sector. In Heterotic string theory, we select states with odd fermionic numbers in the NS sector and states with even fermionic numbers in the R sector.

$$\begin{aligned} P_{\text{NS}}^H &:= \frac{1 - (-1)^F}{2}, \\ P_{\text{R}}^H &:= \frac{1 + (-1)^F}{2}. \end{aligned} \quad (11.10)$$

Additionally, we must impose the level matching condition:

$$N - a = \bar{N} - \bar{a} \quad (11.11)$$

where the normal ordering constants in the left-moving sector are

$$\bar{a}_{\text{NS}} = \frac{8}{24} + \frac{32}{48} = 1, \quad \bar{a}_{\text{R}} = \frac{8}{24} - \frac{32}{24} = -1.$$

Here the first term comes from the left-moving bosonic field \bar{X}^i whereas the second term depends on the boundary condition (7.13) of $\bar{\lambda}^A$. As a result, the R sector contains only massive states.

Under the GSO projection, the Tachyon state $|0\rangle_{\text{NS}}$ in the NS sector is preserved, but there is no corresponding state in the right-moving sector. Therefore, it does not satisfy the level matching condition (11.11), and the left-moving Tachyon is excluded from the spectrum. Then, the first excited states after the GSO projection in the NS sector are

$$\begin{aligned}\bar{\alpha}_{-1}^i |0\rangle_{\text{NS}}, & \quad (\mathbf{8}_v, \mathbf{1}) \\ \bar{\lambda}_{-1/2}^A \bar{\lambda}_{-1/2}^B |0\rangle_{\text{NS}}, & \quad (\mathbf{1}, \mathbf{adj})\end{aligned}$$

where the bold letters denote the representations of $\text{SO}(8) \times \text{SO}(32)$. The adjoint representation **adj** of $\text{SO}(32)$ corresponds to the antisymmetric representation with dimension $32 \times 31/2 = \mathbf{496}$.

The following table summarizes the massless spectrum of Heterotic $\text{SO}(32)$ (HO) theory. The first row represents the $D = 10 \mathcal{N} = 1$ supergravity multiplet, while the second row shows the $\mathcal{N} = 1$ gauge multiplet in the adjoint representation of $\text{SO}(32)$, as we have seen in the bosonic construction.

Left \ Right	$\mathbf{8}_v$	$\mathbf{8}_s$
$(\mathbf{8}_v, \mathbf{1})$	$\mathbf{1} \oplus \mathbf{28} \oplus \mathbf{35}$ $\phi \ B_{\mu\nu} \ G_{\mu\nu}$	$\mathbf{8}_c \oplus \mathbf{56}_s$ $\lambda^+ \ \psi_m^-$
$(\mathbf{1}, \mathbf{496})$	$\text{SO}(32)$ gauge boson $A_{[A,B]}^\mu$	$\text{SO}(32)$ gaugini $\eta_{[A,B]}$

Table 5: Massless field contents of Heterotic string theory

Heterotic $E_8 \times E_8$ (HE)

The second Heterotic string theory is obtained by dividing the $\bar{\lambda}^A$ into two sets of 16 with independent boundary conditions,

$$\bar{\lambda}^A(t, \sigma + 2\pi) = \begin{cases} \epsilon_1 \bar{\lambda}^A(t, \sigma) & A = 1, \dots, 16 \\ \epsilon_2 \bar{\lambda}^A(t, \sigma) & A = 17, \dots, 32 \end{cases}$$

where $\epsilon_i = \pm 1$. Therefore, in the left-moving sector, we need to take the following boundary conditions into account

$$(\text{NS}_1, \text{NS}_2), \quad (\text{R}_1, \text{NS}_2), \quad (\text{NS}_1, \text{R}_2), \quad (\text{R}_1, \text{R}_2).$$

Consequently, the global symmetry is broken to $\text{SO}(16)_1 \times \text{SO}(16)_2$. The GSO projection is applied independently to the two sets of left-movers. We also apply for the level-matching condition (11.11). The normal ordering constant in each boundary condition is

$$\bar{a}_{\text{NS}_1, \text{NS}_2} = 1, \quad \bar{a}_{\text{R}_1, \text{NS}_2} = \bar{a}_{\text{NS}_1, \text{R}_2} = \frac{8}{24} + \frac{16}{48} - \frac{16}{24} = 0, \quad \bar{a}_{\text{R}_1, \text{R}_2} = -1.$$

Similarly to Heterotic $\text{SO}(32)$, the (R_1, R_2) boundary condition contains only massive states. Although the Tachyon state $|0\rangle_{\text{NS}_1, \text{NS}_2}$ in the NS sector is preserved under the GSO

projection, it does not satisfy the level-matching condition (11.11), so it is not present in the spectrum.

As a result, the massless states are then given by

$$\begin{aligned} \bar{\alpha}_{-1}^i |0\rangle_{\text{NS}_1, \text{NS}_2}, & \quad (\mathbf{8}_v, \mathbf{1}, \mathbf{1}) \\ \bar{\lambda}_{-1/2}^A \bar{\lambda}_{-1/2}^B |0\rangle_{\text{NS}_1, \text{NS}_2}, & \quad (\mathbf{1}, \mathbf{adj}, \mathbf{1}) \text{ or } (\mathbf{1}, \mathbf{1}, \mathbf{adj}) \end{aligned} \quad (11.12)$$

where the bold letters are the representations of $\text{SO}(8) \times \text{SO}(16)_1 \times \text{SO}(16)_2$. Note that the GSO projection requires either $1 \leq A, B \leq 16$ or $17 \leq A, B \leq 32$ in (11.12). The adjoint representation of $\text{SO}(16)$ is of $16 \times 15/2 = \mathbf{120}$ dimensions.

In the (R_1, NS_2) and (NS_1, R_2) boundary conditions, the ground states are massless due to the zero normal ordering constant. The 16 $\bar{\lambda}_0^A$ zero modes can be expressed in terms of 8 raising and 8 lowering operators:

$$\bar{\lambda}_0^{K\pm} = 2^{-1/2} (\bar{\lambda}_0^{2K-1} \pm i \bar{\lambda}_0^{2K}), \quad K = 1, \dots, 8 \text{ or } K = 9, \dots, 16,$$

As a result, the $2^8 = \mathbf{256}$ -dimensional spinor representation of $\text{SO}(16)$ becomes massless. However, the GSO projection selects the positive chirality $\mathbf{128}$ out of $\mathbf{256} = \mathbf{128} + \mathbf{128}'$ in the Ramond sector. Therefore, the ground states in the (R_1, NS_2) and (NS_1, R_2) sectors are $(\mathbf{1}, \mathbf{128}, \mathbf{1})$ and $(\mathbf{1}, \mathbf{1}, \mathbf{128})$, respectively, under $\text{SO}(8) \times \text{SO}(16)_1 \times \text{SO}(16)_2$.

All in all, the left-moving massless states form the representations of $\text{SO}(8) \times \text{SO}(16)_1 \times \text{SO}(16)_2$

$$(\mathbf{8}_v, \mathbf{1}, \mathbf{1}) + (\mathbf{1}, \mathbf{120}, \mathbf{1}) + (\mathbf{1}, \mathbf{1}, \mathbf{120}) + (\mathbf{1}, \mathbf{128}, \mathbf{1}) + (\mathbf{1}, \mathbf{1}, \mathbf{128})$$

This spectrum strongly suggests that gauge symmetry is enhanced $\text{SO}(16) \rightarrow E_8$ because the dimension of the adjoint representation E_8 is $\mathbf{120} + \mathbf{128} = \mathbf{248}$. In fact, the $\text{SO}(16)$ subgroup of E_8 acts on the $\mathbf{248}$ adjoint representation as $\mathbf{120} + \mathbf{128}$. Therefore, the massless spectrum of the Heterotic string theory includes the 496-dimensional adjoint representations of both $\text{SO}(32)$ and $E_8 \times E_8$ gauge groups.

In summary, the fermionic construction of the Heterotic string theory reproduces the massless spectrum consisting of gravity, gauge bosons, and gauginos, which is consistent with the bosonic construction.

No supersymmetric D-branes in Heterotic strings

We have seen that D-branes are charged to R-R fields in Type II theories. However, there is no R-R field in Heterotic string theories because there is only world-sheet supersymmetry in the right-moving sector. In other words, although the R-R $(p+2)$ -form field strength $G_{(p+2)}$ in Type II theories can be expressed as

$$G_{(p+2)} \longleftrightarrow \bar{\psi}^L \Gamma^{\mu_1 \dots \mu_{p+2}} \psi^R,$$

there is no left-moving fermionic field $\bar{\psi}^L$ present. Therefore, the R-R $(p+2)$ -form field strength $G_{(p+2)}$ cannot be expressed in terms of fermionic fields in Heterotic string theories. Hence, there is no D-brane in Heterotic string theories. Consequently, Heterotic string theories are the theories of closed strings⁷. Despite the absence of D-branes, Heterotic string theories do feature extended objects known as NS5-branes or Heterotic five-branes in Heterotic string theories and they are magnetically charged under the B -field.

⁷However, Polchinski discusses open Heterotic strings in [Pol06].

12 Supergravity

String theory includes massless states as well as massive states. However, in the energy regions much lower than $1/\ell_s$, we do not see the stringy massive states because the mass $M^2 \sim \frac{1}{\alpha'} \sim \frac{1}{\ell_s^2}$ is assumed to be very heavy. Therefore, in the low-energy region, we can describe the theory by an **effective theory**, which only contains the massless particles (lightest states). The effective theory, of course, does not contain all the information of the original theory, however, it does give us some information about the original theory.

For the bosonic string theory, the low-energy effective action for the massless fields is given in (4.8). Below, we deal with its supersymmetric versions, type **IIA/IIB supergravity**, that are low-energy effective theories of IIA/IIB superstring theory. In §13.2, we will deal with the low-energy effective actions of Type I and Heterotic string theory. The theory of supergravity is very broad, and we refer to [FVP12] for more detail.

12.1 Local supersymmetry

Before going to Type IIA/IIB supergravity, let us see the basic idea of supergravity. Roughly speaking, a global supersymmetry algebra takes the form

$$\{\epsilon_2 Q, \bar{\epsilon}_1 \bar{Q}\} = i\bar{\epsilon}_1 \Gamma^\mu \epsilon_2 P_\mu . \quad (12.1)$$

(For instance, see (17.10).) The key idea in supergravity is that supersymmetry holds locally so that the spinor parameters ϵ are arbitrary functions of the spacetime coordinates $\epsilon \rightarrow \epsilon(x)$. The supersymmetry algebra will then involve local translation parameters $\bar{\epsilon}_1 \gamma^\mu \epsilon_2$ which must be viewed as diffeomorphisms. Thus, local supersymmetry requires gravity. Note that we use x^μ for local coordinates of the spacetime instead of X^μ , which has been adopted in the previous sections.

First of all, in order to define spinors in a curved space, we need to introduce vielbeins $e_\mu^a(x^\rho)$. The metric can be the orthonormal one at one point by coordinate transformations, and vielbeins transforms a local coordinate x^μ into the orthonormal coordinate at the point:

$$x^a = e_\mu^a x^\mu , \quad x^\mu = e_a^\mu x^a ,$$

which is defined through spacetime metric G_{MN} by

$$G_{\mu\nu}(x^\rho) = \eta_{ab} e_\mu^a(x^\rho) e_\nu^b(x^\rho) .$$

Now we can use spinor representations and gamma matrices thanks to the vielbeins:

$$\{\Gamma^a, \Gamma^b\} = 2\eta^{ab} .$$

If we include gravity in a theory, the translational invariance is promoted to invariance under general coordinate transformations. Similarly, global Lorentz symmetry is promoted to local Lorentz symmetry (this is done by the vielbeins):

$$\Lambda_b^a e_a^\mu(x^\rho) e_\nu^b(x^\rho) \equiv \Lambda_\nu^\mu(x^\rho) .$$

In the orthonormal basis, the covariant derivative is expressed by the spin connection

$$\begin{aligned} D_\mu V^a &= \partial_\mu V^a + \omega_\mu{}^a{}_b V^b \\ D_\mu \psi &= \left(\partial_\mu + \frac{1}{4} \omega_\mu^{ab} \Gamma_{ab} \right) \psi \end{aligned} \quad (12.2)$$

Gravity can be interpreted as the gauge field for general coordinate transformations. Since the gauge transformation of a spin-1 gauge field can be generally written as

$$\delta A_\mu = D_\mu \lambda , \quad (12.3)$$

the gauge transformation of the gravity field will take the form

$$\delta \omega_\mu{}^a{}_b = D_\mu \Theta^a{}_b , \quad (12.4)$$

where $\Theta^a{}_b$ is a general coordinate transformation, and D_μ is a covariant derivative.

Similarly, we need to introduce a gauge field for local supersymmetry in supergravity, and the corresponding field is a **gravitino**. (Recall that gravitino appears as a massless field in superstring theory.) Then, the supersymmetry transformations for the gravitino and vielbeins are given by

$$\delta \psi_\mu^\alpha = D_\mu \epsilon^\alpha , \quad \delta_\epsilon e_\mu^a = i\bar{\epsilon} \Gamma^a \psi_\mu . \quad (12.5)$$

Then, the square of the supersymmetry transformation is as we want:

$$\begin{aligned} [\delta_1, \delta_2] e_\mu^a &= iD_\mu (\bar{\epsilon}_1 \Gamma^a \epsilon_2) \\ [\delta_1, \delta_2] \psi_\mu &= i\bar{\epsilon}_1 \Gamma^\rho \epsilon_2 (D_\rho \psi_\mu - D_\mu \psi_\rho) . \end{aligned} \quad (12.6)$$

Supergravity theory always includes the two gauge fields, e_μ^a and ψ_μ^α , and the action is given by

$$S = \frac{1}{2\kappa_D^2} \int d^D x e [R - 2i\psi_\mu \Gamma^{\mu\nu\rho} D_\nu \psi_\rho] , \quad (12.7)$$

where $e = \det e_\mu^a$. It is straightforward to check that the action is invariant under local supersymmetry transformation (12.5). This action can be regarded as the supersymmetric version of the Einstein-Hilbert action.

12.2 $D = 11$ supergravity

Supersymmetry puts a strong constraint on the spacetime dimension. If we limit ourselves to consider fields up to spin-2, then it is known that the highest dimension is $D = 11$ [Nah78]. Roughly speaking, this is because the degrees of freedom for fermions grow exponentially as $2^{[D/2]}$ whereas those of bosons grow by the power law of D (for instance, $\frac{(D-1)(D-2)}{2} - 1$ for gravitons). Therefore, to balance fermions and bosons with spin less than two, we cannot go arbitrarily higher.

The $D = 11$ supergravity [CJS78] consists of three fields;

- one is the graviton G_{MN} (44 states)
- three-form field $M_{(3)}$ (84 states)
- gravitino ψ_M (128 states)

We can see that the numbers of fermions and bosons are balanced. Although the existence of fermions is crucial, we write the bosonic part of the action of the $D = 11$ supergravity:

$$2\kappa_{11}^2 S_{11} = \int d^{11}x \sqrt{-G} \left[R - \frac{1}{2} K_{(4)}^2 \right] - \frac{1}{6} \int d^{11}x M_{(3)} \wedge K_{(4)} \wedge K_{(4)} , \quad (12.8)$$

where $K_{(4)} = dM_{(3)}$ is the field strength, and $K_{(4)}^2 = K_{(4)} \wedge *K_{(4)}$. The fermionic part of the action follows from supersymmetry in principle. Note that there is only one parameter κ_{11} defined as

$$\frac{1}{2\kappa_{11}^2} = \frac{2\pi}{(2\pi\ell_p)^9}, \quad (12.9)$$

which is written in terms of Planck length

$$\ell_P = \sqrt{\frac{\hbar G}{c^3}}.$$

The critical dimension of superstring theory is $D = 10$. Nonetheless, Type IIA $D = 10$ supergravity can be obtained from the dimensional reduction of the $D = 11$ supergravity on S^1 , as we see below. In fact, about the $D = 11$ supergravity, it was mentioned in [GSW87, §13.1.1] as follows:

The ten-dimensional supergravity theories are, of course, the low-energy limits of certain string theories. Further dimensional reduction can give a variety of supergravity theories in various dimensions less than ten. If string theory proves to be correct, it could be regarded as 'explaining' the existence of supergravity theories for $D \leq 10$. Eleven-dimensional supergravity remains an enigma. It is hard to believe that its existence is just an accident, but it is difficult at the present time to state a compelling conjecture for what its role may be in the scheme of things.

The crucial understanding of this enigma was provided in [Wit95b] where the $D = 11$ supergravity indeed gives an important clue to the strong coupling behavior of superstring theory.

12.3 $D = 10$ IIA supergravity

Now we compactify the $D = 11$ supergravity on S^1 of radius R . In the following, $C_{(p+1)}$ is the R-R $(p+1)$ -form and $B_{(2)}$ is the B -field. Also, $G_{(p+2)}$ and $H_{(3)}$ are their field strengths. On the reduction, the three-form field $M_{(3)}$ decomposes into the massless fields of Type IIA string theory in Table 2:

$$M_{(3)} = C_{(3)} + B_{(2)} \wedge d\theta, \quad K_{(4)} = \hat{G}_{(4)} + H_{(3)} \wedge (d\theta + C_{(1)}), \quad (12.10)$$

where

$$\hat{G}_{(4)} = G_{(4)} - C_{(1)} \wedge H_{(3)}. \quad (12.11)$$

Like (9.7), the metric takes the form

$$ds_{11}^2 = G_{MN} dx^M dx^N = G_{\mu\nu} dx^\mu dx^\nu + R^2 (d\theta + C_{(1)})^2,$$

where the compactified S^1 direction is denoted as θ .

Now we rewrite the action (12.8) in terms of the massless fields of Type IIA theory. (See the left of Table 2.) However, we want to obtain the action of the form (4.10) in the NS-NS sector. For this purpose, we consider the radius of the circle depends on the spacetime coordinate and it can be written as the Dilaton fields as

$$R = \ell_p e^{\frac{2}{3}\Phi}. \quad (12.12)$$

Then, the action (12.8) is written as

$$\begin{aligned} S_{A,NS} &= \frac{1}{2\kappa_{10}^2} \int d^{10}x \sqrt{-G} e^{-2\Phi} \left[R + 4\partial_\mu \Phi \partial^\mu \Phi - \frac{1}{2} H_{(3)}^2 \right] , \\ S_{A,R} &= -\frac{1}{4\kappa_{10}^2} \int d^{10}x \sqrt{-G} \left[G_{(2)}^2 + \hat{G}_{(4)}^2 \right] , \\ S_{A,CS} &= -\frac{1}{4\kappa_{10}^2} \int B_{(2)} \wedge G_{(4)} \wedge G_{(4)} , \end{aligned} \quad (12.13)$$

where the coupling constants are related by

$$\frac{2\pi R}{2\kappa_{11}^2} = \frac{e^{-2\Phi}}{2\kappa_{10}^2} . \quad (12.14)$$

Since it is the low-energy effective action of Type IIA string theory, the dimensional analysis of the coupling constant yields

$$\frac{1}{2\kappa_{10}^2} = \frac{2\pi}{(2\pi\ell_s)^8} . \quad (12.15)$$

Recalling that an expectation value of the dilaton is the string coupling $g_s = e^\Phi$, (12.12) gives

$$\left(\frac{R}{\ell_p} \right)^3 = g_s^2 . \quad (12.16)$$

Also, (12.14) provides the relation

$$\frac{R}{\ell_p^3} = \frac{1}{\ell_s^2} . \quad (12.17)$$

Combining these two relations, we have

$$R = g_s \ell_s . \quad (12.18)$$

12.4 $D = 10$ IIB supergravity

As seen in §9, Type IIA and IIB are T-dual to each other so that the action of Type IIB supergravity is consistent with the T-duality. The right of Table 2 lists all the massless fields, and the supergravity action indeed takes a very similar form

$$\begin{aligned} S_{B,NS} &= \frac{1}{2\kappa_{10}^2} \int d^{10}x \sqrt{-G} e^{-2\Phi} \left[R + 4\partial_\mu \Phi \partial^\mu \Phi - \frac{1}{2} H_{(3)}^2 \right] , \\ S_{B,R} &= -\frac{1}{4\kappa_{10}^2} \int d^{10}x \sqrt{-G} \left[G_{(1)}^2 + \hat{G}_{(3)}^2 + \frac{1}{2} \hat{G}_{(5)}^2 \right] , \\ S_{B,CS} &= -\frac{1}{4\kappa_{10}^2} \int C_{(4)} \wedge H_{(3)} \wedge G_{(3)} , \end{aligned} \quad (12.19)$$

where

$$\begin{aligned} \hat{G}_{(3)} &= G_{(3)} - C_{(0)} H_{(3)} , \\ \hat{G}_{(5)} &= G_{(5)} - \frac{1}{2} C_{(2)} \wedge H_{(3)} + \frac{1}{2} B_{(2)} \wedge G_{(3)} . \end{aligned} \quad (12.20)$$

Note that the action for the NS sector is the same as that of Type IIA.

As in (8.5), Type IIB supergravity has the self-dual 5-form $G_{(5)}$. In the notation here, the self-dual condition is

$$*\widehat{G}_{(5)} = \widehat{G}_{(5)},$$

which does not follow from the action. In fact, if the field satisfies the self-dual condition, its kinetic action becomes trivial

$$\int \widehat{G}_{(5)} \wedge *\widehat{G}_{(5)} = 0.$$

Therefore, the self-duality condition needs to be imposed by hand.

As we will see below, Type IIB supergravity exhibits a hidden symmetry that manifests as the S -duality in Type IIB string theory. Moreover, Type IIB supergravity is subject to two types of \mathbb{Z}_2 actions. The first acts on the bosonic fields as follows:

$$G_{(1,5)} \rightarrow -G_{(1,5)}, \quad H_{(3)} \rightarrow -H_{(3)}. \quad (12.21)$$

The second \mathbb{Z}_2 action changes the sign of all the R-R field strengths:

$$G_{\text{odd}} \rightarrow -G_{\text{odd}}. \quad (12.22)$$

Both \mathbb{Z}_2 actions identify two copies of the dilatino and gravitino in Type IIB theory (see Table 2). The first \mathbb{Z}_2 projection results in Type I supergravity, while the second leads to Heterotic supergravity.

13 String dualities

We have introduced superstring theories of five types: Type IIA, IIB, I, and Heterotic $\text{SO}(32)$ and $E_8 \times E_8$. In fact, the seminal paper [Wit95b] reveals that these string theories are related by dualities, which led to the second string revolution. Indeed, we have already learned

- Type IIA and IIB are T-dual to each other
- Type I is the orientifold projection of Type IIB

In this section, we will learn string dualities extensively studied in the second string revolution after [Wit95b].

- Type IIB has $\text{SL}(2, \mathbb{Z})$ symmetry so that it is self-dual under S-duality
- Type I to Heterotic $\text{SO}(32)$ is S-dual
- Heterotic $\text{SO}(32)$ and $E_8 \times E_8$ are T-dual to each other
- The strong coupling regime of Type IIA is described by M-theory on S^1
- The strong coupling regime of Heterotic $E_8 \times E_8$ is described by M-theory on S^1/\mathbb{Z}_2
- Heterotic string on T^4 is dual to Type IIA on K3

Even for these dualities, we can cover only key points in this section. More details can be found in [Pol98, BBS06]. Moreover, we just see the tip of the iceberg, and there are much more string dualities. Thus, we refer to good reviews [Asp96, FL98, OY96,

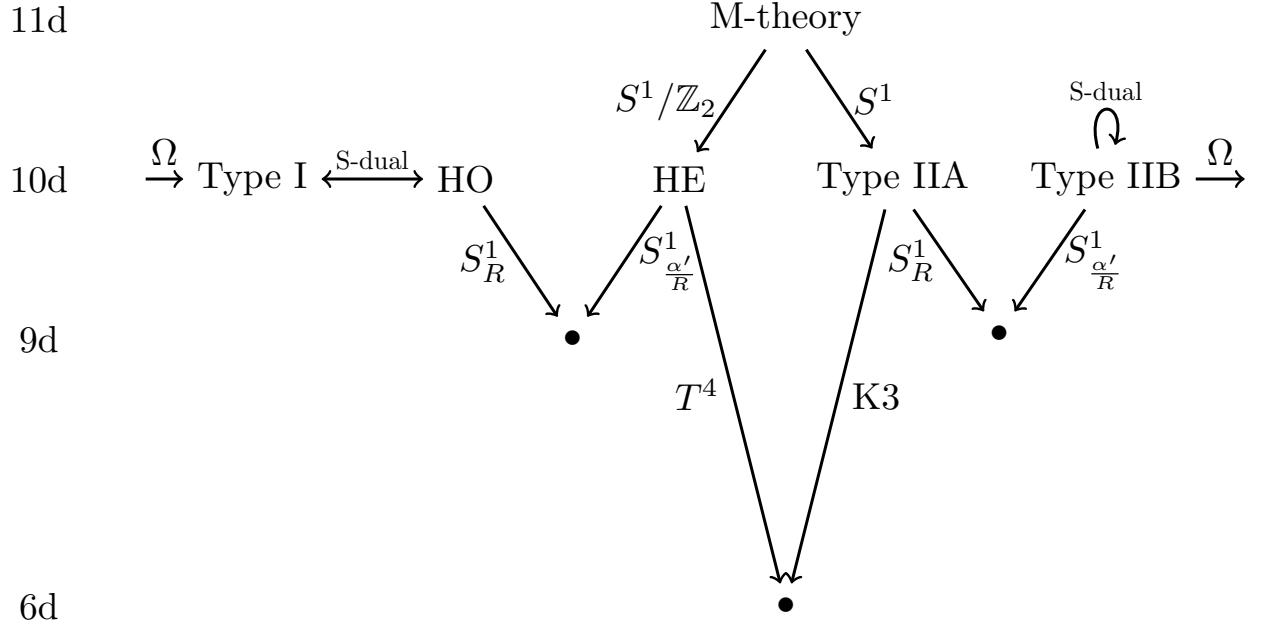


Figure 36: Duality web of string theory

[Pol96a](#), [Pol96b](#), [Tow96](#), [Sch97](#), [Sen97](#), [Dij97](#), [Vaf97](#), [Sen98a](#), [Joh00](#)] written during the second string revolution for this rich subject. All in all, these dualities tell us that quantum strings somehow see geometry from drastically different viewpoints. I hope you will get some feeling of it in this section.

13.1 S-duality of Type IIB supergravity

The electromagnetic duality seen in (8.9) and (8.10) is a basic example of duality. The generalization to the Yang-Mills theory is proposed by Goddard-Nuyts-Olive [[GNO77](#)] and Montonen-Olive [[MO77](#)]. The Montonen-Olive duality exchanges the Yang-Mills coupling by

$$g_{\text{YM}} \leftrightarrow 1/g_{\text{YM}}, \quad (13.1)$$

and it also exchanges a fundamental particle into a soliton. Thus, this duality is also called the **strong-weak duality** or **S-duality**. Although standard techniques in quantum field theory cannot be applied to the strong coupling regime, the S-duality provides new insights into non-perturbative dynamics in the strong-coupling regime. One of the reasons why string theory sheds new light on physical theories is that the strong-weak dualities show up in string theory [[Sen94](#), [Duf95](#)], which often have geometric origins.

In fact, Type IIB string theory enjoys the S-duality. To see it, let us rewrite the action by rescaling the metric $G_{E,\mu\nu} = e^{-\Phi/2} G_{\mu\nu}$, and by introducing

$$\tau = C_{(0)} + ie^{-\Phi}, \quad (13.2)$$

and

$$\mathbb{M} = \frac{1}{\text{Im}\tau} \begin{pmatrix} |\tau|^2 & -\text{Re}\tau \\ -\text{Re}\tau & 1 \end{pmatrix}, \quad \mathbb{F}_{(3)} = \begin{pmatrix} H_{(3)} \\ G_{(3)} \end{pmatrix},$$

Then, the action is rewritten as

$$S_B = \frac{1}{2\kappa_{10}^2} \int d^{10}x \sqrt{-G_E} \left[R_E - \frac{\partial_\mu \tau \partial^\mu \bar{\tau}}{2(\text{Im}\tau)^2} - \frac{1}{2} \mathbb{F}_{(3)}^T \cdot \mathbb{M} \cdot \mathbb{F}_{(3)} - \frac{1}{4} \tilde{G}_{(5)}^2 \right] - \frac{1}{4\kappa_{10}^2} \int \epsilon_{ij} C_{(4)} \wedge \mathbb{F}_{(3)}^i \wedge \mathbb{F}_{(3)}^j , \quad (13.3)$$

where $\epsilon = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$. Remarkably, this action is invariant under the $\text{SL}(2, \mathbb{R})$ transformation:

$$\tau' = \frac{a\tau + b}{c\tau + d}, \quad \mathbb{M}' = (\Lambda^{-1})^T \mathbb{M} \Lambda^{-1}, \quad \mathbb{F}'_{(3)} = \Lambda \mathbb{F}_{(3)}, \quad \Lambda = \begin{pmatrix} d & c \\ b & a \end{pmatrix} \in \text{SL}(2, \mathbb{R}).$$

On the other hand, the metric G_E and $C_{(4)}$ are invariant under the transformation.

Let us consider its physical implications. Since $C_{(4)}$ is invariant, a D3-brane is invariant. On the other hand, since the 2-form fields $B_{(2)}$ and $C_{(2)}$ are transformed by $\text{SL}(2, \mathbb{R})$, the extended objects coupled to these fields are transformed accordingly. Namely, F1 and D1-branes are electrically coupled to the 2-form fields while NS5 and D5-branes are magnetically coupled to them, respectively. Thus, they are transformed under $\text{SL}(2, \mathbb{R})$ as

$$(F1' \quad D1') = (F1 \quad D1) \Lambda^{-1}, \quad \begin{pmatrix} \text{NS5}' \\ \text{D5}' \end{pmatrix} = \Lambda \begin{pmatrix} \text{NS5} \\ \text{D5} \end{pmatrix}, \quad \text{as} \quad \begin{pmatrix} B'_{(2)} \\ C'_{(2)} \end{pmatrix} = \Lambda \begin{pmatrix} B_{(2)} \\ C_{(2)} \end{pmatrix}$$

Due to the Dirac quantization condition, the electric and magnetic charges of D-branes must be integers so that the true symmetry in Type IIB string theory is indeed $\text{SL}(2, \mathbb{Z})$ [HT95].

In particular, if we choose

$$\Lambda = S = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix},$$

the fields transform as $\tau \leftrightarrow -1/\tau$, and the extended objects are exchanged as

$$F1 \leftrightarrow D1, \quad \text{NS5} \leftrightarrow \text{D5}.$$

Assuming that $\langle C_0 \rangle = 0$, this results in the S-duality of the string coupling constant, $g_s \leftrightarrow 1/g_s$. Other elements of $\text{SL}(2, \mathbb{Z})$ lead to infinitely many bound states of F1 and D1: (p, q) -string, and those of NS5 and D5: (p, q) 5-branes.

The combination (13.2) of the fields behaves as a complex structure of a torus and the theory enjoys the modular transformations (6.4) and (6.5) of a torus. Hence, a $D = 12$ dimensional theory, called **F-theory**, is proposed in [Vaf96, MV96a, MV96b] by considering a torus fibration over ten-dimensional spacetime of Type IIB theory, which can be interpreted as the geometric origin of the field combination (13.2).

13.2 S-duality between Type I and Heterotic $\text{SO}(32)$

Now, let us explore the S-duality between Type I theory and Heterotic $\text{SO}(32)$ theory [Wit95b, Dab95, Hul95, PW96] from the perspective of their low-energy effective actions.

As seen in (10.5), the massless fields in Type I string theory are obtained from Type IIB by the orientifold projection, which projects out $C_{(0)}$, $B_{(2)}$, and $C_{(4)}$. This orientifold projection reduces the amount of supersymmetry by half, resulting in a total of 16 supercharges in Type I. Consequently, the gravity part of the Type I supergravity action can be obtained by removing the part of $G_{(1)}$, $H_{(3)}$ and $G_{(5)}$ from (12.19). Additionally, SO(32) gauge fields with appropriate dilaton dependence are included:

$$\begin{aligned} S_I &= S_{\text{grav}} + S_{\text{YM}} \\ S_{\text{grav}} &= \frac{1}{2\kappa_{10}^2} \int d^{10}x \sqrt{-G} \left[e^{-2\Phi} (R + 4\partial_\mu \Phi \partial^\mu \Phi) - \frac{1}{2} |\tilde{G}_{(3)}|^2 \right] \\ S_{\text{YM}} &= -\frac{1}{2g_{10}^2} \int d^{10}x \sqrt{-G} e^{-\Phi} \text{Tr} |F_{(2)}|^2 \end{aligned} \quad (13.4)$$

where $F_{(2)}$ is the SO(32) field strength and the trace is in the adjoint representation. Here $G_{(3)}$ is the field strength of the R-R 2-form $C_{(2)}$

$$\tilde{G}_{(3)} = dC_{(2)} - \frac{\kappa_{10}^2}{g_{10}^2} \omega_3 \quad (13.5)$$

with the Chern-Simons 3-form

$$\omega_3 = \text{Tr} \left(A dA + \frac{2}{3} A^3 \right) - \text{Tr} \left(\omega d\omega + \frac{2}{3} \omega^3 \right). \quad (13.6)$$

The gauge coupling constant g_{10} and the gravitational constant κ_{10} are related by $\kappa_{10}^2/g_{10}^2 = \alpha'/4$, which is determined by anomaly cancellation. Under the gauge transformation $\delta A = d\lambda + [A, \lambda]$ and diffeomorphism $\delta\omega = d\Theta + [\omega, \Theta]$, the Chern-Simons term transforms as

$$\delta\omega_3 = d\text{Tr}(\lambda dA) - d\text{Tr}(\Theta dA) \quad (13.7)$$

up to a topological term. Hence, the $C_{(2)}$ field is sensitive to the gauge transformation as

$$\delta C_{(2)} = \frac{\kappa_{10}^2}{g_{10}^2} [d\text{Tr}(\lambda dA) - d\text{Tr}(\Theta dA)]. \quad (13.8)$$

The difference between the massless fields in Type I and Heterotic strings involves the exchange between $C_{(2)}$ and $B_{(2)}$ (see Table 11.2). As mentioned earlier, S -duality in Type IIB exchanges $C_{(2)}$ and $B_{(2)}$. Therefore, it is natural to expect that Type I and Heterotic SO(32) (HO) are related by S -duality. However, in the absence of open strings or R-R fields, the dilaton dependence should be $e^{-2\Phi}$ throughout:

$$\begin{aligned} S_{\text{Het}} &= S_{\text{grav}} + S_{\text{YM}} \\ S_{\text{grav}} &= \frac{1}{2\kappa_{10}^2} \int d^{10}x \sqrt{-G} e^{-2\Phi} \left[R + 4\partial_\mu \Phi \partial^\mu \Phi - \frac{1}{2} |\tilde{H}_{(3)}|^2 \right] \\ S_{\text{YM}} &= -\frac{1}{2g_{10}^2} \int d^{10}x \sqrt{-G} e^{-2\Phi} \text{Tr} |F_{(2)}|^2 \end{aligned} \quad (13.9)$$

where the 3-form \tilde{H}_3 is the field strength of the B -field equipped with Chern-Simons form

$$\tilde{H}_{(3)} = dB_{(2)} - \frac{\kappa_{10}^2}{g_{10}^2} \omega_3.$$

Indeed the low-energy effective actions of Type I (13.4) and Heterotic $\text{SO}(32)$ (13.9) are related by the following exchange

$$\begin{aligned} G_{\mu\nu}^I &\leftrightarrow e^{-\Phi^H} G_{\mu\nu}^H, & \Phi^I &\leftrightarrow -\Phi^H \\ \tilde{G}_{(3)}^I &\leftrightarrow \tilde{H}_{(3)}^H, & A^I &\leftrightarrow A^H. \end{aligned} \quad (13.10)$$

Recalling that the vacuum expectation value of the dilaton is the string coupling $g_s = e^\Phi$, we see that the strong coupling limit of one theory is related to the weak coupling limit of the other theory and vice versa.

In Type I theory, D1-branes and D5-branes are electrically and magnetically charged under $C_{(2)}$, respectively. In Heterotic $\text{SO}(32)$ theory, fundamental strings and NS5-branes are electrically and magnetically charged under $B_{(2)}$, respectively. The S-duality maps them as [PW96]

Type I	\leftrightarrow	Heterotic $\text{SO}(32)$
D1-branes	\leftrightarrow	F-strings
D5-branes	\leftrightarrow	NS5-branes

One can provide another evidence of this duality by looking at the massless spectrum. We have seen that Heterotic $\text{SO}(32)$ has massless fields:

1. 8_v of $\text{SO}(8)$: bosonic right-moving $X^i(z)$
2. 8_c of $\text{SO}(8)$: fermionic right-moving $\psi^i(z)$
3. 32 of $\text{SO}(32)$: left-moving Majorana-Weyl fermion $\bar{\lambda}^a(\bar{z})$

Correspondingly, one can see the massless BPS excitations from D1-strings stretched in the x_1 -direction in Type I theory (Homework):

1. 8_v of $\text{SO}(8)$: normal bosonic excitations of D1-D1 strings
2. 8_c of $\text{SO}(8)$: right-moving fermionic excitations of D1-D1 strings
3. 32 of $\text{SO}(32)$: left-moving fermionic excitations of D1-D9 strings

Further evidence of this duality has been assembled by comparing tensions, $F_{(2)}^4$ interactions, and so on [Pol98, §14.3].

13.3 Heterotic T-duality

Let us consider the T-duality in Heterotic strings on a circle S^1 in [Nar86, NSW87, Gin87]. In the bosonic construction, the bosonic left-moving sector is compactified on an even self-dual Euclidean lattice of 16-dimensions. There are only two such lattices: the weight lattice $\Gamma_{\text{SO}(32)}$ of $\text{SO}(32)$ and the root lattice $\Gamma_{E_8} \oplus \Gamma_{E_8}$ of $E_8 \times E_8$ as we have seen in §11.1.

One may also describe the compactification on a circle S^1 in terms of lattices. As we have seen, the left-moving and right-moving momenta compactified boson takes the value on the lattice $\Gamma^{1,1}$ of the Lorentzian signature. Hence, the circle compactification results in adding (\oplus) the lattice $\Gamma^{1,1}$ to the original lattice.

It is a useful mathematical fact that for Lorentzian lattices, there is only unique even unimodular Lorentzian lattice for each rank. Therefore, the theorem implies

$$\Gamma_{\text{SO}(32)} \oplus \Gamma^{1,1} \cong \Gamma^{1,17} \cong \Gamma_{E_8} \oplus \Gamma_{E_8} \oplus \Gamma^{1,1}.$$

Together with the metric G and the B -field, they parameterize the moduli space

$$\mathcal{M} = \frac{\mathrm{O}(1, 17)}{\mathrm{O}(1) \times \mathrm{O}(17)} \Big/ \mathrm{O}(1, 17; \mathbb{Z}), \quad (13.11)$$

where $\mathrm{O}(1, 17; \mathbb{Z})$ is the T-duality group. It is called **Narain moduli space**. Different points in the moduli space correspond to physically distinct compactifications, e.g. the gauge groups can be different, although always of rank 18. At generic points, it is $\mathrm{U}(1)^{18}$ that corresponds to the fact that Wilson loops generically break the 10d gauge group to $\mathrm{U}(1)^{18}$.

However, there are special subspaces of the moduli space where it is enhanced. This moduli space has exactly two asymptotic boundary points, one associated to the decomposition $\Gamma^{1,17} \cong \Gamma_{E_8} \oplus \Gamma_{E_8} \oplus \Gamma^{1,1}$, and the other to the decomposition $\Gamma^{1,17} \cong \Gamma_{\mathrm{SO}(32)} \oplus \Gamma^{1,1}$. Therefore, these boundary points are associated to HE and HO string, or large and small radii. T-duality will relate these boundary points.

In fact, starting from either Heterotic theory, there is a simple choice of Wilson line which breaks the gauge group to $\mathrm{SO}(16) \times \mathrm{SO}(16) \times \mathrm{U}(1) \times \mathrm{U}(1)$. If we leave this group unbroken, then the only remaining parameter is the radius. An analysis of the massive states shows that if we map $R \rightarrow 1/R$ while exchanging KK momenta and winding modes, then the two Heterotic theories are exchanged [Pol98, §11.6].

More generally, upon a compactification of Heterotic strings on a d -dimensional torus T^d , momenta take the values on an even self-dual Lorentzian lattice $\Gamma^{d,d+16}$. Therefore, the Narain moduli space becomes

$$\mathcal{M} = \frac{\mathrm{O}(d, d+16)}{\mathrm{O}(d) \times \mathrm{O}(d+16)} \Big/ \mathrm{O}(d, d+16; \mathbb{Z}) . \quad (13.12)$$

which is $d(d+16)$ -dimensional.

13.4 M-theory

We would also like to understand the strong coupling region of Type IIA theory. As we obtain Type IIA supergravity by compactification of the $D = 11$ supergravity on S^1 , the radius is proportional to the string coupling constant (12.18) in the unit of the string length. This suggests that in the strong coupling region $g_s \gg 1$ of Type IIA string theory, another spacetime direction emerges (dimensional decompactification). Although the strong coupling behavior may look rather bizarre, it turns out that a number of phenomena in string theory can be explained “more naturally” from the $D = 11$ strongly-coupled Type IIA “string theory”. Witten calls the $D = 11$ “string theory” **M-theory** where M stands for Magic, Mystery, or Membrane, according to him. Thus, the $D = 11$ supergravity is the low-energy effective theory of M-theory.

Since the $D = 11$ supergravity is endowed with the 3-form fields $M_{(3)}$, there must be corresponding objects coupled to the field electrically and magnetically. They are called **M2-brane** and **M5-brane**, which are $(1+2)$ - and $(1+5)$ -dimensional objects, respectively. Let us assume that they have the following tensions and charges

$$T_{M2} = \mu_{M2} = \frac{2\pi}{(2\pi\ell_p)^3}, \quad T_{M5} = \mu_{M5} = \frac{2\pi}{(2\pi\ell_p)^6} .$$

The compactification of M-theory on a circle S^1 leads to the IIA superstring theory. Since it is compactified on S^1 , the momenta of a particle along S^1 are quantized as n/R

($n \in \mathbb{Z}$), and it can be identified with D0-branes. An M2-brane on S^1 becomes a fundamental string while an M2-brane in $D = 10$ reduces to a D2-brane. Similarly, an M5-brane on S^1 becomes a D4-brane whereas an M5-brane in $D = 10$ reduces to an NS5-brane. A D6-brane arises geometrically as a Kaluza-Klein monopole, which we do not deal with in this lecture note. (See [Joh02, §15.2] for instance.) Let us see this by comparing the branes and their tensions in Table 6. One can convince oneself that the tensions of extended objects perfectly agree. Note that the tensions and charges of D-branes in Type IIA theory are related by the string coupling constant

$$g_s T_{Dp}^{\text{eff}} = \mu_p . \quad (13.13)$$

In fact, charges of a Dp and $D(6 - p)$ -brane can be read off from Table 6, and it is easy to check that they are subject to the Dirac quantization condition (8.13). We will come back to this relation in §15.1.

Note that M-theory is *not* even defined in a sense that we do not know how to quantize the M-branes. For instance, the world-volume theory on M5-branes is endowed with $\text{OSp}(2, 6|2)$ superconformal symmetry, called 6d $\mathcal{N} = (2, 0)$ superconformal field theory [Wit95a]. The field contents are as follows

- 2-form tensor field B with self-dual field strength $dB = H = -*H$
- spinors $\psi_{\alpha,a}$ with $\psi_{\alpha,a} = J_{\alpha\beta} J_{ab} \bar{\psi}^{\beta,b}$ ($\alpha, a = 1, 2, 3, 4$)
- 5 scalars ϕ_i ($i = 1, \dots, 5$)

The source for the tensor field is an M2-brane ending on the M5-brane, called self-dual string, $*dH = J$. However, we do not know how to write down Lagrangian of the theory for nonabelian cases because of the self-dual 3-form H :

$$\frac{1}{g^2} \int d^6x H \wedge *H = -\frac{1}{g^2} \int d^6x H \wedge H = 0 . \quad (13.14)$$

It is believed that 6d $\mathcal{N} = (2, 0)$ SCFTs are classified by $A_n, D_n, E_{6,7,8}$ root system.

Nevertheless, the existence of such theories tells us a lot. Especially, even though no effective description of M5-branes in flat space is known, M5-branes wrapped on manifolds give rise to surprising dualities between a d -dim topological theory and $(6 - d)$ -dim supersymmetric gauge theories. A salient example is the AGT relation [AGT10] where M5-branes are wrapped on Riemann surface. Also, the highest dimension of SCFTs is $D = 6$ [Nah78], and string theory is indispensable for the study of 6d SCFTs. The reader is referred to [HR19] for this subject.

Dimension	0	1	2	4	5	6
M on S^1	KK-mom.	$M2/S^1$	M2	$M5/S^1$	M5	KK-mono.
	$\frac{1}{R}$	$\frac{2\pi \cdot 2\pi R}{(2\pi\ell_p)^3}$	$\frac{2\pi}{(2\pi\ell_p)^3}$	$\frac{2\pi \cdot 2\pi R}{(2\pi\ell_p)^6}$	$\frac{2\pi}{(2\pi\ell_p)^6}$	$\frac{2\pi(2\pi R)^2}{(2\pi\ell_p)^9}$
IIA	D0	F1	D2	D4	NS5	D6
	$\frac{2\pi}{g_s(2\pi\ell_s)}$	$\frac{2\pi}{(2\pi\ell_s)^2}$	$\frac{2\pi}{g_s(2\pi\ell_s)^3}$	$\frac{2\pi}{g_s(2\pi\ell_s)^5}$	$\frac{2\pi}{g_s^2(2\pi\ell_s)^6}$	$\frac{2\pi}{g_s(2\pi\ell_s)^7}$

Table 6: Wrapped/unwrapped M-branes and the corresponding extended objects in Type IIA theory with their effective tensions.

13.5 Heterotic $E_8 \times E_8$ string from M-theory

Now we shall consider the strong-coupling behavior of Heterotic $E_8 \times E_8$ theory. Taking T-duality and S-duality, Heterotic $E_8 \times E_8$ theory is dual to Type I theory. As seen in §10.3, the T-dual to Type I theory is the Type I' theory, which is Type IIA theory on a line segment S^1/\mathbb{Z}_2 where $O8^-$ -planes sit at the two ends, and $(16 + 16)$ D8-branes are distributed on S^1/\mathbb{Z}_2 . In the strong coupling regime, the M-theory circle will emerge, and it is described as M-theory on $S^1 \times S^1/\mathbb{Z}_2$.

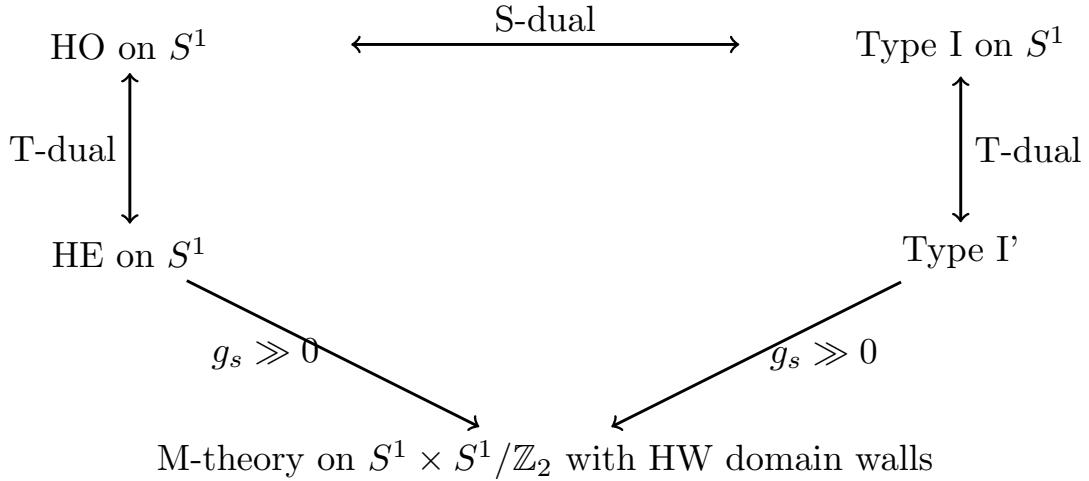


Figure 37: Duality web for Heterotic M-theory

Interestingly enough, the relative position of $O8^-$ -planes and D8-branes in Type I' string theory may be adjusted. This freedom goes away in the M-theory limit; the D8-branes have to be stuck at the $O8^-$ -planes, and they become the domain walls of M-theory, which are called **Hořava-Witten domain wall** or **M9-branes** [HW96b, HW96a].

Its low-energy effective description is given by $D = 11$ supergravity on S^1/\mathbb{Z}_2 which gives rise to gravitational anomaly [AGW84]. In order to cancel such anomaly, non-Abelian gauge fields have to be present at the boundaries in order to employ a Green-Schwarz mechanism [GS84, GS85]. This mechanism that bulk anomaly cancels with boundary anomaly is called **anomaly inflow**. Indeed the low-energy effective theory at the Hořava-Witten domain wall is described by 10d $\mathcal{N} = 1$ SYM with E_8 gauge group that cancels anomaly.

As in Type IIA, the distance between the two boundaries is related to Heterotic coupling $R = g_{\text{het}}^{\frac{3}{2}} \ell_p$. Hence, the line segment S^1/\mathbb{Z}_2 shrinks at the weak coupling regime, leading to Heterotic $E_8 \times E_8$ string theory. The reason to pick $E_8 \times E_8$ is that the anomalies must be canceled on both boundaries, and there is no way to distribute $SO(32)$ between two boundaries (it's a simple group with no factors). Using the previous terminology Heterotic $E_8 \times E_8$ string theory can be viewed as M-theory compactified on S^1/\mathbb{Z}_2 . This setup is called **Hořava-Witten M-theory** or **Heterotic M-theory** [HW96b, HW96a].

13.6 Duality between Heterotic on T^4 and Type IIA on K3

Let us now see one more non-trivial duality. Although we have studied only toroidal compactifications, we have seen a rich web of dualities. In string theory, a theory is

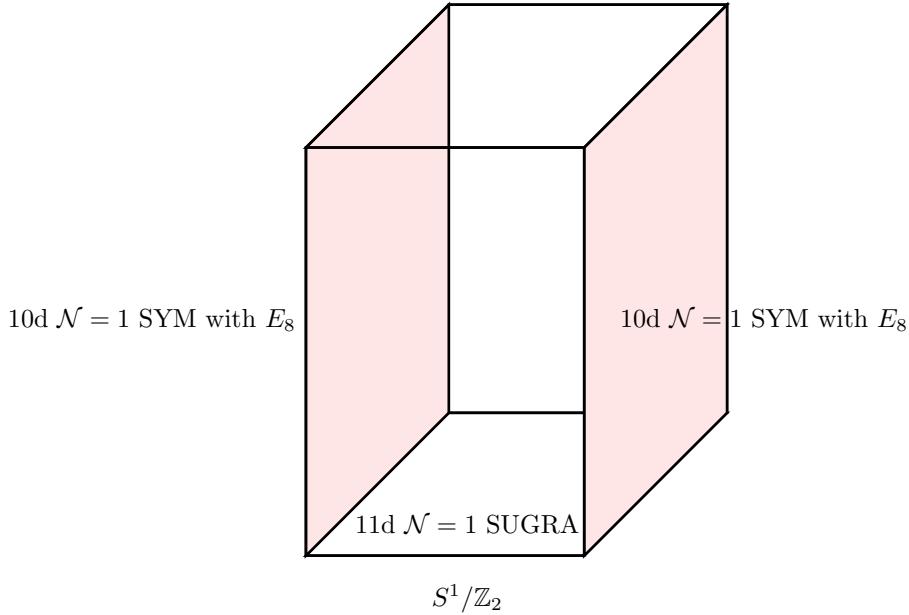


Figure 38: Low-energy effective description of Heterotic M-theory. Hořava-Witten domain walls at the two boundaries give rise to $\mathcal{N} = 1$ SYM with E_8 gauge group and they cancel bulk anomaly.

consistent if we compactify it on a Calabi-Yau manifold. Since there are wide varieties of Calabi-Yau manifolds, string dualities involving them are much richer. It has still been an active research area both in physics and mathematics. Here, we deal with the next simplest Calabi-Yau manifold called **K3 surface**.

K3 surface

A K3 surface is a resolution of T^4/\mathbb{Z}_2 . We write a 4-torus as

$$T^4 = \mathbb{R}^4/\mathbb{Z}^4 = \{\mathbf{x} = (x_1, x_2, x_3, x_4) \in \mathbb{R}^4 | x_i \sim x_i + 1\}$$

and the \mathbb{Z}_2 action is a reflection $x_i \rightarrow -x_i$. Note that this action has $2^4 = 16$ fixed points given by choice of midpoints or the origin in any of the four x_i . Thus, the resulting space T^4/\mathbb{Z}_2 is singular at any of these 16 fixed points. The neighborhood of a singular point is indeed a cone of $\mathbb{R}P^3$. To make it smooth, let us consider the set of vectors of length ≤ 1 in the tangent bundle of TS^2

$$V = \{(v_1, v_2) \in S^2 \times T_{v_1}S^2 | |v_2| \leq 1\}.$$

Then the boundary of V is $\partial V = \mathbb{R}P^3$ so that you can replace the neighborhood of each singular point by V . Since V is a smooth manifold, the resulting space is smooth, and it is a K3 surface. This smoothing procedure is called **resolution** or **blow-up**.

Although the construction of a K3 surface is rather simple, its geometry is surprisingly fertile. First of all, it is a Calabi-Yau manifold, namely a Ricci-flat Kähler manifold. In real four dimensions, there are only two topologically equivalent compact closed

Calabi-Yau manifolds, T^4 and K3. Moreover, it is a hyper-Kähler manifold. (Let us not go into detail about hyper-Kähler manifolds.)

Let us now briefly look at the topological property of K3 surfaces. The resolution of the 16 singular points provides 16 elements of $H^2(K3, \mathbb{Z})$ in addition to 6 = ${}_4C_2$ tori in T^4 . Therefore, we have $H^2(K3, \mathbb{Z}) \cong \mathbb{Z}^{22}$. Moreover, the Hodge diamond as a complex manifold turns out to be

$$\begin{matrix} & & h^{0,0} & & & & 1 \\ & h^{1,0} & & h^{0,1} & & & 0 & 0 \\ h^{2,0} & & h^{1,1} & & h^{0,2} & = & 1 & 20 & 1 \\ & h^{2,1} & & h^{1,2} & & & 0 & 0 \\ & & h^{2,2} & & & & & 1 \end{matrix}.$$

Since it is a real 4-dimensional manifold, one can consider the intersection matrix of rank 22

$$Q(\alpha_i, \alpha_j) = \alpha_i \cap \alpha_j \quad \alpha_i \in H_2(K3; \mathbb{Z})$$

In fact, in a certain nice basis, the intersection matrix can be written as follows

$$Q(\alpha_i, \alpha_j) \sim 2(-E_8) \oplus 3 \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

where $-E_8$ denotes the 8×8 matrix given by minus the Cartan matrix of the Lie algebra E_8 . Hence, we may decompose

$$H^2(K3, \mathbb{R}) = H^+ \oplus H^-,$$

where H^\pm represents the cohomology of the space of (anti-)self-dual 2-forms. We then see that

$$\dim H^+ = 3, \quad \dim H^- = 19.$$

The moduli space of non-trivial metric deformations on a K3 is 58-dimensional and given by the coset space [Asp96]

$$\mathcal{M}_{K3} = \mathbb{R}^+ \times \frac{\mathrm{O}(3, 19)}{\mathrm{O}(3) \times \mathrm{O}(19)} / \mathrm{O}(3, 19, \mathbb{Z}),$$

where the second factor is the Teichmüller space for Ricci-flat metrics of volume one on a K3 surface and the first factor is associated with the size of the K3.

This is not the end of the story if we consider string propagation on K3. For each element of $H_2(K3, \mathbb{Z})$, we can turn on the B -field. Because of $H_2(K3, \mathbb{Z}) = \mathbb{Z}^{22}$, we have 22 additional real parameters and that makes the total dimension of moduli space $58 + 22 = 80$. It turns out that this moduli space is isomorphic to

$$\mathcal{M}_{K3}^{\text{stringy}} = \frac{\mathrm{O}(4, 20)}{\mathrm{O}(4) \times \mathrm{O}(20)} / \mathrm{O}(4, 20, \mathbb{Z}). \quad (13.15)$$

Substituting $D = 4$ into (13.12), one can see that this is exactly the same as the Narain moduli space for Heterotic string on T^4 !

Heterotic on T^4 /Type IIA on K3

The duality between Heterotic on T^4 and Type IIA on K3 can be seen by comparing the effective actions in $D = 6$.

On the Heterotic side, when generic Wilson lines are considered, the gauge symmetry $E_8 \times E_8$ or $\text{SO}(32)$ is broken to $\text{U}(1)^{16}$. By including the Kaluza-Klein (KK) gauge bosons resulting from the compactification on T^4 , the gauge group is enhanced to $\text{U}(1)^{24}$. The effective action for the six-dimensional supergravity in Heterotic string theory is given by

$$S_{\text{Het}} = \frac{1}{2\kappa_6^2} \int d^6x \sqrt{-G} e^{-2\Phi} \left[R + 4\partial_\mu \Phi \partial^\mu \Phi - \frac{1}{2} |\tilde{H}_{(3)}|^2 - \frac{\kappa_6^2}{2g_6^2} \sum_{I=1}^{24} |F_{(2)}^I|^2 \right].$$

where κ_6 represents the six-dimensional gravitational coupling, G is the determinant of the six-dimensional metric, Φ denotes the dilaton field, $\tilde{H}_{(3)}$ is the field strength associated with the antisymmetric tensor field $B_{(2)}$, and $F_{(2)}^I$ corresponds to the field strength of the $\text{U}(1)^{24}$ gauge fields.

On the Type IIA side, the compactification of the superstring theory on K3 leads to the breaking of half of the supersymmetries, resulting in 16 preserved supercharges, which is the same as in the Heterotic theory. The compactification also gives rise to $\text{U}(1)^{24}$ gauge fields, all originating from the R-R sector. One comes from the one-form field $C_{(1)}$ with indices in the non-compact six-dimensional spacetime, while the other gauge fields arise from the three-form field $C_{(3)}$, which is Hodge-dual to a massless vector in six dimensions. Since $H_2(K3, \mathbb{Z}) = \mathbb{Z}^{22}$, the three-form field $C_{(3)}$ on the two-cycles of K3 provides 22 vectors, resulting in a total of $1 + 1 + 22 = 24$ $\text{U}(1)$ gauge fields. The effective action for Type IIA superstring theory compactified on K3 is given by

$$S_{\text{IIA}} = \frac{1}{2\kappa_6^2} \int d^6x \sqrt{-G} \left[e^{-2\Phi} \left(R + 4\partial_\mu \Phi \partial^\mu \Phi - \frac{1}{2} |\tilde{H}_{(3)}|^2 \right) - \frac{\kappa_6^2}{2g_6^2} \sum_{I=1}^{24} |F_{(2)}^I|^2 \right].$$

It is straightforward to see that these two actions are equivalent through a straightforward field redefinition,

$$\begin{aligned} \Phi^H &= -\Phi^{\text{IIA}}, & G^H &= e^{-2\Phi^{\text{IIA}}} G^{\text{IIA}} \\ A^H &= A^{\text{IIA}}, & \tilde{H}^H &= e^{-2\Phi^{\text{IIA}}} * \tilde{H}^{\text{IIA}}. \end{aligned}$$

More dualities

Indeed, the examples presented above provide a glimpse into the rich landscape of string dualities. However, these are just a few representative instances of the many dualities that have been discovered in string theory. By considering different compactifications of M-theory on tori with orbifold actions, one can uncover numerous other dualities. Similar arguments and techniques as discussed earlier have been employed to establish these dualities. Here are some conjectured dualities in this context, supported by references [DM96a, Wit96, Sen96]:

M-theory on		
K3	\leftrightarrow	Heterotic/Type I on T^3
T^5/\mathbb{Z}_2	\leftrightarrow	IIB on K3
T^8/\mathbb{Z}_2	\leftrightarrow	Type I/Heterotic on T^7
T^9/\mathbb{Z}_2	\leftrightarrow	Type IIB on T^8/\mathbb{Z}_2

In each case, the action of the \mathbb{Z}_2 orbifold is to reverse the sign of all the coordinates of T^D . These proposed dualities have been extensively examined and verified by taking the Type IIA limit of M-theory.

14 Green-Schwarz anomaly cancellation

In §10, we observed from the perspective of string amplitudes that unoriented superstrings can be well-defined only for $\text{SO}(32)$ gauge groups. In this section, we consider the celebrated **Green-Schwarz anomaly cancellation** for Type I supergravity from the viewpoint of anomaly polynomials.

Quantum anomalies arise when the partition function is not invariant under gauge transformations in the presence of background gauge fields. More concretely, this is expressed as:

$$Z_{\text{QFT}}[A^g] = e^{i\alpha(A, \epsilon)} Z_{\text{QFT}}[A], \quad (14.1)$$

where A^g and A are background gauge fields of the symmetry group G , related by a gauge transformation ϵ . An anomaly is identified when $\alpha(A, \epsilon)$ cannot be eliminated by local counterterms, indicating a global topological (cohomological) issue.

Anomalies for global symmetries, known as 't Hooft anomalies, are a fundamental aspect of quantum field theory (QFT). According to [tH80, Fre23], an 't Hooft anomaly is a feature, not a bug. It also represents an obstruction to gauging the symmetry G . Therefore, the gauge anomaly must vanish for a well-defined QFT. This section focuses on the cancellation of gauge anomalies in Type I and heterotic string theory.

The anomaly for a QFT on a manifold M_{2n} is given by the integral of a local density. Mathematically, this is expressed as:

$$\alpha(A, \epsilon) = \delta_\epsilon W[A] = 2\pi \int_{M_{2n}} \omega_{2n}, \quad (14.2)$$

where ω_{2n} is the local density contributing to the anomaly. In this section, we introduce the powerful framework of **anomaly polynomials** to derive ω_{2n} . The anomaly polynomial is a characteristic class of a degree 2 higher than the spacetime dimension of the theory under consideration. By performing a simple algebraic operation called the **descent equation** on it, the anomaly ω_{2n} of the theory can be obtained. For details, we refer to [GSW87, §13], [Bil08, AGVM24].

Our objective is to study chiral anomalies in string theory. A necessary preliminary is an understanding of chiral gauge theories.

14.1 Anomalies and the Wess-Zumino condition

Let us focus on an even-dimensional space-time $D = 2n$. The Clifford algebra

$$\{\Gamma_\mu, \Gamma_\nu\} = 2\eta_{\mu\nu}, \quad (14.3)$$

has a spinor representation of dimension 2^n , which corresponds to the Dirac spinor. Defining the chirality operator

$$\frac{1 \pm \Gamma_{2n+1}}{2}, \quad \Gamma_{2n+1} = i^{n-1} \Gamma_0 \Gamma_1 \cdots \Gamma_{2n-1}, \quad (14.4)$$

we can decompose it into a sum of irreducible representations:

$$\mathbf{2}^n = \mathbf{2}_+^{n-1} \oplus \mathbf{2}_-^{n-1}. \quad (14.5)$$

Thus, if ψ is a $\mathbf{2}^n$ component Dirac spinor, it can be decomposed into a sum of chiral (Weyl) spinors ψ_{\pm} where

$$\psi_{\pm} = \frac{1 \pm \Gamma_{2n+1}}{2} \psi. \quad (14.6)$$

When chiral fermions propagate in the background of gauge and gravitational fields, it is essential to consider the potential for gauge and gravitational anomalies.

Now we consider a chiral theory in which the gauge field A couples only to the Weyl fermion ψ_+ where the effective action $W[A]$ is given by

$$e^{-W[A]} = \int D\psi D\bar{\psi} \exp \left(- \int dx \bar{\psi}_+ iD\psi_+ \right).$$

The covariant derivative $D_{\mu} = \partial_{\mu} + A_{\mu}$ depends on the representation of the chiral fermion ψ_+ . In this section, unless it is specified, we assume the adjoint representation. Let us consider an infinitesimal gauge transformation

$$\delta A_{\mu}^a = D_{\mu} \theta^a \quad (14.7)$$

where T_a is the basis of the gauge Lie algebra \mathfrak{g} , and $\theta = \theta^a T_a$ be an infinitesimal gauge transformation parameter. Under this transformation, we have

$$\begin{aligned} \delta W &= \int d^{2n}x \delta A_{\mu}^a(x) \frac{\delta W}{\delta A_{\mu}^a(x)} \\ &= \int d^{2n}x D_{\mu} \theta^a \frac{\delta W}{\delta A_{\mu}^a(x)} \\ &= - \int d^{2n}x \theta^a \langle D_{\mu} j_a^{\mu} \rangle \end{aligned} \quad (14.8)$$

where

$$j^{\mu}_a = i\bar{\psi}_+ \gamma^{\mu} T_a \psi_+ \quad (14.9)$$

If $\langle D_{\mu} j_a^{\mu} \rangle \neq 0$, then the effective action $W[A]$ is not gauge invariant, and the theory is anomalous. Our aim is to find an expression for and examine the conditions under which the quantum theory maintains the gauge symmetries of the classical theory defined by the action S .

On the other hand, the effective action changes as

$$\begin{aligned} \delta W &= \int d^{2n}x \delta A_{\mu}^a(x) \frac{\delta W}{\delta A_{\mu}^a(x)} \\ &= \int d^{2n}x (\partial_{\mu} \theta^a + f^{abc} A_{\mu}^b \theta^c) \frac{\delta W}{\delta A_{\mu}^a(x)} \\ &= \int d^{2n}x \theta^a \left(-\partial_{\mu} \frac{\delta W}{\delta A_{\mu}^a(x)} - f^{abc} A_{\mu}^b(x) \frac{\delta W}{\delta A_{\mu}^c(x)} \right) \end{aligned} \quad (14.10)$$

Comparing this with (14.8), we have the equality

$$-\delta W = \int d^{2n}x \theta^a(x) \frac{\delta}{\delta \theta^a}(x) W = \int d^{2n}x \theta^a(x) \langle D_{\mu} j_a^{\mu}(x) \rangle \quad (14.11)$$

where we define

$$\frac{\delta}{\delta \theta^a} = \partial_\mu \frac{\delta}{\delta A_\mu^a(x)} + f_{ab}{}^c A_\mu^b(x) \frac{\delta}{\delta A_\mu^c(x)}$$

This gives

$$\langle D_\mu J_a^\mu(x) \rangle = \frac{\delta}{\delta \theta^a}(x) W := G_a(x). \quad (14.12)$$

It is easy to see that

$$\left[\frac{\delta}{\delta \theta^a}(x), \frac{\delta}{\delta \theta^b}(x') \right] = f_{abc} \frac{\delta}{\delta \theta^c}(x) \delta_{2n}(x - x'). \quad (14.13)$$

It follows that

$$\frac{\delta}{\delta \theta^a}(x) G_b(x') - \frac{\delta}{\delta \theta^b}(x') G_a(x) = f_{abc} G_c(x) \delta_{2n}(x - x') \quad (14.14)$$

This is known as the **Wess-Zumino consistency condition**, and consistent anomalies obey these equations.

Rewriting Using BRST Transformation

The Wess-Zumino consistency condition can be more concisely written by using the Grassmann odd ghost field $c = c^a T^a$. This will lead to the BRST transformation for a small gauge transformation.

The BRST transformation δ for a gauge field A is defined using the ghost field $c = c^a T^a$ as follows:

$$\begin{aligned} \delta A^a &= c_x^b \frac{\delta}{\delta \theta_x^b} A^a \\ \delta c^a &= -\frac{1}{2} f^{abc} c^b c^c \end{aligned} \quad (14.15)$$

Let us check that the BRST transformation is nilpotent $\delta^2 A = 0$. Applying the BRST transformation twice, we obtain

$$\delta^2 A^a = \delta c_x^b \frac{\delta}{\delta \theta_x^b} A^a - c_x^b \frac{\delta}{\delta \theta_x^b} \delta A^a \quad (14.16)$$

For the first term, we have

$$\delta c_x^b \frac{\delta}{\delta \theta_x^b} A^a = -\frac{1}{2} f^{bcd} c_x^c c_x^d \frac{\delta}{\delta \theta_x^b} A^a = -\frac{1}{2} f^{dbc} c_x^b c_x^c \frac{\delta}{\delta \theta_x^d} A^a \quad (14.17)$$

For the second term, we have

$$c_x^b \frac{\delta}{\delta \theta_x^b} \delta A^a = c_x^b \frac{\delta}{\delta \theta_x^b} c_y^c \frac{\delta}{\delta \theta_y^c} A^a = \frac{1}{2} c_x^b c_x^c \left[\frac{\delta}{\delta \theta_x^b}, \frac{\delta}{\delta \theta_x^c} \right] A^a. \quad (14.18)$$

Therefore, due to the Wess-Zumino integrability condition, it becomes zero

$$\delta^2 A^a = \frac{1}{2} c_x^b c_x^c \left(-f^{dbc} \frac{\delta}{\delta \theta_x^d} + \left[\frac{\delta}{\delta \theta_x^b}, \frac{\delta}{\delta \theta_x^c} \right] \right) A^a = 0. \quad (14.19)$$

The above calculations can be similarly applied to any functional $F[A]$. By formally replacing A with $F[A]$ in the above calculations:

$$\delta^2 F[A] = \frac{1}{2} c_x^b c_x^c \left(-f^{dbc} \frac{\delta}{\delta \theta_x^d} + \left[\frac{\delta}{\delta x^b}, \frac{\delta}{\delta x^c} \right] \right) F[A] = 0. \quad (14.20)$$

Thus, the Wess-Zumino consistency condition (14.13) is equivalent to the nilpotency $\delta^2 = 0$. Introducing the BRST transformation, the Wess-Zumino consistency condition becomes surprisingly simple.

Consequently, the anomaly is obtained by the BRST transformation of the effective action, and this will lead to a cohomological problem for the anomaly as follows:

$$\delta W[A] = 2\pi \int_{M_{2n}} \omega_{2n}(A, c). \quad (14.21)$$

Note that due to the nilpotency, we have

$$\delta^2 W[A] = 2\pi \int_{M_{2n}} \delta \omega_{2n}(A, c) = 0. \quad (14.22)$$

14.2 Anomaly polynomials and decent equations

If the spacetime M_{2n} is closed (i.e., without boundary) and oriented, it can be bounded by a $(2n+1)$ -dimensional oriented manifold M_{2n+1} . By extending the fields (including c) defined on M_{2n} to M_{2n+1} , and the Stokes theorem tells us

$$\delta_c W = 2\pi \int_{M_{2n+1}} d\omega_{2n}(c).$$

The Wess-Zumino consistency condition now states

$$[\delta_{c_1}, \delta_{c_2}] W = 2\pi \int_{M_{2n+1}} [\delta_{c_1} d\omega_{2n}(c_2) - \delta_{c_2} d\omega_{2n}(c_1)] = 2\pi \int_{M_{2n+1}} d\omega_{2n}([c_1, c_2]).$$

Since the above equation holds for any arbitrary extension of the fields into the interior, we have

$$\delta_{c_1} d\omega_{2n}(c_2) - \delta_{c_2} d\omega_{2n}(c_1) = d\omega_{2n}([c_1, c_2]).$$

This implies the existence of a $2n+1$ -form $\omega_{2n+1}^{\text{CS}}$ (independent of c) defined over M_{2n+1} that satisfies

$$\delta_c \omega_{2n+1}^{\text{CS}} = d\omega_{2n}(c).$$

This $\omega_{2n+1}^{\text{CS}}$ is commonly referred to as a Chern-Simons form. Taking an exterior derivative, we define a $(2n+2)$ -form

$$I_{2n+2} = d\omega_{2n+1}^{\text{CS}}, \quad (14.23)$$

which is closed dI_{2n+2} . Moreover, the form I_{2n+2} is gauge invariant

$$\delta I_{2n+2} = d\delta \omega_{2n+1}^{\text{CS}} = d^2 \omega_{2n}(c) = 0$$

since $2n$ and δ_B commute.

The natural candidate for I_{2n+2} is a characteristic class, called **anomaly polynomial**, constructed from the curvatures $F = dA + A^2$ for the gauge field A . More precisely, the anomaly polynomial is a specific polynomial expression (see (14.29)) built from the

curvatures whose specific form depends on the space-time dimension, the global symmetries, and the field content of the theory.

In summary, starting from the anomaly polynomial I_{2n+2} , the Wess-Zumino consistency condition ensures that we can obtain ω_{2n} by the **descent equation**

$$\begin{aligned} I_{2n+2} &= d\omega_{2n+1}^{\text{CS}} \\ \delta\omega_{2n+1}^{\text{CS}} &= d\omega_{2n} \end{aligned} \quad (14.24)$$

The anomaly under the variation of the gauge field is obtained by integrating ω_{2n} over the space-time M_{2n} as in (14.21). The descent equation can be illustrated by the following diagram:

$$\begin{array}{ccccc} I_{2n+2} & \xleftarrow{d} & \omega_{2n+1}^{\text{CS}} & & \\ \downarrow \delta & & \downarrow \delta & & \\ 0 & \xleftarrow{d} & d\omega_{2n} & \xleftarrow{d} & \omega_{2n} \\ & & \downarrow \delta & & \downarrow \delta \\ & & 0 & & \end{array} \quad (14.25)$$

As a toy example, we consider the Chern character

$$I_{2n+2} = \text{ch}_{n+1}(F) = \frac{i^{n+1}}{(2\pi)^{n+1}(n+1)!} \text{Tr } F^{n+1} \quad (14.26)$$

as an anomaly polynomial. It is easy to find the explicit forms for $n = 1$:

$$\begin{aligned} \text{ch}_2(F) &= \frac{1}{8\pi^2} \text{Tr } F^2 \xleftarrow{d} \omega_3^{\text{CS}} = \frac{1}{8\pi^2} \text{Tr} (AdA + \frac{2}{3}A^3) \\ &\qquad\qquad\qquad \downarrow \delta \\ &\qquad\qquad\qquad \omega_2(c) = \frac{1}{8\pi^2} \text{Tr}(c \cdot dA) \end{aligned} \quad (14.27)$$

In this case, ω_3^{CS} is literally the Chern-Simons 3-form so that the variation under the gauge transformation is given by (13.7).

More generally, we have the following expressions

$$\begin{aligned} \text{Tr } F^n &= \int_0^1 dt \frac{d}{dt} [\text{Tr } F_t^n] \\ &= \int_0^1 dt n \cdot \text{Tr} [(D_t A) F_t^{n-1}] \\ &= \int_0^1 dt n \cdot \text{Tr} [D_t (AF_t^{n-1})] \\ &= \int_0^1 dt n \cdot \text{Tr} [d(AF_t^{n-1}) + tA(AF_t^{n-1}) + (AF_t^{n-1}) tA] \\ &= \int_0^1 dt n \cdot d[\text{Tr } AF_t^{n-1}] \end{aligned}$$

Here, let $A_t = tA$ and the field strength be $F_t = tF + (t^2 - t)A^2$. Also, the covariant derivative can be written as $D_t C = dC + A_t C + CA_t$, and the Bianchi identity is $D_t F_t = 0$.

$$\frac{d}{dt} F_t = dA + 2tA^2 = D_t A$$

By noting this, the above expressions can be confirmed. More generally,

$$\omega_{2n+1}^{\text{CS}} = \frac{i^{n+2}}{(2\pi)^{n+2}(n+1)!} \int_0^1 dt \text{Tr} [AF_t^{n+1}] \quad (14.28)$$

is given.

To transition from the first step to the second step of the descent equation (14.24), one should apply the BRST transformation δ to the Chern-Simons form $\omega_{2n+1}^{\text{CS}}$. In the case of a $U(N)$ gauge theory, by applying the BRST transformation to (14.28), we obtain

$$\omega_{2n} = -\frac{i^{n+2}}{(2\pi)^{n+2}(n+1)!} \int_0^1 dt (1-t) \text{Tr} \sum_{r=0}^{n-1} [c \cdot d (F_t^r A F_t^{n-r-1})].$$

So far, we treat the case that the chiral fermion is coupled to the Yang-Mills gauge field, but if the manifold M_{2n} has a curvature, the covariant derivative involves the spin connection

$$D_\mu \psi_+ = \left(\partial_\mu + \frac{1}{4} \omega_\mu^{ab} \Gamma_{ab} + A_\mu \right) \psi_+$$

Then, the anomaly polynomial involves characteristic classes with the spacetime curvature so that it generally can be written as

$$\begin{aligned} I_{2n+2}(R, F) &= P \left(\text{Tr} R^2, \dots, \text{Tr} R^{n+1}; \text{Tr} F, \dots, \text{Tr} F^{n+1} \right) \\ &= \sum_{k=0}^{\lfloor \frac{n+1}{2} \rfloor} \beta_k \text{Tr} R^{2k} \text{Tr} F^{n+1-2k}, \end{aligned} \quad (14.29)$$

where β_k are some constants. Note that the spacetime holonomy of M_{2n} is $\text{SO}(2n)$ in general, so the curvature form R takes

$$\frac{R}{2\pi} = \begin{pmatrix} 0 & x_1 & 0 & 0 & \cdots \\ -x_1 & 0 & 0 & 0 & \cdots \\ 0 & 0 & 0 & x_2 & \cdots \\ 0 & 0 & -x_2 & 0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix} \quad (14.30)$$

where x_j are Chern roots. These generators are $2n \times 2n$ antisymmetric matrices, which implies that $\text{Tr} R^{2k+1} = 0$.

14.2.1 Anomaly polynomials

The evaluation of anomaly polynomials involves the use of the index theorem. Instead of elaborating on the details here, we refer the reader to [Har05, Bil08] for the derivation. To facilitate the expression of anomaly polynomials, we briefly introduce the definition of characteristic classes.

Given the Chern roots (14.30), the Pontryagin classes are defined by

$$p(R) = \prod_{i=1}^n (1 + x_i^2) \equiv 1 + p_1(R) + p_2(R) + \dots + p_n(R), \quad (14.31)$$

where the ℓ -th Pontryagin class (4ℓ -form) is defined as the homogeneous polynomial

$$p_\ell(R) = \sum_{i_1 < \dots < i_\ell}^n x_{i_1}^2 \dots x_{i_\ell}^2. \quad (14.32)$$

For the first few cases, it is straightforward to express $p_\ell(R)$ in terms of the curvature two-form

$$\begin{aligned} p_1 &= -\frac{1}{8\pi^2} \text{Tr } R^2, \\ p_2 &= \frac{1}{128\pi^4} \left[(\text{Tr } R^2)^2 - 2 \text{Tr } R^4 \right], \\ p_3 &= -\frac{1}{3072\pi^6} \left[(\text{Tr } R^2)^3 - 6(\text{Tr } R^2)(\text{Tr } R^4) + 8 \text{Tr } R^6 \right], \end{aligned} \quad (14.33)$$

Moreover, let us introduce another important characteristic class, the \widehat{A} -roof genus:

$$\begin{aligned} \widehat{A}(R) &\equiv \prod_{a=1}^n \frac{x_a/2}{\sinh(x_a/2)} \\ &= 1 - \frac{1}{24} p_1 + \frac{1}{5760} (7p_1^2 - 4p_2) - \frac{1}{967680} (31p_1^3 - 44p_1p_2 + 16p_3) + \dots \end{aligned} \quad (14.34)$$

where we have expressed it in terms of the Pontryagin classes in the second line.

Now, let us provide the contribution of each chiral fermion to the anomaly polynomial using these characteristic classes. For a complex Weyl gravitino of a given chirality, the anomaly polynomial is given by:

$$\begin{aligned} I^{\frac{3}{2}}(R) &= \widehat{A}(R) \left(\text{Tr } e^{R/2\pi} - 1 \right) \\ &= 2n - 1 + \frac{25 - 2n}{24} p_1 + \frac{1}{5760} [(233 + 14n)p_1^2 - 4(239 + 2n)p_2] \\ &\quad + \frac{1}{967680} [(535 - 62n)p_1^3 - 4(559 - 22n)p_1p_2 + 16(505 - 2n)p_3] + \dots \end{aligned} \quad (14.35)$$

For a complex Weyl spinor of a given chirality coupled to a gauge field, the anomaly polynomial is given by:

$$I^{\frac{1}{2}}(R, F) = \widehat{A}(R) \text{ch}(F) \quad (14.36)$$

In these expressions, R represents the curvature 2-form in the fundamental representation of $\text{SO}(2n)$. In $D = 2n$ dimensions, we extract the rank $2n + 2$ terms from the above expressions. The trace $\text{Tr}_R e^{\frac{iF}{2\pi}}$ should be computed in the representation to which the fermions belong.

In $D = 4k + 2$ dimensions, in addition to chiral fermions, self-dual (or anti-self-dual) antisymmetric tensor fields also contribute to pure gravitational anomalies. These fields are of the form

$$G_{(2k+1)} = dC_{(2k)} = \pm * G_{(2k+1)}, \quad (14.37)$$

where the Hodge star operator picks up a minus sign under parity. The (anti)self-dual condition is not parity invariant, and these fields must be considered when analyzing anomalies. The contribution of a self-dual tensor field to the anomaly is given in terms

of the Hirzebruch polynomial $L(R)$ as:

$$\begin{aligned} I^{\text{sd}}(R) &= -\frac{1}{8}L(R) \\ &= -\frac{1}{8} \prod_{a=1}^n \frac{x_a}{\tanh x_a} \\ &= -\frac{1}{8} - \frac{1}{24}p_1 + \frac{1}{360}(p_1^2 - 7p_2) - \frac{1}{7560}(2p_1^3 - 13p_1p_2 + 62p_3) + \dots \end{aligned} \tag{14.38}$$

14.2.2 Anomaly cancellation of IIB supergravity

As an illustrative example of anomaly cancellation, let us examine $D = 10$ $\mathcal{N} = 2$ supergravity. As we recall from §12, there are two $\mathcal{N} = 2$ supergravity models in $D = 10$, corresponding to the low-energy limits of type IIA and type IIB superstring theories. The type IIA theory is non-chiral in terms of the spacetime fermions and thus free from anomalies. (See Table 2.) In contrast, the type IIB theory is chiral in terms of the spacetime fermions, and this chiral theory lacks gauge interactions but can exhibit gravitational anomaly. The fields that contribute to gravitational anomaly include:

- two left-handed spin- $\frac{3}{2}$ gravitini ψ_m^-
- two right-handed spin- $\frac{1}{2}$ dilatinos λ^+
- a self-dual four-form field $C_{(4)}^+$

All chiral fermions also satisfy the Majorana condition. In ten dimensions, the relevant anomaly polynomial is a 12-form. For the gravitino, we have:

$$I_{12}^{\frac{3}{2}}(R) = \frac{1}{107520}(25p_1^3 - 180p_1p_2 + 880p_3), \tag{14.39}$$

For the spin- $\frac{1}{2}$ fields:

$$I_{12}^{\frac{1}{2}}(R) = -\frac{1}{967680}(31p_1^3 - 44p_1p_2 + 16p_3), \tag{14.40}$$

For the self-dual four-form field:

$$I_{12}^{\text{sd}}(R) = -\frac{1}{7560}(2p_1^3 - 13p_1p_2 + 62p_3). \tag{14.41}$$

As a result, the total anomaly polynomial I_{12} is given by:

$$I_{12} = 2 \times \frac{1}{2}I_{12}^{\frac{3}{2}}(R) - 2 \times \frac{1}{2}I_{12}^{\frac{1}{2}}(R) + I_{12}^{\text{sd}}(R) \tag{14.42}$$

The factors of 2 account for the two copies of gravitino and dilatino, while the $\frac{1}{2}$ terms reflect the Majorana-Weyl condition, halving the degrees of freedom compared to a complex Weyl fermion. The signs correspond to their chiralities. Using the above expressions, we can verify the vanishing of the anomaly

$$I_{12}^{\frac{3}{2}}(R) - I_{12}^{\frac{1}{2}}(R) + I_{12}^{\text{sd}}(R) = 0. \tag{14.43}$$

Thus, all gravitational anomalies cancel in Type IIB supergravity. Since this theory does not include gauge fields, gauge or mixed anomalies need not be considered.

14.3 Green-Schwarz mechanism

Finally, let us consider Type I supergravity. As shown in (10.5), the orientifold projection of the Type IIB theory removes the self-dual field. Consequently, in Type I theory, only the left-handed gravitino and the right-handed dilatino contribute to the gravitational anomaly. Therefore, this theory is not free from gravitational anomalies [AGW84]. Specifically, using (14.40) and (14.39), the total anomaly polynomial is found to be nonzero:

$$\begin{aligned} I_{12} &= \frac{1}{2} I_{12}^{\frac{3}{2}}(R) - \frac{1}{2} I_{12}^{\frac{1}{2}}(R) \\ &= \frac{1}{15120} (2p_1^3 - 13p_1p_2 + 62p_3) \neq 0, \end{aligned} \tag{14.44}$$

where the $\frac{1}{2}$ factors arise from the Majorana condition. Nonetheless, the massless spectrum of Type I superstrings includes, besides an $\mathcal{N} = 1$ supergravity multiplet, an $\mathcal{N} = 1$ super-Yang-Mills multiplet (13.4). This multiplet includes its gaugino, a Majorana-Weyl left-handed fermion, which transforms in the adjoint representation of the gauge group G .

To verify anomaly cancellation in Type I string theory, we must add the gaugino's contribution

$$I_{12}^{\frac{1}{2}}(R, F) = \text{ch}_6 - \frac{1}{24} p_1 \text{ch}_4 + \frac{1}{5760} (7p_1^2 - 4p_2) \text{ch}_2 - \frac{\dim G}{967680} (31p_1^3 - 44p_1p_2 + 16p_3). \tag{14.45}$$

with all Chern characters evaluated in the adjoint representation. By incorporating these contributions, the anomaly polynomial for Type I superstrings amounts to

$$\begin{aligned} 2I_{12} &= \text{ch}_6 - \frac{1}{24} p_1 \text{ch}_4 + \frac{1}{5760} (7p_1^2 - 4p_2) \text{ch}_2 \\ &+ \frac{256 - 31 \dim G}{967680} p_1^3 + \frac{11 \dim G - 416}{241920} p_1 p_2 + \frac{496 - \dim G}{60480} p_3, \end{aligned} \tag{14.46}$$

At first glance, it might still seem impossible to construct an anomaly-free $\mathcal{N} = 1$ supergravity plus Yang-Mills theory in $D = 10$. However, this conclusion is incorrect! The Green-Schwarz mechanism is the way out.

The astute reader immediately notice that the p_3 term drops when either $G = \text{SO}(32)$ or $G = E_8 \times E_8$ since $\dim G = 496$. (See §11.) Moreover, for $\text{SO}(n)$ gauge group, the characteristic classes in the adjoint representation and fundamental (vector) representation are related by

$$\text{Tr}_{\text{adj}} F^6 = (n - 32) \text{Tr}_{\square} F^6 + 15 (\text{Tr}_{\square} F^2) (\text{Tr}_{\square} F^4) \tag{14.47}$$

so that the $\text{Tr}_{\square} F^6$ term drops for $\text{SO}(32)$. Even for E_8 , we have

$$\text{Tr}_{\text{adj}} F^6 = \frac{1}{7200} (\text{Tr}_{\text{adj}} F^2)^3. \tag{14.48}$$

Consequently, for $G = \text{SO}(32)$ or $G = E_8 \times E_8$, the ch_6 term in (14.46) splits into

$$\text{ch}_6 = \frac{1}{720} \left(\text{ch}_2 \text{ch}_4 - \frac{1}{1800} \text{ch}_2^3 \right). \tag{14.49}$$

Therefore, for these gauge groups, the anomaly polynomial (14.46) can be factorized into an exterior product of a 4-form and 8-form

$$I_{12} = X_{(4)} \wedge Y_{(8)}, \tag{14.50}$$

where

$$X_{(4)} = -\frac{1}{30} \text{ch}_2 + p_1, \quad Y_{(8)} = -\frac{1}{48} \left(\text{ch}_4 - \frac{1}{1800} \text{ch}_2^2 - \frac{1}{60} p_1 \text{ch}_2 + \frac{3}{8} p_1^2 - \frac{1}{2} p_2 \right). \quad (14.51)$$

This factorization is the key point of the anomaly cancellation as we will see below.

Since both $X_{(4)}$ and $Y_{(8)}$ are closed form, they can be written locally as

$$X_{(4)} = d\Omega_{(3)}, \quad Y_{(8)} = d\Omega_{(7)}, \quad I_{12} = d\omega_{11}^{\text{CS}} \quad (14.52)$$

so that

$$\begin{aligned} I_{12} &= d\Omega_{(3)} \wedge Y_{(8)} = X_{(4)} \wedge d\Omega_{(7)} \\ &= d \left(c \Omega_{(3)} \wedge Y_{(8)} + (1 - c) X_{(4)} \wedge \Omega_{(7)} \right) \end{aligned} \quad (14.53)$$

where c is any real constant.

$$\omega_{11}^{\text{CS}} = c \Omega_{(3)} \wedge Y_{(8)} + (1 - c) X_{(4)} \wedge \Omega_{(7)}$$

Now, the descent equation for each Chern-Simons term yields

$$\delta\Omega_{(3)} = d\mu_{(2)}, \quad \delta\Omega_{(7)} = d\mu_{(6)}. \quad (14.54)$$

The anomaly can be obtained from the gauge variation of ω_{11}^{CS}

$$\delta\omega_{11}^{\text{CS}} = d\omega_{10} = c d\mu_{(2)} \wedge Y_{(8)} + (1 - c) X_{(4)} \wedge d\mu_{(6)}$$

Consequently, the gauge variation of the effective action is given by

$$\delta W = \int \omega_{10} = \int \left[c \mu_{(2)} \wedge Y_{(8)} + (1 - c) X_{(4)} \wedge \mu_{(6)} \right]. \quad (14.55)$$

To the effective action, we can add a local counterterm defined by

$$W_{\text{ct}} = -(1 - c) \int \Omega_{(3)} \wedge \Omega_{(7)} + \frac{4}{\alpha'} \int C_{(2)} \wedge Y_{(8)} \quad (14.56)$$

where $C_{(2)}$ is the R-R 2-form in the $\mathcal{N} = 1 D = 10$ supergravity. Now, let us examine the response of $\tilde{W} = W + W_{\text{ct}}$ with respect to gauge transformation:

$$\begin{aligned} \delta\tilde{W} &= \delta W - (1 - c) \int (d\mu_{(2)} \wedge \Omega_{(7)} + \Omega_{(3)} \wedge d\mu_{(6)}) + \frac{4}{\alpha'} \int \delta C_{(2)} \wedge Y_{(8)} \\ &= \int \left(\mu_{(2)} \wedge Y_{(8)} + \frac{4}{\alpha'} \delta C_{(2)} \wedge Y_{(8)} \right). \end{aligned} \quad (14.57)$$

If we choose

$$\delta C_{(2)} = -\frac{\alpha'}{4} \mu_{(2)}, \quad (14.58)$$

we obtain $\delta\tilde{W} = 0$. Thus, the theory defined by \tilde{W} is anomaly-free. This method of achieving anomaly cancellation is called the **Green-Schwarz mechanism**. Recalling that $\delta\Omega_{(3)} = d\mu_{(2)}$, we have the gauge-invariant combination

$$\tilde{G}_{(3)} = dC_{(2)} + \Omega_{(3)} = dC_{(2)} - \frac{\alpha'}{4} \left[\frac{1}{30} \text{Tr}_{\text{adj}} \left(A dA + \frac{2}{3} A^3 \right) - \text{Tr} \left(\omega d\omega + \frac{2}{3} \omega^3 \right) \right]. \quad (14.59)$$

In fact, the gauge-invariant combination (13.5) that appears in the action of Type I supergravity can be obtained by using the relation of the Chern-Simons terms in the adjoint and fundamental representation of $\text{SO}(32)$:

$$\text{Tr}_{\text{adj}} \left(AdA + \frac{2}{3} A^3 \right) = 30 \text{Tr}_{\square} \left(AdA + \frac{2}{3} A^3 \right). \quad (14.60)$$

The counterterm (14.56) added is often called the Green-Schwarz counterterm, which is a higher derivative correction to the low-energy effective action.

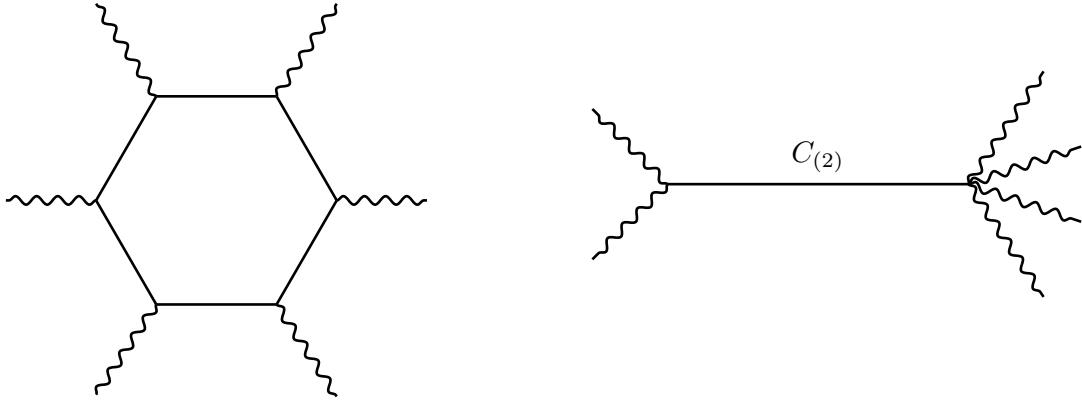


Figure 39: Green-Schwarz anomaly cancellation. The left diagram gives rise to the gravitational anomaly, and the right diagram is from the counterterm to cancel the anomaly.

In fact, the anomaly arises from the hexagon one-loop diagram in $D = 10$. This is canceled by the tree diagram where the $C_{(2)}$ field propagates in the intermediate state (see Figure 39). The kinetic term of $\tilde{G}_{(3)}$ in the supergravity action (13.4) gives the left vertex with two gravitons and/or gauge bosons, while the Green-Schwarz term (14.56) gives rise to the right vertex with four gravitons and/or gauge bosons.

The Green-Schwarz anomaly cancellation predicts string theory with an $E_8 \times E_8$ gauge group. This leads to the formulation of Heterotic string theory [GHMR85b, GHMR85a, GHMR86]. In fact, as we will see in §13.2, the low-energy effective action of the Heterotic string theory is given in (13.9), which is related to Type I supergravity by (13.10).

15 D-brane dynamics

So far, D-branes are treated as static rigid objects that are associated to boundary conditions of an open string. Actually, it has dynamics like a fundamental string, and this section introduces the dynamics of D-branes. After we derive the action for the world-volume theory of Dp -branes in the first half, we discuss the dynamics of D-branes. Branes control non-perturbative dynamics in string theory, which reveals the richness and depth of string theory. Therefore, the study of D-branes is remarkably broad and we cover a tiny part of it. The reader is referred to [Joh06] for more detail.

15.1 D-brane action

When we quantize a world-sheet in §2, we use the string sigma action (2.2) instead of the Nambu-Goto action (2.1) to avoid the complexity of the square root. There, the

dimension of a world-sheet to be two is crucial for quantization, and we cannot simply follow the same procedure for the fundamental string to quantize a Dp -brane, in general. Thus, we will first learn the effective action of a D-brane in this section.

Let us summarize the properties the D-brane action should be endowed with:

1. it contains scalars as the spacetime coordinates of the world-volume of a D-brane
2. it includes gauge fields living on a D-brane, which arises from an open string massless spectrum
3. it involves B -field because open strings can end on a D-brane
4. it couples to the R-R-field $C_{(p+1)}$ via (8.12)
5. it can possess supersymmetry (though we focus only on the bosonic part in this section)

The first point will be addressed by the Nambu-Goto action (2.1) for a D-brane, namely the volume of the D-brane. We write the world-volume coordinates of Dp -branes by $\sigma^a (a = 0, 1, \dots, p)$, and $X^\mu(\sigma^a)$ maps from the world-sheet to the spacetime. Then, the Nambu-Goto action (2.1) for a D-brane can be written as

$$S_{Dp} = -T_{Dp}^{\text{eff}} \int d^{p+1}\sigma \sqrt{-\det \left(G_{\mu\nu} \frac{\partial X^\mu}{\partial \sigma^a} \frac{\partial X^\nu}{\partial \sigma^b} \right)}$$

where T_{Dp}^{eff} is an effective Dp -brane tension which we will discuss shortly. Starting the Nambu-Goto action with the R-R coupling (8.12), we will generalize the action by using the T-duality and gauge invariance. In this way, we will incorporate the second and third point.

Let us consider the simple setup as in Figure 24 where D1 and D2-brane are related by the T-duality. Here we are interested in a part $\mathbb{R}_t \times \mathbb{R} \times S^1$ of the spacetime where the D-brane is located, and we assume that its metric is flat $ds^2 = \eta_{\mu\nu} dx^\mu dx^\nu$ (i.e. $G_{\mu\nu} = \eta_{\mu\nu}$). As explained in §9.2, the position $X_2 (\sim X_2 + 2\pi R)$ of the D1-brane, and the gauge field $A_2 (\sim A_2 + 1/R)$ on the D2-brane are related under the T-duality by

$$X_2 = 2\pi\alpha' A_2 .$$

Now let us consider a situation in which the D1-brane has dynamics, namely it vibrates as $X^2 = X^2(X^1)$ (see Figure 40). Then, the vibration of the D1-brane is translated as

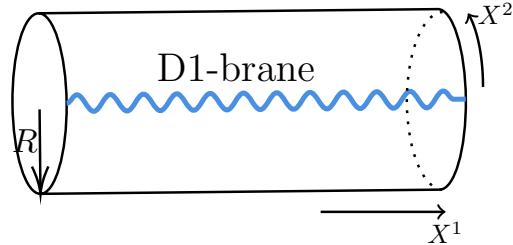


Figure 40: Vibrating D1-brane.

non-trivial field strength $F_{12} = \partial_1 A_2(X^1) \neq 0$ on the D2-brane. Parametrizing the world-volume coordinates as $X^0 = \sigma^0$, $X^1 = \sigma^1$, the Nambu-Goto action of the D1-brane

becomes

$$\begin{aligned} S_{D1} &= -T_{D1}^{\text{eff}} \int d^2\sigma \sqrt{-\det \left(G_{\mu\nu} \frac{\partial X^\mu}{\partial \sigma^a} \frac{\partial X^\nu}{\partial \sigma^b} \right)} \quad \text{with ,} \\ &= -T_{D1}^{\text{eff}} \int d\sigma^0 d\sigma^1 \sqrt{1 + \left(\frac{\partial X^2}{\partial \sigma^1} \right)^2}. \end{aligned}$$

In order for the D-brane action to be invariant under the T-duality, we incorporate the field strength of the gauge field in such a way that

$$\begin{aligned} S_{D2} &= -T_{D2}^{\text{eff}} \int d^3\sigma \sqrt{-\det \left(G_{\mu\nu} \frac{\partial X^\mu}{\partial \sigma^a} \frac{\partial X^\nu}{\partial \sigma^b} + 2\pi\alpha' F_{ab} \right)} \\ &= -T_{D2}^{\text{eff}} \cdot 2\pi\tilde{R} \cdot \int d\sigma^0 d\sigma^1 \sqrt{1 + (2\pi\alpha' F_{12})^2}. \end{aligned}$$

D-brane tension

The dimension analysis tells us that a D-brane tension should be proportional to

$$T_{Dp} \sim \frac{\text{mass}}{p\text{-dim vol}} \Rightarrow T_{Dp} \sim \frac{1}{\ell_s^{p+1}}.$$

From the argument above in order for the two actions to coincide we need $T_{D1}^{\text{eff}} = 2\pi\tilde{R}T_{D2}^{\text{eff}}$. On the other hand, T_{Dp} should be independent of the spacetime geometry including the radius R . Note that D-brane effective theory is supposed to reproduce open string amplitude, whose leading contribution is the disk amplitude $\sim e^{-\langle\Phi\rangle}$. Thus, we reach the following form

$$S_{Dp} = -T_{Dp} \int d^{p+1}\sigma e^{-\Phi(X)} \sqrt{-\det (G_{\mu\nu} \partial_a X^\mu \partial_b X^\nu + 2\pi\alpha' F_{ab})}. \quad (15.1)$$

Then, the ratio of the tensions of D1 and D2-brane is

$$2\pi\tilde{R} = \frac{T_{D1}^{\text{eff}}}{T_{D2}^{\text{eff}}} = \frac{T_{D1}e^{-\Phi}}{T_{D2}e^{-\tilde{\Phi}}} = \frac{T_{D1}}{T_{D2}} \cdot \frac{\tilde{R}}{\ell_s} \Rightarrow T_{D1} = 2\pi\ell_s \cdot T_{D2}.$$

Note that the dilation field transforms under the T-duality as $e^{-\tilde{\Phi}} = e^{-\Phi} \frac{\tilde{R}}{\ell_s}$ (exercise). For a D-brane in superstring theory, the correct normalization is

$$T_{Dp} = \frac{2\pi}{(2\pi\ell_s)^{p+1}}. \quad (15.2)$$

The B -field

In (4.8), we see the action of the bosonic closed string. For an open string, we need to add a boundary term in order for the action to be invariant under gauge transformations:

$$\begin{aligned} S_{\text{open}} &= \frac{1}{4\pi\alpha'} \int_{\Sigma} d^2\sigma \left(\sqrt{h} h^{ab} \partial_a X^\mu \partial_b X^\nu G_{\mu\nu}(X) + i\epsilon^{ab} \partial_a X^\mu \partial_b X^\nu B_{\mu\nu}(X) + \alpha' \sqrt{h} R^{(2)} \Phi(X) \right) \\ &\quad + i \int_{\partial\Sigma} d\sigma^0 \partial_0 X^\mu A_\mu \\ &= \dots + \frac{1}{2\pi\alpha'} \int_{\Sigma} B_{(2)} + \int_{\partial\Sigma} A_{(1)}. \end{aligned} \quad (15.3)$$

Here the gauge transformation (4.9) of the B -field $\delta_B B_{(2)} = d\Lambda_{(1)}$ is compensated by that of the gauge field $A_{(1)}$, which is

$$\delta_B A_{(1)} = -\frac{\Lambda_{(1)}}{2\pi\alpha'}.$$

Therefore, (15.1) can be generalized by incorporating the B -field as

$$S_{Dp} = -T_{Dp} \int d^{p+1}\sigma e^{-\Phi(X)} \sqrt{-\det(G_{ab} + 2\pi\alpha' F_{ab} + B_{ab})}. \quad (15.4)$$

This is called the **Dirac-Born-Infeld (DBI) action**.

Generalization of R-R coupling and the DBI action

In addition, a Dp -brane couples to the R-R field $C_{(p+1)}$ via the action (8.12), which can be written in terms of local coordinates as

$$S_{R-R Dp} = \mu_p \cdot \int_{Dp} C_{(p+1)} = \mu_p \cdot \int d^{p+1}\sigma C_{\mu_1 \dots \mu_{p+1}}(X) \frac{\partial X^{\mu_1}}{\partial \sigma^1} \dots \frac{\partial X^{\mu_{p+1}}}{\partial \sigma^{p+1}}.$$

As above, let us consider the T-duality for the vibrating D1-brane in the R-R coupling. Recalling that R-R fields are transformed under the T-duality by (9.14), the T-duality connects the following two expressions

$$\begin{aligned} S_{R-R D1} &= \mu_1 \cdot \int d\sigma^0 d\sigma^1 \left(C_{01} + C_{02} \frac{\partial X^2}{\partial \sigma^1} \right), \\ S_{R-R D2} &= \mu_2 \cdot \int d\sigma^0 d\sigma^1 d\sigma^2 \left(\tilde{C}_{012} + \tilde{C}_0 \cdot 2\pi\alpha' F_{12} \right), \end{aligned}$$

where $C_{01} \leftrightarrow \tilde{C}_{012}$, $C_{02} \leftrightarrow \tilde{C}_0$, and $X^2 \leftrightarrow 2\pi\alpha' A_2$. This can be understood as follows. The vibrating D1-brane consists of a straight D1-brane along X^1 and local vibration along X^2 . After T-duality along X^2 , the vibration part gives non-trivial flux $F_{12} \neq 0$ or equivalently gives D0-branes. (See left Figure 42.) Therefore, the generalization of the R-R coupling (9.14) is

$$S_{Dp} = \mu_p \cdot \int C_{RR} \wedge \exp(2\pi\alpha' F_{(2)} + B_{(2)}), \quad (15.5)$$

where $C_{RR} = \sum_n C_{(n)}$. Note that the B -field appears with $F_{(2)}$ due to the gauge invariance.

Finally, the fully general form of Dp -brane action is given as follows:

$$\begin{aligned} S_{Dp} &= -T_{Dp} \int d^{p+1}\sigma e^{-\Phi(X)} \sqrt{-\det(G_{ab} + 2\pi\alpha' F_{ab} + B_{ab})} \\ &\quad + \mu_p \cdot \int C_{RR} \wedge \exp(2\pi\alpha' F_{(2)} + B_{(2)}), \end{aligned} \quad (15.6)$$

where $G_{ab} = G_{\mu\nu} \partial_a X^\mu \partial_b X^\nu$ and $B_{ab} = B_{\mu\nu} \partial_a X^\mu \partial_b X^\nu$.

Though we have focused on the bosonic part so far, there is a fermionic part so that they form spacetime supersymmetry. Here, we only write down the leading fluctuation:

$$-i \int d^{p+1}\sigma \text{Tr}(\bar{\psi} \Gamma^a D_a \psi).$$

For the full nonlinear supersymmetric form, one should consult with, for example, [Tse99].

So far, we consider the world-volume effective action of a single D-brane, and non-Abelian generalization for multiple D-branes remains an open problem although there is a proposal [Tse97, Mye99].

D-brane tensions and charges

As seen in (8.13), D-brane charges must satisfy the Dirac quantization condition. Similarly, the D-brane tension (15.2) is also governed by the Dirac quantization condition. Thus, we can express μ_p as $\mu_p = \frac{2\pi}{(2\pi\ell_s)^{p+1}} = T_{Dp}$. The tension between Dp -branes induces an attractive gravitational force (mediated by gravitons and dilatons), whereas the R-R charge induces a repulsive force. Therefore, the equality of the charge and the tension is essential for the stable existence of multiple Dp -branes.

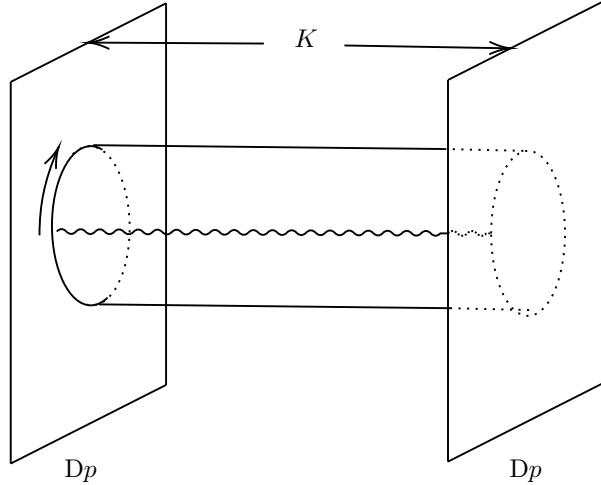


Figure 41:

Let us evaluate the gravitational force between two Dp -branes by following the seminal paper [Pol95]. As illustrated in Figure 27, the closed string amplitude between the Dp -branes can be interpreted as the one-loop amplitude of an open string connecting the Dp -branes. The contribution of a particle with mass m to the free energy at one-loop can be evaluated as follows:

$$\begin{aligned} F &= -V_{p+1} \frac{1}{2} \int \frac{d^{p+1}k}{(2\pi)^{p+1}} \log(k^2 + m^2) \\ &= -V_{p+1} \int \frac{dt}{2t} \int \frac{d^{p+1}k}{(2\pi)^{p+1}} e^{-2\pi\alpha'(k^2 + m^2)t} \\ &= -iV_{p+1} \int \frac{dt}{2t} (8\pi^2 \ell_s^2 t)^{-\frac{p+1}{2}} e^{-2\pi\alpha' m^2 t} \end{aligned} \quad (15.7)$$

Since the mass of an open string mode between the branes at distant K is given by

$$\alpha' m^2 = L_0 + \frac{K^2}{\ell_s^2}, \quad (15.8)$$

the free energy takes the form

$$F = -iV_{p+1} \int \frac{dt}{2t} (8\pi^2 \ell_s^2 t)^{-\frac{p+1}{2}} q^{\frac{K^2}{\ell_s^2}} \text{Tr}(q^{L_0}). \quad (15.9)$$

where

$$\text{Tr}(q^{L_0}) = q^{-\frac{1}{2}} \prod_{m=1}^{\infty} \frac{(1 - q^{m-1/2})^8}{(1 - q^m)^8} = \frac{\eta(it/2)^8}{\eta(it)^{16}}. \quad (15.10)$$

Note that the denominator is the bosonic and the numerator is the fermionic contribution. As we learn in (10.7), the leading contribution from a closed string can be read off in the limit $t \rightarrow 0$ when the cylinder becomes a thin tube. Therefore, we make the modular transformation (7.36) and we take the leading order

$$\begin{aligned} F &= -iV_{p+1} \frac{1}{(2\pi\ell_s)^{p+1}} \int \frac{dt}{2t} (2t)^{\frac{7-p}{2}} e^{-2\pi\frac{K^2}{\ell_s^2}t} \\ &= -iV_{p+1} \frac{1}{(2\pi\ell_s)^{p+1}} \left(\frac{\ell_s^2}{\pi K^2} \right)^{\frac{7-p}{2}} \int \frac{dx}{x} x^{\frac{7-p}{2}} e^{-x} \\ &= -iV_{p+1} \frac{1}{(2\pi\ell_s)^{p+1}} \left(\frac{\ell_s^2}{\pi K^2} \right)^{\frac{7-p}{2}} \Gamma\left(\frac{7-p}{2}\right) \\ &= \text{const} \cdot \ell_s^{6-2p} G_{9-p}(K^2) \end{aligned} \quad (15.11)$$

where $G_d(K)$ is the massless scalar Green's function in d dimensions, the inverse of $-\nabla^2$. In the $D = 10$ supergravity §12, the Newton constant is given by $2\kappa_{10}^2 = (2\pi\ell_s)^8/2\pi$ so that the Newton's potential is given by

$$V_G = -\text{const} \cdot \kappa_{10}^2 T_{Dp}^2 G_{9-p}(r). \quad (15.12)$$

Therefore, we can conclude that $T_{Dp} \propto \ell_s^{-(p+1)}$, which justifies (15.2).

For the B -field, the normalization of the quantization condition is different from (8.13)

$$T_{F1} \cdot T_{NS5} \cdot 2\kappa_{10}^2 g_s^2 \in 2\pi\mathbb{Z}.$$

Since $T_{F1} = \frac{2\pi}{(2\pi\ell_s)^2}$, we have

$$T_{NS5} = \frac{2\pi}{T_{F1} \cdot 2\kappa_{10}^2 g_s^2} = \frac{2\pi}{(2\pi\ell_s)^6 g_s^2}.$$

See also Table 6.

15.2 Branes and strings ending on branes

Now we consider D-brane systems in Type IIA theory. Let us recall that Maxwell equations lead to the charge conservation law:

$$\begin{aligned} d * F_{(2)} &= J_e & \Rightarrow & \quad dJ_e = 0 \\ dF_{(2)} &= J_m & \Rightarrow & \quad dJ_m = 0 \end{aligned}.$$

Because of the charge conservation, a world-line of a charged particle cannot have an endpoint and therefore it must be either a closed path or an infinitely long line. If we apply this logic to branes, we may find the same result for branes. However, the charge conservation in supergravity is quite non-trivial due to the non-linearity of the equation of motions. In this subsection, we will study the physics of D-branes from the equation of motions in (massive) Type IIA supergravity.

Massive IIA supergravity

When Type IIA supergravity action is introduced in (12.13), we do not include the R-R field corresponding to D8-brane. As explained in §8.2, D8-branes are non-dynamical, and its R-R field is a 9-form so that its field strength is 10-form $G_{(10)}$. Consequently, $d *G_{(10)} = 0$ leads to $*G_{(10)} = G_{(0)} \equiv m$, which clearly shows that D8-brane is non-dynamical. However, it actually has a constant but non-trivial contribution to the action, called **massive IIA supergravity**. Since it is a constant, it contributes as a mass term, called **Romans mass** to the action [Rom86].

Let us write the massive IIA supergravity action (we omit the wedge product \wedge and the Hodge star $*$ here):

$$\begin{aligned} S_{A,NS} &= \frac{1}{2\kappa_{10}^2} \int d^{10}x \sqrt{-G} e^{-2\Phi} \left[R + 4\partial_\mu \Phi \partial^\mu \Phi - \frac{1}{2} H_{(3)}^2 \right] , \\ S_{A,R} &= \frac{1}{2\kappa_{10}^2} \int d^{10}x \sqrt{-G} \left[-\frac{1}{2}m^2 - \frac{1}{2}G_{(2)}^2 - \frac{1}{2}G_{(4)}^2 \right] , \\ S_{A,CS} &= \frac{1}{2\kappa_{10}^2} \int \left[-\frac{1}{2}B_{(2)}G_{(4)}G_{(4)} + \frac{1}{2}B_{(2)}^2G_{(2)}G_{(4)} - \frac{1}{6}B_{(2)}^3G_{(2)}^2 \right. \\ &\quad \left. - \frac{m}{6}B_{(2)}^3G_{(4)} + \frac{m}{8}B_{(2)}^4G_{(2)} - \frac{m^2}{40}B_{(2)}^5 \right] , \end{aligned}$$

where the field strengths of the R-R fields receive some modifications (here we omit the tilde notation for simplicity)

$$\begin{aligned} H_{(3)} &= dB_{(2)} , \\ G_{(2)} &= dC_{(1)} + mB_{(2)} , \\ G_{(4)} &= dC_{(3)} + dC_{(1)}B_{(2)} + \frac{1}{2}mB_{(2)}^2 . \end{aligned} \tag{15.13}$$

From the action, we obtain the following equations of motion for the fields $C_{(1)}$ and $C_{(3)}$

$$\begin{aligned} -d *G_{(2)} &= H_{(3)} *G_{(4)} , \\ d *G_{(4)} &= H_{(3)}G_{(4)} . \end{aligned} \tag{15.14}$$

Given the relationship between the field strengths and their corresponding gauge fields (15.13), we derive the following Bianchi identities

$$\begin{aligned} dG_{(2)} &= mH_{(3)} , \\ dG_{(4)} &= H_{(3)}G_{(2)} . \end{aligned} \tag{15.15}$$

If we relabel m as $G_{(0)}$ and define the dual field strengths:

$$G_{(10)} = *G_{(0)} , \quad G_{(8)} = -*G_{(2)} , \quad G_{(6)} = *G_{(4)} ,$$

then the equations of motion (15.14) and the Bianchi identities (15.15) can be unified into

$$dG_{\text{even}} = H_{(3)}G_{\text{even}} ,$$

where

$$\begin{aligned} C_{\text{odd}} &= C_{(1)} + C_{(3)} + C_{(5)} + C_{(7)} + C_{(9)} , \\ G_{\text{even}} &= G_{(0)} + G_{(2)} + G_{(4)} + G_{(6)} + G_{(8)} + G_{(10)} . \end{aligned} \tag{15.16}$$

This equation can be solved by setting

$$G_{\text{even}} = e^{B_{(2)}} (m + dC_{\text{odd}}) .$$

The equation of motion for the B -field is given by

$$d \left(e^{-2\Phi} * H_{(3)} \right) = m * G_{(2)} + * G_{(4)} G_{(2)} - \frac{1}{2} G_{(4)}^2 .$$

Defining the dual field strength $H_{(7)} = e^{-2\Phi} * H_{(3)}$, the equation of motion becomes

$$dH_{(7)} = \frac{1}{2} [(*G_{\text{even}})G_{\text{even}}]_{(8)} . \quad (15.17)$$

The Bianchi identity is straightforward: $dH_{(3)} = 0$. Note that these field strengths are invariant under the gauge transformations of the B -field and the R-R fields

$$\begin{aligned} \delta_B B_{(2)} &= d\lambda_{(1)}, & \delta_B C_{\text{odd}} &= -\lambda_{(1)} (m + dC_{\text{odd}}), \\ \delta_C B_{(2)} &= 0, & \delta_C C_{\text{odd}} &= d\lambda_{\text{even}}, \end{aligned}$$

where we introduced a formal sum of gauge parameters

$$\lambda_{\text{even}} = \lambda_{(0)} + \lambda_{(2)} + \lambda_{(4)} + \lambda_{(6)} + \lambda_{(8)} .$$

Now, let us introduce brane currents $J_{(8)}^{\text{F1}}$, $J_{(4)}^{\text{NS5}}$, and

$$J_{\text{odd}} = J_{(1)}^{\text{D8}} + J_{(3)}^{\text{D6}} + J_{(5)}^{\text{D4}} + J_{(7)}^{\text{D2}} + J_{(9)}^{\text{D0}} . \quad (15.18)$$

We can then include these in the equations of motion

$$\begin{aligned} dH_{(3)} &= J_{(4)}^{\text{NS5}}, \\ dH_{(7)} &= J_{(8)}^{\text{F1}} + \frac{1}{2} [(*G_{\text{even}})G_{\text{even}}]_{(8)}, \\ dG_{\text{even}} &= J_{\text{odd}} + H_{(3)}G_{\text{even}} . \end{aligned}$$

From these, we can derive the following “conservation” laws

$$\begin{aligned} dJ_{(4)}^{\text{NS5}} &= 0, \\ dJ_{(8)}^{\text{F1}} &= -[J_{\text{odd}}(*G_{\text{even}})]_{(9)}, \\ dJ_{\text{odd}} &= -J_{(4)}^{\text{NS5}}G_{\text{even}} - J_{\text{odd}}H_{(3)} . \end{aligned}$$

These equations imply several important facts:

- NS5-branes cannot have boundaries.
- F1 strings can end on any D-branes.
- Dp -branes can end on NS5-branes for $p \leq 6$ (D8 cannot).
- Dp -branes can end on $D(p+2)$ -branes.

A similar analysis can be performed in Type IIB supergravity, or we can apply T-duality appropriately. The results are summarized in Table 15.2.

Brane	Branes end on	Brane	Branes end on
F1	nothing	F1	nothing
NS5-brane	D0, D2, D4, D6	NS5-brane	D1, D3, D5
D0-brane	F1	D1-brane	F1
D2-brane	F1, D0	D3-brane	F1, D1
D4-brane	F1, D2	D5-brane	F1, D3
D6-brane	F1, D4	D7-brane	F1, D5
D8-brane	F1, D6		

Table 7: Which branes can end on a brane in Type IIA (left) and IIB (right).

D0-D2 bound states and Myers effect

The general R-R coupling (15.5) includes all possible R-R fields C_{RR} in the theory. This implies that Dp -branes can include lower-dimensional $D(p - 2n)$ -branes ($n \in \mathbb{Z}_{\geq 0}$), forming a bound state.

Let us consider a concrete example of a D2-brane supported on Σ where the R-R coupling at $B_{(2)} = 0$ is

$$S \sim \frac{1}{2\pi} \int_{\mathbb{R}_t \times \Sigma} (C_{(3)} + F_{(2)} C_{(1)}) ,$$

Note that the $F_{(2)}$ flux needs to be quantized

$$\frac{1}{2\pi} \int_{\Sigma} F_{(2)} = n \in \mathbb{Z} .$$

If $C_{(3)}$ is zero, the D2-brane tries to shrink to a point due to its tension. On the other hand, $C_{(1)}$ part remains finite because the flux $F_{(2)}$ is quantized as

$$S \sim \frac{1}{2\pi} \int_{\mathbb{R}_t \times \Sigma} (F_{(2)} C_{(1)}) \rightarrow n \int_{\mathbb{R}_t} C_{(1)} .$$

This is equivalent to n D0-branes. Namely, when the D2-brane with n flux shrinks to a point, n D0-branes remain. This can be understood as a bound state of D2- and D0-branes (see Figure 42).

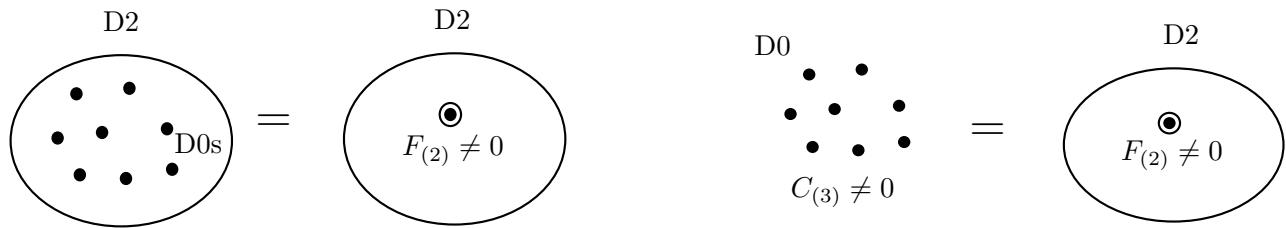


Figure 42: Myers effect: transition from D0-branes to D0-D2 .

Let us consider the opposite process of the previous argument. When there are n D0-branes in the non-trivial $C_{(3)}$ background flux, they can become a D2-brane with the $F_{(2)}$ flux. Due to the non-trivial $C_{(3)}$ flux, being a D2-brane is energetically more preferable than being the D0-branes (see Figure 42). This is similar to the polarization phenomenon in electromagnetism, and in this case, it is called **Myers effect** [Mye99].

15.3 Brane creations/annihilations

In fact, the last two Type IIB D-brane configurations in Table 4 preserve a quarter of supersymmetries, and they give rise to 3d $\mathcal{N} = 4$ theories in X^{012} . Let us first consider the third configuration of Table 4. As in Figure 43, N D3-branes suspended by NS5-branes give rise to $U(N)$ gauge group (or vector multiplet), and F1-strings between D3- and D5-branes give rise to matters (hypermultiplet) in the fundamental representation. Also, bifundamental matter (hypermultiplet) arises from F1-strings between two adjacent D3-branes. The resulting theory is often represented by a quiver diagram. 3d $\mathcal{N} = 4$ supersymmetry is endowed with $SU(2)_C \times SU(2)_H$ R-symmetry where $SU(2)_C$ acts on the X^{345} direction and $SU(2)_H$ acts on the X^{789} direction in the Type IIB setup. Also, 3d $\mathcal{N} = 4$ supersymmetric theories have moduli spaces of vacua, called Coulomb and Higgs branch. Both are hyper-Kähler manifolds. In the brane realization, the Coulomb branch corresponds to the motion of D3-branes along X^{345} and the Higgs branch corresponds to the motion of D3-branes along X^{789} . Therefore, a large class of 3d $\mathcal{N} = 4$ theories can be studied by using Type IIB brane configurations in the third configuration of Table 4.

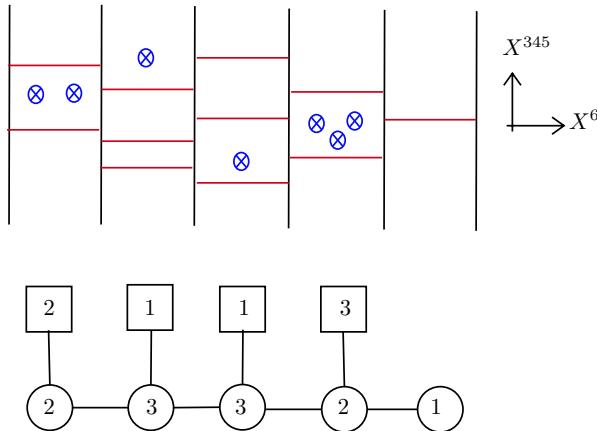


Figure 43: Type IIB brane system where NS5/D5/D3-branes are colored by black/blue/red, and the corresponding 3d $\mathcal{N} = 4$ quiver theory where circles/squares represent unitary gauge/flavor groups.

What makes this system very intriguing is a brane creation/annihilation process called **Hanany-Witten transition** [HW97]. As a D5-brane crosses an NS5-brane, a D3-brane is created or annihilated. See Figure 44. In the Hanany-Witten transition, the s-rule constrains that at most one D3-brane can be suspended between an NS5 and a D5-brane.

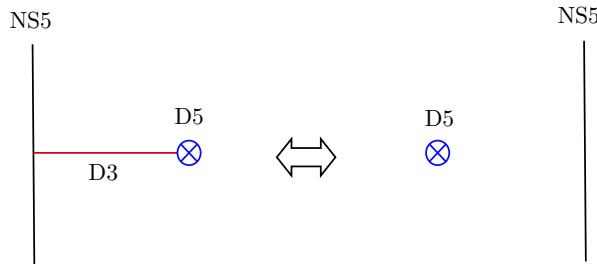


Figure 44: The crossing of NS5 and D5-brane leads to D3-brane creation/annihilation.

As in Table 4, Type IIB theory enjoys the S-duality which exchanges NS5 and D5-branes. The combination of the S-duality and the Hanany-Witten transition predicts non-trivial duality in 3d $\mathcal{N} = 4$ theories, called the 3d $\mathcal{N} = 4$ **mirror symmetry** [IS96]. One example of mirror symmetry is illustrated in Figure 45. Since the S-duality exchanges NS5 and D5-branes, so does the roles of the Coulomb and Higgs branch (namely the roles of X^{345} and X^{789}). Therefore, 3d $\mathcal{N} = 4$ supersymmetric theories can be studied from many perspectives such as branes, dualities, geometry of moduli spaces, and quantum field theories. They also admit mathematically very deep interpretations in representation theory, the geometry of hyper-Kahler manifolds, and equivalences of categories.

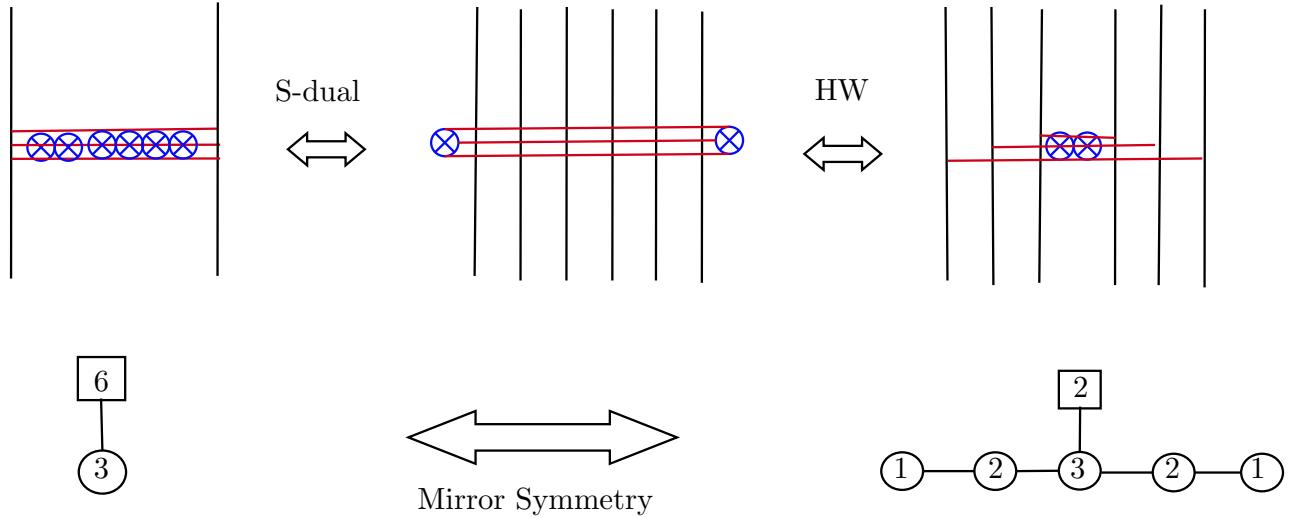


Figure 45: The S-duality maps from the left to middle brane configuration. Moving the two D5-branes into the middle involves the Hanany-Witten transition, and D3-branes are annihilated. The 3d $\mathcal{N} = 4$ theories from the left and the right brane configuration are dual to each other.

The study of supersymmetric theories by D-branes is very fruitful and successful. Brane dynamics provides profound insights into supersymmetric theories. Since the subject is very rich and broad, the reader is referred to, for instance, [HW97, DM96b, Wit97, AHK98, GK99, Joh00] as a starting point.

16 Black holes in string theory

In the previous section, we explored the remarkable properties of D-branes, including their dynamical nature and their ability to form bound states. These features make D-branes an ideal tool for studying black holes.

When a large number of D-branes is present, their collective mass can give rise to a black hole. In particular, wrapping a cycle in a compact manifold can generate a black hole configuration. The presence of open strings attaching to the D-branes leads to a significant degeneracy, which provides a statistical interpretation of the thermodynamic entropy associated with the black hole. This remarkable result, first derived by Strominger-Vafa [SV96], provides a microscopic explanation for the Bekenstein-Hawking entropy of supersymmetric black holes.

The investigation of black holes within the framework of string theory using D-branes has played a crucial role in the development of the AdS/CFT correspondence [Mal99]. This groundbreaking duality, first proposed by Maldacena, establishes a deep connection between the quantum theory of gravity in Anti-de Sitter (AdS) spacetime and conformal field theories (CFTs) defined on the boundary of the AdS space. Maldacena's Ph.D. thesis [Mal96] is an excellent reference for further exploration of this correspondence.

Before delving further into these aspects, it is important to review the basics of black holes in the framework of general relativity and the laws of black hole thermodynamics established in the early 1970s [Bek72, Bek73, BCH73, Haw75]. A comprehensive understanding of these concepts can be found in an excellent lecture note by Townsend [Tow97], and for a historical account of black hole entropy, I recommend the work by Weinstein [Wei21].

16.1 Black holes

To begin with, we consider the Einstein-Maxwell action

$$\frac{1}{16}\pi \int d^4x \sqrt{g} \left(\frac{1}{G}R - F_{\mu\nu}F^{\mu\nu} \right), \quad (16.1)$$

where G is Newton's constant. In this subsection, we shall review black hole solutions to the action (16.1) and see that they are characterized by mass M , charge Q and angular momentum J .

Schwarzschild metric

If there is no electromagnetic fields $F = 0$ in the action (16.1), the equation of motion is $R_{\mu\nu} - \frac{1}{2}g_{\mu\nu} = 0$, which has a spherically symmetric, static solution

$$ds^2 \equiv g_{\mu\nu}dx^\mu dx^\nu = -\left(1 - \frac{2GM}{r}\right)dt^2 + \left(1 - \frac{2GM}{r}\right)^{-1}dr^2 + r^2d\Omega^2,$$

where t is the time, r is the radial coordinate, and $d\Omega$ is the canonical metric of a 2-sphere. This metric describes the spacetime outside a gravitationally collapsed non-rotating star with zero electric charge, called **Schwarzschild metric**. It is well-known that the **event horizon** appears at

$$g^{rr} = 0,$$

and the sphere $r = 2GM$ is indeed the event horizon of the Schwarzschild black hole with mass M .

It turns out that much of the interesting physics having to do with the quantum properties of black holes comes from the region near the event horizon. To examine the region *near* $r = 2GM$, we analytically continued to the Euclidean metric $t = -it_E$, and we set

$$r - 2GM = \frac{x^2}{8GM}.$$

Then, the metric near the event horizon $r = 2GM$

$$ds_E^2 \approx (\kappa x)^2 dt_E^2 + dx^2 + \frac{1}{4\kappa^2}d\Omega^2,$$

where $\kappa = \frac{1}{4GM}$ is called the **surface gravity** because it is indeed the acceleration of a static particle near the horizon as measured at spatial infinity. Note that the surface gravity is defined by using Killing vector at the horizon, precisely speaking [Tow97]. The first part of the metric is just \mathbb{R}^2 with polar coordinates if we make the periodic identification

$$t_E \sim t_E + \frac{2\pi}{\kappa}.$$

Using the relation between Euclidean periodicity and temperature, we can deduce **Hawking temperature** of the Schwarzschild black hole

$$k_B T_H = \frac{\hbar \kappa}{2\pi c} = \frac{\hbar c^3}{8\pi GM}. \quad (16.2)$$

Here we restore the Boltzman constant k_B , and the speed of light c . This is a very heuristic way to introduce the Hawking temperature which was not originally found in this way.

Reissner-Nordström black hole

The most general static, spherically symmetric, charged solution of the Einstein-Maxwell theory (16.1) is

$$ds^2 = - \left(1 - \frac{2GM}{r} + \frac{GQ^2}{r^2} \right) dt^2 + \left(1 - \frac{2GM}{r} + \frac{GQ^2}{r^2} \right)^{-1} dr^2 + r^2 d\Omega^2, \quad (16.3)$$

with the electromagnetic field strength

$$F_{tr} = \frac{Q}{r^2}.$$

This solution is called the **Reissner-Nordström (RN) black hole** with mass M and charge Q . From the metric (16.3) we see that there are two event horizon for this solution where $g^{rr} = 0$ at

$$r_{\pm} = GM \pm \sqrt{(GM)^2 - GQ^2}.$$

Thus, r_+ defines the outer horizon of the black hole and r_- defines the inner horizon of the black hole. The area of the black hole is $4\pi r_+^2$. It turns out that the Hawking temperature of the RN black hole is

$$T_H = \frac{\sqrt{(GM)^2 - GQ^2}}{2\pi G \left(GM + \sqrt{(GM)^2 - GQ^2} \right)^2}.$$

For a physically sensible definition of temperature, the mass must satisfy the bound $GM^2 \geq Q^2$, and the two horizons coincide $r_+ = r_- = GM$ when this bound is saturated. In this case, the temperature of the black hole is zero and it is called an **extremal black hole**.

Kerr-Newman black hole

If we relax the static condition, black holes can have angular momentum. Hence, general stationary solutions, called **Kerr-Newman black holes**, to the action (16.1) are

described with three parameters. In **Boyer-Linquist coordinates**, the KN metric is

$$ds^2 = -\frac{(\Delta - a^2 \sin^2 \theta)}{\Sigma} dt^2 - 2a \sin^2 \theta \frac{(r^2 + a^2 - \Delta)}{\Sigma} dt d\phi \\ + \left(\frac{(r^2 + a^2)^2 - \Delta a^2 \sin^2 \theta}{\Sigma} \right) \sin^2 \theta d\phi^2 + \frac{\Sigma}{\Delta} dr^2 + \Sigma d\theta^2$$

where

$$\Sigma = r^2 + a^2 \cos^2 \theta \\ \Delta = r^2 - 2Mr + a^2 + e^2.$$

The three parameters are M , a , and e . It can be shown that

$$a = \frac{J}{M}$$

where J is the total angular momentum, while

$$e = \sqrt{Q^2 + P^2}$$

where Q and P are the electric and magnetic (monopole) charges, respectively. The Maxwell 1-form of the KN solution is

$$A_\mu dx^\mu = \frac{Qr(dt - a \sin^2 \theta d\phi) - P \cos \theta [adt - (r^2 + a^2)d\phi]}{\Sigma}.$$

16.2 Black hole thermodynamics and Bekenstein-Hawking entropy

Classically, a stationary black hole is characterized by its mass M , angular momentum J , and charge Q . This property is known as the “black hole no hair theorem”. However, Bekenstein raised an intriguing question in [Bek72]: if we consider a black hole as a purely geometric object, what happens to the entropy of the universe when we throw an object with entropy (such as a cup of tea) into the black hole? Naively, the total entropy outside the black hole would seem to decrease, which contradicts the second law of thermodynamics that states the total entropy of a closed system never decreases. To reconcile this contradiction, it suggests that black holes must possess their own entropy. The challenge then becomes how to quantify this entropy.

The key insight came from the area theorem of black holes [Haw71, Haw72], which states that the total area of the black hole horizons never decreases in any physical process. For example, when two Schwarzschild black holes with masses M_1 and M_2 merge to form a larger black hole with mass $M = M_1 + M_2$, the total area of the resulting black hole is greater than or equal to the sum of the areas of the individual black holes, as the area is proportional to the square of the mass, namely, $(M_1 + M_2)^2 \geq M_1^2 + M_2^2$. Conversely, the reverse process of a larger black hole splitting into two smaller ones is not allowed by this theorem. Motivated by the area theorem, Bekenstein proposed in [Bek72] that black holes possess entropy proportional to their horizon area.

Shortly thereafter, Bardeen, Carter, and Hawking established the analogy between the laws of black hole mechanics and the laws of thermodynamics in [BCH73]. They found precise correspondences between the laws of black hole dynamics and the three laws of thermodynamics:

- (0) Zeroth Law: In thermodynamics, the zeroth law states that the temperature T of a thermal equilibrium object is constant throughout the body. Correspondingly, for a stationary black hole, its surface gravity $\kappa = 1/4GM$ is constant over the event horizon.
- (1) First Law: The first law of thermodynamics states that energy is conserved, and the variation of energy is given by

$$dE = TdS + \mu dQ + \Omega dJ \quad (16.4)$$

where E is the energy, Q is the charge with chemical potential μ and J is the angular momentum with chemical potential Ω in the system. Correspondingly, for a black hole, the variation of its mass is given by

$$dM = \frac{\kappa}{8\pi G} dA + \mu dQ + \Omega dJ \quad (16.5)$$

where A is the area of the horizon, and κ is the surface gravity, μ is the chemical potential conjugate to Q , and Ω is the angular velocity conjugate to J .

- (2) Second Law: The second law of thermodynamics states that the total entropy S never decreases, $\delta S \geq 0$. Correspondingly, for a black hole, the area theorem states that the total area of a black hole in any process never decreases, $\delta A \geq 0$.

Laws of thermodynamics	Laws of black hole mechanics
Temperature is constant throughout a body at equilibrium. $T=\text{constant}$.	Surface gravity is constant on the event horizon. $\kappa=\text{constant}$.
Energy is conserved. $dE = TdS + \mu dQ + \Omega dJ$.	Energy is conserved. $dM = \frac{\kappa}{8\pi G} dA + \mu dQ + \Omega dJ$.
Entropy never decreases. $\delta S \geq 0$.	Area never decreases. $\delta A \geq 0$.

Table 8: Laws of black hole thermodynamics

This result can be understood as one of the highlights of general relativity. Classically, a black hole is a not only geometric but also thermodynamic object. If a black hole has energy E and entropy S , then it must also have temperature T given by

$$\frac{1}{T} = \frac{\partial S}{\partial E}.$$

For example, for a Schwarzschild black hole, the area and the entropy are proportional to M^2 . Hence, we can derive

$$\frac{1}{T} = \frac{\partial S}{\partial M} \sim \frac{\partial M^2}{\partial M} \sim M.$$

Therefore, black hole temperature is inversely proportional to mass M . The smaller a black hole is, the hotter it is! Moreover, if the black hole has temperature, it must thermally radiate like any hot body. The understanding of the thermal properties of black holes requires treatment beyond classical general relativity.

Hawking has applied techniques of quantum field theories on a curved background to the near-horizon region of a black hole and showed that a black hole indeed radiates [Haw75]. This can be intuitively understood as follows: in a quantum theory, particle-antiparticle creations constantly occur in the vacuum. Around the horizon, after pairs are created, antiparticles fall into a black hole due to the gravitational attraction whereas particles escape to the infinity. Although we do not deal with Hawking's calculation unfortunately (see [Tow97]), it indeed justifies this picture. Moreover, it revealed that the spectrum emitted by the black hole is precisely subject to the thermal radiation with temperature (16.2). Indeed, a black hole is not black at quantum level. Hence, we can treat a black hole as a thermal object, and the analogy of the laws in Table 8 can be understood as the natural consequence of the laws of thermodynamics. As a result, the formula for the Hawking temperature (16.2) and the first law of thermodynamics

$$c^2 dM = T_H dS = \frac{\kappa c^2}{8\pi G} dA,$$

lead to the precise relation between entropy and the area of the black hole:

$$S = \frac{k_B c^3 A}{4G\hbar}. \quad (16.6)$$

This is a universal result for any black hole, and this remarkable relation between the thermodynamic properties of a black hole and its geometric properties is called the celebrated **Bekenstein-Hawking entropy formula**. This formula involves all four fundamental constants of nature; (G, c, k_B, \hbar) . Also, this is the first place where the Newton constant G meets with the Planck constant \hbar . Thus, this formula shows a deep connection between black hole geometry, thermodynamics and quantum mechanics.

For ordinary objects, Boltzmann has given the statistical interpretation of the thermodynamic entropy of a system. We fix the macroscopic parameters (e.g. total electric charge, energy etc.) and count the number Ω of quantum states, known as microstates, each of which has the same values for the macroscopic parameters, and the entropy is expressed as

$$S = k_B \log \Omega.$$

Since the Bekenstein-Hawking entropy (16.6) behaves as the ordinary thermodynamic entropy in every aspect, it is therefore natural to ask whether the black hole entropy admits a statistical interpretation in the same way.

Furthermore, one of the most profound implications of Hawking's work is the notion that black holes are associated with information loss. In quantum mechanics, information is associated with pure states. If we throw a pure quantum state, such as an s-wave, into a black hole, it eventually emerges as a thermal (mixed) state. As a result, the evolution of a pure quantum state into a mixed state violates the law of quantum mechanics known as unitarity. This conundrum is referred to as the **information paradox** [Haw76]. The paradox arises due to Hawking's semi-classical analysis, where the background is fixed and particles are quantized. In fact, the information paradox stems from the absence of such a microscopic description of gravity.

In order to investigate the microscopic description of black hole entropy, we need quantum theory of gravity. This is precisely what string theorists have attempted to do and have been partially successful.

16.3 Black holes in string theory

In string theory on a d -dimensional compact manifold, branes can be wrapped in a cycle of the compact manifold and it looks like a point-like object in $(10 - d)$ -dimensional spacetime. In the regime where supergravity approximation is valid, configurations of this kind give rise to black hole solutions of the corresponding low-energy supergravity theory. Moreover, if a brane configuration preserves supersymmetry, then the corresponding solution will be an extremal supersymmetric black hole. Extremal black holes are interesting because they are stable against Hawking radiation and nevertheless have a large entropy. On the other hand, configurations without supersymmetry yield non-extremal black holes.

In general, the regime of the parameter space in which supergravity is valid is different from the regime in which the microstates counting can be performed. Thus, even if we know that a given brane configuration becomes a black hole when we go from weak to strong coupling, it is generally difficult to extract microscopic information of the black hole from the brane configuration.

For supersymmetric black holes, however, one can count the number of states at weak coupling and extrapolate the result to the black hole phase due to the BPS property. We will see that in this way, one derives the Bekenstein-Hawking entropy formula (including the precise numerical coefficient) for a 5d supersymmetric black hole [SV96]. (For more detail, I refer to [DMW02, DN12].)

D1-D5-P brane system

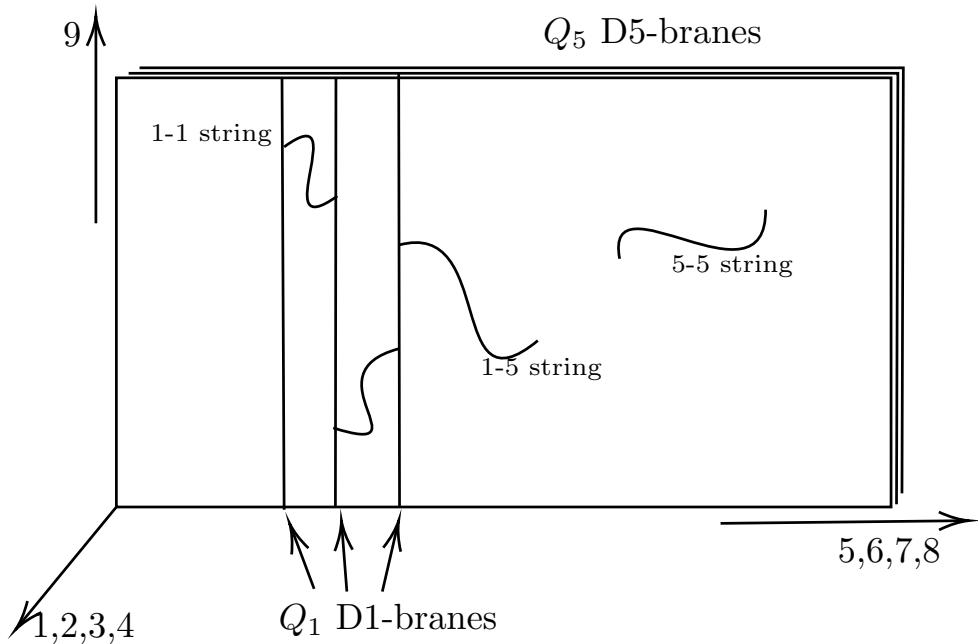


Figure 46:

Here we follow the example treated in [CM96], called the D1-D5-P brane system on $\mathbb{R}^{1,4} \times T^5$. Let us consider Type IIB compactified on a five-torus $T^5 = T^4 \times S^1$, which spans the $(x_5 \cdots x_9)$ coordinates, with Q_1 D1-branes and Q_5 D5-branes in the following

	0	1	2	3	4	5	6	7	8	9
Q_1 D1	×								×	
Q_5 D5	×					×	×	×	×	×
Q_P mom										\rightsquigarrow

configuration. We consider that the volume of T^4 is $(2\pi)^4 V$ and the radius of S^1 is R . Here we also assume that there is an excitation by open strings carrying momenta Q_P/R in the x_9 -direction. This system preserves 4 real supercharges since each constituent breaks a half of supersymmetry.

Black hole in 5d supergravity

If there are large enough D-brane charges (Q_1, Q_5, Q_P) and the five-torus is sufficiently small, the configuration produces a 5d black hole. We would like to compute the Bekenstein-Hawking entropy (16.6) of the black hole by evaluating the area of the event horizon. When 5d supergravity analysis is valid, this brane system gives rise to a 1/8-BPS black hole configuration. Ignoring the R-R field and the B -field configuration, the 5d Einstein frame metric of this solution then becomes

$$ds_5^2 = -\lambda(r)^{-\frac{2}{3}} dt^2 + \lambda(r)^{\frac{1}{3}} [dr^2 + r^2 d\Omega_3^2] ,$$

where the harmonic functions are

$$\lambda(r) = H_1(r)H_5(r)H_P(r) = \left(1 + \frac{r_1^2}{r^2}\right)\left(1 + \frac{r_5^2}{r^2}\right)\left(1 + \frac{r_P^2}{r^2}\right),$$

with

$$r_1^2 = \frac{g_s Q_1 \ell_s^6}{V}, \quad r_5^2 = g_s Q_5 \ell_s^2 \quad r_P^2 = \frac{g_s^2 Q_P \ell_s^8}{R^2 V}.$$

Let us briefly evaluate the validity of the supergravity analysis. In order for the α' corrections to geometry to be small, the radius parameters have to be large with respect to the string unit, $r_{1,5,m} \gg \ell_s$. Since we assume $V^{1/4}, R$ are an order of the string length, this implies

$$g_s Q_1 \gg 1, \quad g_s Q_5 \gg 1, \quad g_s^2 Q_P \gg 1. \quad (16.7)$$

To suppress string loop corrections, we need g_s to be small (but finite) so that the D-brane charges must be sufficiently large for supergravity analysis.

It turns out that the surface gravity and therefore the Hawking temperature of this black hole is zero, $T_H = 0$, as expected. The metric shows that the event horizon is located at $r = 0$ and the Bekenstein-Hawking entropy (16.6) is

$$\begin{aligned} S_{\text{macro}} &= \frac{A}{4G_5} = \frac{1}{4G_5} 2\pi^2 \left[r^2 \lambda(r)^{\frac{1}{3}} \right]^{\frac{3}{2}} \text{ at } r = 0 \\ &= 2 \frac{\pi^2}{4 [\pi g_s^2 \ell_s^8 / (4VR)]} (r_1 r_5 r_P)^{\frac{1}{2}} = \frac{2\pi VR}{g_s^2 \ell_s^8} \left(\frac{g_s Q_1 \ell_s^6}{V} g_s Q_5 \ell_s^2 \frac{g_s^2 Q_P \ell_s^8}{R^2 V} \right)^{\frac{1}{2}} \\ &= 2\pi \sqrt{Q_1 Q_5 Q_P}, \end{aligned} \quad (16.8)$$

where we use $G_5 = \frac{G_{10}}{(2\pi)^5 VR}$ and $16\pi G_{10} = (2\pi)^7 g_s^2 \ell_s^8$. Notice that it is also independent of R and of V while the ADM mass depends on R, V explicitly.

$$M = \frac{Q_P}{R} + \frac{Q_1 R}{g_s \ell_s^2} + \frac{Q_5 R V}{g_s \ell_s^6}.$$

Counting microstates

The next step is to identify the degeneracy of open string states of the D1-D5-P system, which can be analyzed at the limit opposite to (16.7), i.e.

$$g_s Q_1 \ll 1, \quad g_s Q_5 \ll 1, \quad g_s^2 Q_P \ll 1. \quad (16.9)$$

Further simplification can be made by taking the limit that the volume of T^4 is small as compared to the radius of the circle S^1 ,

$$V^{\frac{1}{4}} \ll R.$$

In this limit, the theory on the D-branes is an effective 2d theory on (x_1, x_9) -direction. Moreover, the smeared D1-branes plus D5-branes have a symmetry group $SO(1, 1) \times SO(4)_\parallel \times SO(4)_\perp$ where $SO(4)_\parallel \times SO(4)_\perp$ becomes R -symmetry of the 2d theory which we call $\mathcal{N} = (4, 4)$ 2d CFT. In the supersymmetric configuration, the left-movers are in their ground states so that we count excited right-movers.

Because the D1-branes are instantons in the D5-brane theory, the low-energy theory of interest is actually a σ -model on the moduli space of instantons

$$\mathcal{M} = \text{Sym}^{Q_1 Q_5}(T^4) = (T^4)^{Q_1 Q_5} / S_{Q_1 Q_5}.$$

The central charge of this 2d CFT is

$$c = n_{\text{bose}} + \frac{1}{2} n_{\text{fermi}} = 6 Q_1 Q_5. \quad (16.10)$$

Roughly, this central charge c can be thought of as coming from having $Q_1 Q_5$ 1-5 strings that can move in the 4 directions of T^4 . Although this orbifold theory has many twisted sectors, the special point of the moduli space corresponds to a single string winding $Q_1 Q_5$ times. It turns out that counting the excitations of this **long string** is only relevant in the limit of large D-brane charges. For this long string, the level-matching condition is

$$N - \bar{N} = Q_P W, \quad W = Q_1 Q_5, \quad \rightarrow \quad N = Q_P Q_1 Q_5$$

where the left-movers are in the ground states $\bar{N} = 0$.

If N_m^i and n_m^i denote occupation numbers of the four compact (T^4) bosonic and fermionic oscillators, respectively, then evaluation of N gives

$$Q_P W = \sum_{i=1}^4 \sum_{m=1}^{\infty} m (N_m^i + n_m^i) \quad (16.11)$$

The degeneracy $\Omega(Q_1, Q_5, Q_P)$ is then given by the number of choices for N_m^i and n_m^i subject to (16.11).

The partition function of this system is the partition function for 4 bosons and an equal number of fermions

$$Z = \left[\prod_{m=1}^{\infty} \frac{1+q^m}{1-q^m} \right]^4 \equiv \sum \Omega(Q_1, Q_5, Q_P) q^N,$$

where $\Omega(Q_1, Q_5, Q_P)$ is the degeneracy of states at the level $N = Q_P Q_1 Q_5$. The Cardy formula [Car86] can be applied in the regime in which the KK momenta $Q_P \gg Q_1 Q_5$ much larger than the central charge (16.10) while assuming (16.9):

$$\Omega(Q_1, Q_5, Q_P) \sim \exp \left(2\pi \sqrt{Q_1 Q_5 Q_P} \right) = \exp \left(2\pi \sqrt{\frac{c}{6} Q_P} \right)$$

Therefore the microscopic D-brane statistical entropy is

$$S_{\text{micro}} = \log(\Omega(Q_1, Q_5, Q_P)) = 2\pi\sqrt{Q_1 Q_5 Q_P}.$$

This agrees exactly with the black hole result (16.8)!

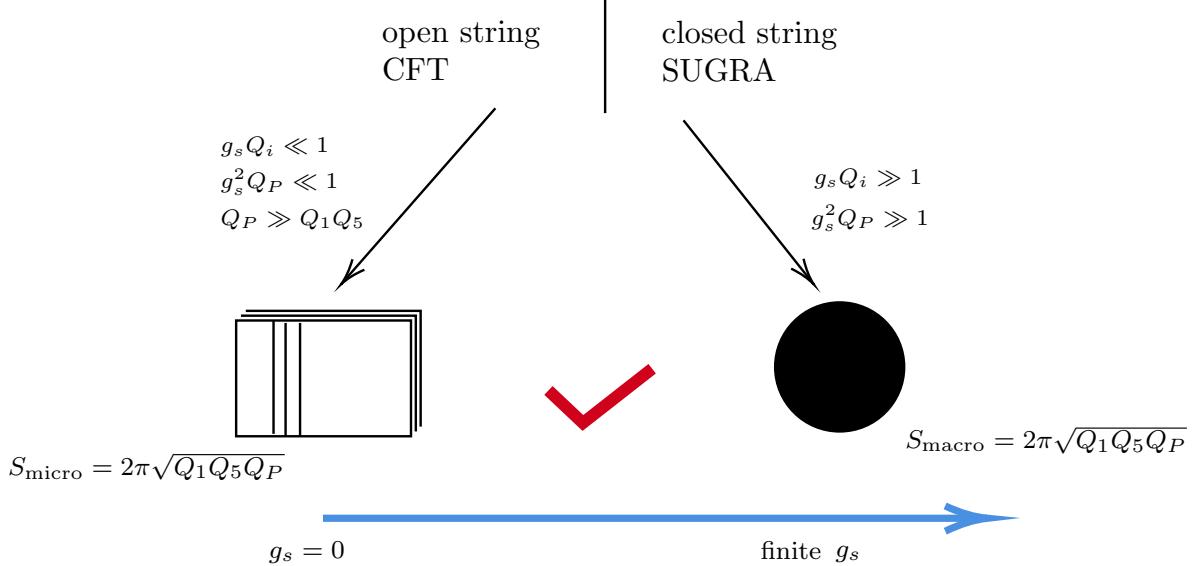


Figure 47:

More results

By coupling the low energy degrees of freedom in the D1-D5-p system to supergravity modes (therefore perturbing the extremal condition), one can also compute the rate of Hawking radiation from a black hole that agrees precisely with the Hawking calculation. Thus this provides a microscopic explanation of Hawking radiation. (See [DMW02, §8].)

In fact, vigorous research in the last decade has shown that one can show the exact match between macroscopic and microscopic calculations of black hole entropy even in finite D-brane charges.

Moreover, a generalization of Bekenstein-Hawking entropy has been proposed in [RT06] that connects quantum theory of gravity and quantum information theory. A recent study has clearly suggested that quantum entanglement must have something to do with quantum physics of spacetime.

17 Introduction to AdS/CFT correspondence

The study of black holes in string theory by using D-branes has led to the celebrated AdS/CFT correspondence [Mal99]. The AdS/CFT correspondence is the equivalence between a string theory or M-theory on an anti-de Sitter background and a conformal field theory. It has shed new light on quantum gravity as well as strongly coupled quantum field theories. Although it was proposed in the framework of string theory, it has already been studied beyond string theory, influencing other physical theories. It has attracted

a large number of researchers, and it is connected to many branches of physics. For the basics of the AdS/CFT correspondence, the most famous review is [AGM⁺00], and there are many other reviews and books, e.g. [DF02, Mal03, Ram15, Năs15].

We shall first study the basic properties of conformal field theories in general dimensions and geometry of anti-de Sitter space. Then, we will deal with the most famous example, Type IIB on $\text{AdS}_5 \times S^5 / 4\text{d } \mathcal{N} = 4 \text{ SYM}$.

17.1 Conformal field theory

Conformal group

A conformal field theory (CFT) is a quantum field theory that is invariant under conformal transformations. In §3, we have studied the conformal transformation for two dimensions, which is a special case. Here, we study a conformal group for arbitrary dimensions (assume $d \geq 3$).

A conformal group is defined by transformations that preserve the metric up to a local scale factor:

$$g_{\mu\nu}(x) \rightarrow g'_{\mu\nu}(x') = \Omega^2(x) g_{\mu\nu}(x) .$$

In an infinitesimal form ($x'^\mu = x^\mu + \epsilon^\mu$) it is (compare with (3.10))

$$\partial_\mu \epsilon_\nu + \partial_\nu \epsilon_\mu = \frac{2}{d} \eta_{\mu\nu} \partial \cdot \epsilon . \quad (17.1)$$

Applying ∂^μ to (17.1) we have

$$\left(1 - \frac{2}{d}\right) \partial_\nu \partial \cdot \epsilon + \square \epsilon_\nu = 0 ,$$

where $\square = \partial \cdot \partial$. From this expression, we can see that $d = 2$ is quite special, and it leads $\partial^2 \epsilon = \partial_- \partial_+ \epsilon$, which gives infinitely many transformations. We further apply ∂^ν and reach

$$(d-1) \square \partial \cdot \epsilon = 0 .$$

This expression implies ϵ_μ is up to quadratic order of x :

$$\epsilon_\mu = a_\mu + b_{\mu\nu} x^\nu + c_{\mu\nu\rho} x^\nu x^\rho .$$

Plugging these expressions back to the definition equations above (and their variants), we have the following constraints

$$\begin{aligned} b_{\mu\nu} &= \alpha \eta_{\mu\nu} + M_{\mu\nu} \quad (M_{\mu\nu} = -M_{\nu\mu}) , \\ c_{\mu\nu\rho} &= \eta_{\mu\nu} f_\rho + \eta_{\mu\rho} f_\nu - \eta_{\nu\rho} f_\mu \quad (f_\mu = \frac{1}{d} c_{\rho\mu}^\rho) . \end{aligned}$$

The parameters above correspond to transformations you are familiar with except f_μ , which is called special conformal transformation (SCT). See Table 17.1 for the summary.

The generators summarized in the table form conformal group commutation relations

$$\begin{aligned} [J_{ab}, J_{cd}] &= i (\eta_{ad} J_{bc} + \eta_{bc} J_{ad} - \eta_{ac} J_{bd} - \eta_{bd} J_{ac}) , \\ J_{\mu\nu} &= L_{\mu\nu} , \quad J_{(d+1)d} = D , \\ J_{\mu d} &= \frac{1}{2} (K_\mu - P_\mu) , \quad J_{\mu(d+1)} = \frac{1}{2} (K_\mu + P_\mu) , \end{aligned} \quad (17.2)$$

Names	Finite transf.	Generators	Dim.
Translation	$x'^\mu = x^\mu + a^\mu$	$P_\mu = -i\partial_\mu$	+1
Dilatation	$x'^\mu = \alpha x^\mu$	$D = -ix \cdot \partial$	0
Lorentz/Rotation	$x'^\mu = M^\mu_\nu x^\nu$	$L_{\mu\nu} = -i(x_\mu \partial_\nu - x_\nu \partial_\mu)$	0
SCT	$x'^\mu = \frac{x^\mu - (x \cdot x) f^\mu}{1 - 2f \cdot x + (f \cdot f)(x \cdot x)}$	$K_\mu = -i(2x_\mu x \cdot \partial - (x \cdot x)\partial_\mu)$	-1

Table 9: Generators of the conformal group.

where note that $a, b, c, d = 0, 1, \dots, d+1$ and $\mu, \nu = 0, 1, \dots, d-1$, and η_{ab} is $\text{diag}(-, +, \dots, +, -)$ for Lorentzian spacetime and $\text{diag}(+, +, \dots, +, -)$ for Euclidean space. The algebras are isomorphic to those of $\text{SO}(2, d)$ and $\text{SO}(1, d+1)$, respectively. Note that the SCT can be understood as an inversion ($x^\mu \rightarrow \frac{x^\mu}{x \cdot x}$) with translation $\frac{x'^\mu}{x' \cdot x'} = \frac{x^\mu}{x \cdot x} - f^\mu$. However, the inversion is not included in the algebra (since the inversion is a discrete transformation). The commutation relations with the dilatation D indicate that P^μ has conformal dimension 1, K^μ has -1, and therefore $L_{\mu\nu}$ has 0:

$$[D, P_\mu] = iP_\mu, \quad [D, K_\mu] = -iK_\mu, \quad [P_\mu, K_\nu] = 2i(\eta_{\mu\nu}D - L_{\mu\nu}). \quad (17.3)$$

The other non-trivial ones are

$$\begin{aligned} [P_\rho, L_{\mu\nu}] &= i(\eta_{\rho\mu}P_\nu - \eta_{\rho\nu}P_\mu), & [K_\mu, L_{\nu\rho}] &= i(\eta_{\mu\nu}K_\rho - \eta_{\mu\rho}K_\nu), \\ [L_{\mu\nu}, L_{\rho\sigma}] &= i(\eta_{\nu\rho}L_{\mu\sigma} + \eta_{\mu\sigma}L_{\nu\rho} - \eta_{\mu\rho}L_{\nu\sigma} - \eta_{\nu\sigma}L_{\mu\rho}), \end{aligned}$$

Primary fields

In 2-dimensions, we defined the primary field by using OPE (3.20) with energy-momentum tensor (it leads to a representation of the conformal group). For general dimensions, we define it by a formal representation of the conformal group. It is known that the representation is characterized by an eigenvalue of the dilatation operator $-i\Delta$ (Δ is called the **scaling dimension** of the field, rather than **weight**), and representation of the Lorentz group. The former statement means that $\Phi(x) \rightarrow \Phi'(\lambda x) = \lambda^{-\Delta}\Phi(x)$. The commutation relations (17.3)) tell us that P_μ is the raising operator, while K_μ is the lowering operator. Therefore, there are operators annihilated by K_μ in each finite-dimensional representation of the conformal group. Such an operator is called **primary operator/field** (we use operator and field interchangeably). The action of the conformal group on the primary field is

$$\begin{aligned} [P_\mu, \Phi(x)] &= -i\partial_\mu\Phi(x), \\ [L_{\mu\nu}, \Phi(x)] &= [-i(x_\mu \partial_\nu - x_\nu \partial_\mu) + \Sigma_{\mu\nu}]\Phi(x), \\ [D, \Phi(x)] &= -i(\Delta + x^\mu \partial_\mu)\Phi(x), \\ [K_\mu, \Phi(x)] &= [-i(x^2 \partial_\mu - 2x_\mu x^\nu \partial_\nu + 2x_\mu \Delta) - 2x^\nu \Sigma_{\mu\nu}]\Phi(x), \end{aligned}$$

where $\Sigma_{\mu\nu}$ are the matrices of a finite-dimensional representation of the Lorentz group, acting on the indices of the Φ field (e.g. it is $\frac{i}{2}\Gamma_{\mu\nu}$ for a spinor). There are some comments on primary fields and others:

- Fields created by acting P_μ on a primary field are called **descendant fields**.

- Fields are not, in general, by eigenfunctions of the Hamiltonian P^0 , or the mass operator $-P \cdot P$, and hence, they have a continuous spectrum.
- In unitary field theories, the scale dimension is bounded from below (**unitary bound**). It is $\Delta \geq (d-2)/2$ for scalars, $\Delta \geq (d-1)/2$ for spinors, and $\Delta \geq d+s-2$ for spin- s fields for $s \geq 1$. (We refer to the derivation of the bounds and other details to [Qua15, Sec. 2].)

17.2 Anti-de Sitter space

An anti-de Sitter (AdS) space is a maximally symmetric manifold with constant negative scalar curvature. It is a solution of Einstein's equations for an empty universe with negative cosmological constant. The easiest way to understand it is as follows.

A Lorentzian AdS_{d+1} space can be illustrated by the hyperboloid in $(2, d)$ Minkowski space:

$$X_0^2 + X_{d+1}^2 - \sum_{i=1}^d X_i^2 = R^2. \quad (17.4)$$

The metric can be naturally induced from the Minkowski space

$$ds^2 = -dX_0^2 - dX_{d+1}^2 + \sum_{i=1}^d dX_i^2.$$

By construction, it has $\text{SO}(2, d)$ isometry, which is the first connection to the conformal group in d -dim.

Global coordinate

A simple solution to (17.4) is given as follows.

$$\begin{aligned} X_0^2 + X_{d+1}^2 &= R^2 \cosh^2 \rho, \\ \sum_{i=1}^d X_i^2 &= R^2 \sinh^2 \rho. \end{aligned}$$

Or, more concretely,

$$\begin{aligned} X_0 &= R \cosh \rho \cos \tau, & X_{d+1} &= R \cosh \rho \sin \tau, \\ X_i &= R \sinh \rho \Omega_i \quad (i = 1, \dots, d, \text{ and } \sum_i \Omega_i^2 = 1). \end{aligned}$$

These are S^1 and S^{d-1} with radii $R \cosh \rho$ and $R \sinh \rho$, respectively. The metric is

$$ds^2 = R^2 \left(-\cosh^2 \rho d\tau^2 + d\rho^2 + \sinh^2 \rho d\Omega_{(d-1)}^2 \right).$$

Note that τ is a periodic variable and if we take $0 \leq \tau < 2\pi$, the coordinate wraps the hyperboloid precisely once. This is why this coordinate is called the **global coordinate**. The manifest sub-isometries are $\text{SO}(2)$ and $\text{SO}(d)$ of $\text{SO}(2, d)$. To obtain a causal space-time, we simply unwrap the circle S^1 , namely, we take the region $-\infty < \tau < \infty$ with no identification, which is called the **universal cover** of the hyperboloid.

In literature, another global coordinate is also used, which can be derived by redefinitions $r \equiv R \sinh \rho$ and $dt \equiv R d\tau$:

$$ds^2 = -f(r) dt^2 + \frac{1}{f(r)} dr^2 + r^2 d\Omega_{(d-1)}^2, \quad f(r) = 1 + \frac{r^2}{R^2}.$$

Poincaré coordinates

There is yet another coordinate system, called the **Poincaré coordinates**. As opposed to the global coordinate, this coordinate covers only half of the hyperboloid. It is most easily (but naively) seen in $d = 1$ case:

$$x^2 - y^2 = R^2 ,$$

which is the hyperbolic curve. The curve consists of two isolated parts in regions $x > R$ and $x < -R$. We simply use one of them to construct the coordinate.

Let us get back to general d -dim. We define the coordinate as follows.

$$\begin{aligned} X_0 &= \frac{1}{2u} (1 + u^2 (R^2 + x_i^2 - t^2)) , \\ X_i &= Rux_i \quad (i = 1, \dots, d-1) , \\ X_d &= \frac{1}{2u} (1 - u^2 (R^2 - x_i^2 + t^2)) , \\ X_{d+1} &= Rut , \end{aligned}$$

where $u > 0$. As it is stated, the coordinate covers half of the hyperboloid; in the region, $X_0 > X_d$. The metric is

$$ds^2 = R^2 \left(\frac{du^2}{u^2} + u^2 (-dt^2 + dx_i^2) \right) = R^2 \left(\frac{du^2}{u^2} + u^2 dx_\mu^2 \right) . \quad (17.5)$$

The coordinates (u, t, x_i) are called the **Poincaré coordinates**. This metric has manifest $ISO(1, d-1)$ and $SO(1, 1)$ sub-isometries of $SO(2, d)$; the former is the Poincaré transformation and the latter corresponds to the dilatation

$$(u, t, x_i) \rightarrow (\lambda^{-1}u, \lambda t, \lambda x_i) .$$

If we further define $z = 1/u$ ($z > 0$), then,

$$ds^2 = \frac{R^2}{z^2} (dz^2 + dx_\mu^2) . \quad (17.6)$$

This is called the **upper (Poincaré) half-plane model**. The hypersurface given by $z = 0$ is called the **(asymptotic) boundary** of the AdS space, which corresponds to $u \sim r \sim \rho = \infty$.

17.3 Introduction to $\text{AdS}_5/\text{CFT}_4$ correspondence

Now let us study the most well-studied example of AdS/CFT correspondence, which is the equivalence between 4d $U(N)$ $\mathcal{N} = 4$ super-Yang-Mills (SYM) and Type IIB string theory on $\text{AdS}_5 \times S^5$, which arises from the large number of D3-branes. Type IIB string theory with D3-branes contains two kinds of perturbative excitations, closed strings and open strings. If we consider the system at low energies, energies lower than the string scale $1/\ell_s$, then only the massless string states can be excited. The closed string massless states give a gravity supermultiplet in $D = 10$ in Type IIB supergravity as in §12.4. The open string massless states give an $\mathcal{N} = 4$ vector multiplet in $D = 4$, and their low-energy effective theory is $\mathcal{N} = 4$ $U(N)$ SYM. Therefore, the duality can also be understood as **open/closed duality**.

$\mathcal{N} = 4$ super-Yang-Mills theory

The low-energy effective theory of N D3-branes is 4d $\mathcal{N} = 4$ U(N) SYM theory so we describe the basic properties of the $\mathcal{N} = 4$ SYM. The action can be obtained by the dimensional reduction from the 10d $\mathcal{N} = 1$ SCFT on $\mathbb{R}^{1,3} \times T^6$ where the 10d Lorentz group SO(1, 9) is decomposed to $\text{SO}(1, 3) \times \text{SO}(6) \subset \text{SO}(1, 9)$:

$$\begin{aligned} S &= -\frac{1}{g_{YM}^2} \int d^{10}x \text{Tr} \left[\frac{1}{4} F_{MN}^2 + \frac{i}{2} \bar{\lambda} \Gamma^M D_M \lambda \right] \\ &= -\frac{1}{g_{YM}^2} \int d^4x \text{Tr} \left[\frac{1}{4} F_{\mu\nu}^2 + \frac{1}{2} (D_\mu X_m)^2 + \frac{i}{2} \bar{\lambda} \Gamma^\mu D_\mu \lambda - \frac{g_{YM}}{2} \bar{\lambda} \Gamma^m [X_m, \lambda] - \frac{g_{YM}^2}{4} [X_m, X_n]^2 \right] \end{aligned} \quad (17.7)$$

where the ten-dimensional gauge fields A_M , $M = 0, \dots, 9$ split the 4d gauge field A_μ , $\mu = 0, \dots, 3$ and 6 scalars X_m , $m = 1, \dots, 6$, and λ is a 10d Majorana-Weyl spinor dimensionally reduced to 4d. We can also add the topological term

$$S_{\text{top}} = \frac{i\theta}{32\pi^2} \int d^4x \epsilon^{\mu\nu\rho\sigma} \text{Tr}(F_{\mu\nu} F_{\rho\sigma})$$

The action is invariant under the supersymmetry transformation

$$\begin{aligned} \delta X^m &= -\bar{\epsilon} \Gamma^m \lambda \\ \delta A^\mu &= -\bar{\epsilon} \Gamma^\mu \lambda \\ \delta \lambda &= \left(\frac{1}{2} F_{\mu\nu} \Gamma^{\mu\nu} + D_\mu X_m \Gamma^{\mu m} + \frac{i}{2} [X_m, X_n] \right) \epsilon. \end{aligned} \quad (17.8)$$

It is easy to see that 4d $\mathcal{N} = 4$ SYM is classically **conformal invariant**. because the mass dimensions of the fields

$$[A_\mu] = [X^i] = 1, \quad [\lambda_a] = \frac{3}{2}, \quad (17.9)$$

so that the coupling constant is dimensionless: $[g] = [\theta] = 0$. However, one has to be careful at the quantum level because quantum correction generally breaks the conformal invariance. To be conformal invariant at the quantum level, the beta function of the coupling constant has to vanish $\beta = 0$. It turns out that 4d $\mathcal{N} = 4$ SYM is the case, and hence, it is quantum mechanically conformal. The $\mathcal{N} = 4$ supersymmetry combined with conformal symmetry forms the superconformal group SU(2, 2|4) which consists of the following generators

- **Conformal Symmetry** is $\text{SO}(2, 4) \cong \text{SU}(2, 2)$ in $d = 4$, as we have seen in §17.1. The generators consist of translations P^μ , Lorentz transformations $L_{\mu\nu}$, dilatations D and special conformal transformations K^μ with the relations (17.2);
- **R-symmetry** is $\text{SO}(6)_R \cong \text{SU}(4)_R$ which is manifest from the 10d viewpoint, and R-symmetry rotates the 6 scalar X^m ($m = 1, \dots, 6$);
- **Poincaré supersymmetries** are generated by the supercharges $Q_\alpha^I, \bar{Q}_{\dot{\alpha}}^I$, $I = 1, \dots, 4$ that transform under the 4 of $\text{SU}(4)_R$. They can be understood as a “square root” of P_μ . Type IIB string theory has 32 supercharges, and D3-branes break half of the supersymmetries. Consequently, the 16 preserved supercharges are indeed $Q_\alpha^I, \bar{Q}_{\dot{\alpha}}^I$, which form $\mathcal{N} = 4$ Poincaré supersymmetry;

- **Conformal supersymmetries:** are generated by the fermionic generators S_α^I and $\bar{S}_{\dot{\alpha}}^I$ that are superconformal partners of $Q_\alpha^I, \bar{Q}_{\dot{\alpha}}^I$. They can be understood as a “square root” of K_μ .

Therefore, there are 32 supercharges Q, \bar{Q}, S, \bar{S} in total and they obey the anti-commutation relations

$$\begin{aligned} \{Q_A^\alpha, \bar{Q}^{\dot{\alpha}B}\} &= P^{\alpha\dot{\alpha}} \delta_A^B, \\ \{S_\alpha^A, \bar{S}_{\dot{\alpha}B}\} &= K_{\alpha\dot{\alpha}} \delta_B^A, \\ \{S_\alpha^A, Q_B^\beta\} &= \delta_B^A M_\alpha^\beta + \delta_\alpha^\beta R_B^A + \delta_B^A \delta_\alpha^\beta \frac{D}{2}, \\ \{\bar{S}_{\dot{\alpha}A}, \bar{Q}^{\dot{\beta}B}\} &= \delta_A^B \bar{M}_{\dot{\alpha}}^{\dot{\beta}} - \delta_{\dot{\alpha}}^{\dot{\beta}} R_A^B + \delta_A^B \delta_{\dot{\alpha}}^{\dot{\beta}} \frac{D}{2}. \end{aligned} \quad (17.10)$$

The $\mathcal{N} = 4$ SYM enjoys **S-duality** [GNO77, MO77] that is the $\text{SL}(2, \mathbb{Z})$ action on the complexified coupling constant $\tau \equiv \frac{\theta}{2\pi} + \frac{4\pi i}{g_{YM}^2}$

$$\tau \rightarrow \frac{a\tau + b}{c\tau + d} \quad \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \text{SL}(2, \mathbb{Z}). \quad (17.11)$$

The $\mathcal{N} = 4$ SYM is the world-volume theory on a stack of D3-branes, and D3-branes are invariant under the $\text{SL}(2, \mathbb{Z})$ symmetry of Type IIB theory as seen in §13.1. In fact, this is the origin of the $\text{SL}(2, \mathbb{Z})$ symmetry (17.11).

There is another way to realize the $\mathcal{N} = 4$ SYM from string theory. Indeed, the M5-branes wrapped on a torus with complex structure τ give rise to the $\mathcal{N} = 4$ SYM with the complexified coupling constant τ . Here, τ manifestly admits a geometric origin as the complex structure of a torus. Note that when $\theta = 0$, the S-duality transformation amounts to $g_{YM} \rightarrow 1/g_{YM}$, thereby exchanging strong and weak coupling.⁸

Near-horizon geometry of D3-branes

To gain a deeper understanding of geometry, let us now study the closed string side of the system. A system of N coincident D3-branes is a classical solution of the low-energy string effective action, characterized by the metric and R-R field.

The metric of the D3-brane solution is given by:

$$ds^2 = H(y)^{-\frac{1}{2}} \eta_{\mu\nu} dx^\mu dx^\nu + H(y)^{\frac{1}{2}} (dy^2 + d\Omega_5^2), \quad (17.12)$$

where y is the coordinate transverse to the brane and $d\Omega_5^2$ represents the metric on a five-sphere. The function $H(y)$ is given by:

$$H(y) = 1 + \frac{R^4}{y^4}, \quad \text{where} \quad R^4 = 4\pi g_s N(\alpha')^2, \quad (17.13)$$

with g_s being the string coupling constant and α' representing the Regge slope.

In the region where $y \gg R$, the D3-brane solution smoothly transitions to flat space-time $\mathbb{R}^{1,9}$, which is the Minkowski space.

On the other hand, when $y < R$, the geometry undergoes a significant transformation and is often referred to as the “throat”. Initially, it may seem that the geometry becomes

⁸Precisely speaking, the electromagnetic duality of the $\mathcal{N} = 4$ SYM depends on a choice of gauge groups, and the duality group is usually a congruence subgroup of $\text{SL}(2, \mathbb{Z})$. For more detail, we refer to [AST13].

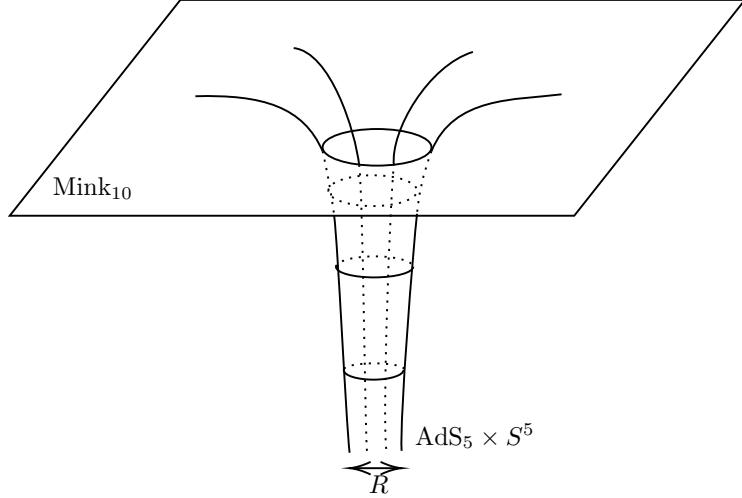


Figure 48: Minkowski region and near-horizon region

singular as y becomes much smaller than R . However, this is not the case. By examining the near-horizon region described by the following limit

$$y \rightarrow 0 \quad \alpha' \rightarrow 0 \quad u \equiv y/R^2 \quad (17.14)$$

where α' is taken to zero while u remains fixed. In this limit, we can safely neglect the factor 1 in the function $H(y)$ appearing in (17.13). Consequently, the metric simplifies to:

$$ds^2 = R^2 \left[u^2 \eta_{\mu\nu} dx^\mu dx^\nu + \frac{du^2}{u^2} + d\Omega_5^2 \right] \quad (17.15)$$

As we have previously observed in (17.5), the first part of the metric corresponds to AdS_5 , while the second part corresponds to S^5 . Remarkably, in the vicinity of the brane where $y \sim 0$ or equivalently $u \sim 0$, the geometry exhibits remarkable regularity and symmetry, manifesting as $\text{AdS}_5 \times S^5$ with the same radii

$$R_{\text{AdS}_5}^2 = R_{S^5}^2 = \alpha' \sqrt{4\pi N g_s} .$$

where N represents the number of coincident D3-branes and g_s denotes the string coupling constant.

In this limit described by (17.14), the dynamics in the asymptotically flat region of the D3-brane geometry decouples from the theory. The only surviving region is the AdS portion of the geometry. Moreover, the interactions between the bulk and the brane dynamics become negligibly small. This limiting regime is aptly referred to as the “decoupling limit”. By studying the decoupling limit, we can focus solely on the AdS region of the D3-brane and unravel its remarkable connections to conformal field theory (CFT) through the AdS/CFT correspondence.

The AdS/CFT correspondence

As mentioned, the world volume theory of N coincident D3-branes is 4d $\mathcal{N} = 4$ SYM with $U(N)$ gauge group. On the other hand, the classical solution in (17.15) is a good

approximation when the radii of AdS_5 and S^5 are very big:

$$\frac{R^2}{\alpha'} \gg 1 \implies N g_{YM}^2 \equiv \lambda \gg 1 \quad (17.16)$$

The fact that those two descriptions are simultaneously consistent for large values of the coupling constant λ brought Maldacena to formulate the conjecture that the strongly interacting $\mathcal{N} = 4$ SYM with gauge group $U(N)$ at large N is equivalent to Type IIB supergravity compactified on $\text{AdS}_5 \times S^5$. However, supergravity is not a consistent quantum theory and it is just a low-energy effective theory of string theory. Hence, the natural way to extend the equivalence at any value of λ is therefore that $\mathcal{N} = 4$ SYM is equivalent to Type IIB string theory on $\text{AdS}_5 \times S^5$ [Mal99]. Namely, the following two theories are dual to each other:

- $\mathcal{N} = 4$ super-Yang-Mills theory in 4-dimensions with gauge group $U(N)$.
- Type IIB superstring theory on $\text{AdS}_5 \times S^5$ with the same radius R as in (17.16), where the 5-form G_5^+ has integer flux $N = \int_{S^5} G_5^+$ on S^5 .

The coupling constants in the two theories are related by $g_s = g_{YM}^2$. The precise formulation of this duality will follow in the next subsection. In these two theories, we can immediately find the following correspondence as in Table 10.

4d $\mathcal{N} = 4$ SYM	Type IIB on $\text{AdS}_5 \times S^5$
32 supercharges	32 supercharges
$SO(2, 4)$ conformal group	$SO(2, 4)$ isometry of AdS_5
$SU(4)_R$ symmetry	$SO(6)$ isometry of S^5
$SL(2, \mathbb{Z})$ symmetry of coupling constants	$SL(2, \mathbb{Z})$ symmetry of axio-dilaton

Table 10: Dictionary for the $\text{AdS}_5/\text{CFT}_4$ correspondence

This conjecture is the most general statement, which is valid at any values of coupling constant $g_s = g_{YM}^2$ and rank N . However, it is still difficult to quantize string theory at any value of g_s on a general manifold, including asymptotic AdS space. Hence, it is still an open problem to prove this general statement of the conjecture. Nevertheless, taking various limits of the conjecture, we can show a variety of non-trivial evidence for the conjecture, providing new physical insight.

The ‘t Hooft Limit: The ‘t Hooft limit [tH74] is the limit in which we keep the **‘t Hooft coupling** $\lambda \equiv g_{YM}^2 N = g_s N$ fixed and letting $N \rightarrow \infty, g_s \rightarrow 0$. As in Figure 49, the planar diagrams become dominant in this limit on the Yang-Mills side. On the AdS side, since the string coupling can be re-expressed in terms of the ‘t Hooft coupling as $g_s = \lambda / N$, the ‘t Hooft limit corresponds to the regime where weak coupling string perturbation theory is valid. Put differently, this limit of the AdS/CFT correspondence can be understood as the incarnation of the idea of ‘t Hooft [tH74].

Supergravity limit: While we take $N \rightarrow \infty$, in the regime that ‘t Hooft coupling is large $\lambda = g_s N \gg 1$, supergravity description becomes reliable. On the gauge theory side, the theory is strongly-coupled so that perturbation techniques cannot be applied. The two theories are conjectured to be the same, but when one side is weakly coupled, the

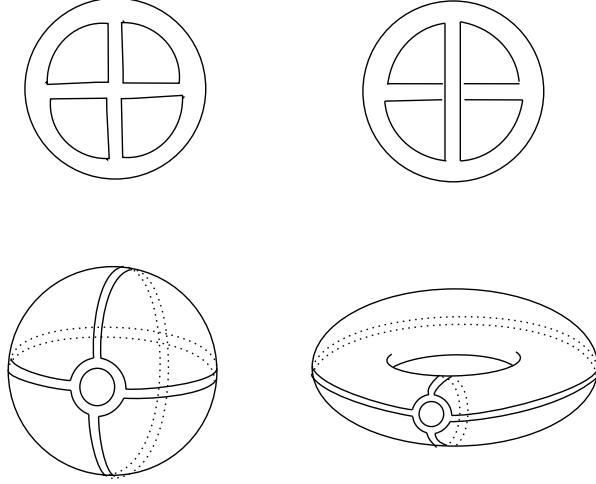


Figure 49: Planar and non-planar diagram

other is strongly coupled. This is a salient feature of **duality**. Thus, using the AdS/CFT correspondence, analyses of supergravity provide new insights to strongly-coupled SYM theory, such as quark confinement and mass gap.

17.4 GKPW relation

After Maldacena's proposal [Mal99], a more precise formulation was provided in [GKP98, Wit98a]. It states that the gravitational partition function on asymptotically AdS space is equal to the generating function of correlation functions in the corresponding conformal field theory (CFT). This relation is known as the Gubser-Klebanov-Polyakov-Witten (GKPW) relation and can be expressed as follows:

$$Z_{\text{grav}}[\phi \rightarrow \phi_0] = \left\langle \exp \left(\int_{\partial AdS} \bar{\phi}_0 \mathcal{O} \right) \right\rangle_{\text{CFT}} \quad (17.17)$$

where $Z_{\text{grav}}[\phi \rightarrow \phi_0]$ denotes the gravitational partition function, ϕ represents a bulk field in the gravity theory on AdS, and \mathcal{O} corresponds to the dual operator in the CFT. The generating function of correlation functions of \mathcal{O} is evaluated in the CFT.

The gravitational partition function can be schematically written as

$$Z_{\text{grav}}[\phi \rightarrow \phi_0] = \int_{\phi \rightarrow \phi_0} \mathcal{D}\phi e^{-S_{\text{string}}[\phi]} .$$

For instance, in the regime $\lambda \gg 1$, we can use supergravity description

$$Z_{\text{grav}}[\phi \rightarrow \phi_0] = \sum_{\text{saddle point}} e^{-S_{\text{SUGRA}}[\phi \rightarrow \phi_0]} ,$$

where the sum is taken over the saddle points of the supergravity action.

Bulk field/boundary operator

The GKPW relation tells us that each field propagating in the bulk AdS space is in one-to-one correspondence with an operator in CFT. Also, the spin of the bulk field is

equal to that of the CFT operator. Moreover, the mass of the bulk field fixes the scaling dimension of the CFT operator. Here are some examples:

- By definition, every gravitational theory has the graviton $g_{\mu\nu}$, a massless spin-2 particle. The dual operator must be the universal one with spin-2 in CFT. In fact, there is a natural candidate for it: the energy-momentum tensor $T_{\mu\nu}$ in CFT. The fact that the graviton is massless corresponds to the fact that the CFT stress tensor is conserved.
- If the theory of gravity includes a spin-1 vector field A_μ , the dual operator in the CFT is also a spin-1 operator J_μ . When A_μ is a gauge field, J_μ represents a conserved current in the CFT. The GKPW relation naturally gives rise to the coupling $e^{i \int A_\mu J^\mu}$, indicating that gauge symmetries in the bulk correspond to global symmetries in the CFT.
- A bulk scalar field is dual to a scalar operator in the CFT. The boundary value of the bulk scalar field acts as a source in the CFT.

There is a relation between the mass of the field ϕ and the scaling dimension of the corresponding operator in CFT. Considering the AdS_{d+1} metric in the form of (17.6), the wave equation for a field with mass m in AdS_{d+1} space has two independent solutions that behave as $z^{d-\Delta}$ (**non-normalizable**) and z^Δ (**normalizable**) as z approaches zero (close to the boundary of AdS) where

$$\Delta = \frac{d}{2} + \sqrt{\frac{d^2}{4} + R^2 m^2}. \quad (17.18)$$

Therefore, as the massive field approaches the boundary $\epsilon \rightarrow 0$, the field on the right-hand side of (17.17) behaves as

$$\phi(\vec{x}, \epsilon) = \epsilon^{d-\Delta} \phi_0(\vec{x}). \quad (17.19)$$

This implies that ϕ_0 has a conformal dimension of $d - \Delta$, which in turn means that the dual operator \mathcal{O} in (17.17) has a conformal dimension of Δ as given by (17.18). This correspondence is consistent with the interpretation that the radial direction z in (17.6) of the bulk AdS space corresponds to the scaling behavior of the boundary CFT.

Examples

Let us now explore the correspondence between a massless scalar field in $\text{AdS}_5 \times S^5$ and operators in $\mathcal{N} = 4$ super Yang-Mills (SYM) theory in four dimensions.

The Klein-Gordon equation for the massless field in $\text{AdS}_5 \times S^5$ can be expressed as

$$\nabla^2 \phi = \nabla_{\text{AdS}_5}^2 + \nabla_{S^5}^2 \phi = 0.$$

The eigenfunctions of the Laplacian $\nabla_{S^5}^2$ on a sphere are known as the spherical harmonics $Y_l(\Omega)$, similar to the theory of angular momenta in quantum mechanics. These spherical harmonics satisfy

$$\nabla_{S^5}^2 Y_l(\Omega) = -\frac{l(l+4)}{R^2} Y_l(\Omega), \quad l = 0, 1, 2, \dots. \quad (17.20)$$

Writing the ten-dimensional field $\phi = \sum_l \phi_l Y_l$, the AdS_5 fields ϕ_l satisfy the massive Klein-Gordon equation

$$\nabla_{\text{AdS}_5}^2 \phi_l = m_l^2 \phi_l, \quad m_l^2 = \frac{l(l+4)}{R^2}. \quad (17.21)$$

This indicates that the compactification on S^5 leads to a tower of fields with Kaluza-Klein (KK) masses m_l . According to the AdS/CFT correspondence, there should exist operators in the $\mathcal{N} = 4$ SYM theory in four dimensions that are dual to these fields.

To determine the corresponding operators in four-dimensional $\mathcal{N} = 4$ SYM theory, we can extract the conformal dimension of the operator using the relation

$$\Delta_l = 2 + \sqrt{4 + (m_l R)^2} = 2 + \sqrt{4 + l(l+4)} = 4 + l. \quad (17.22)$$

First, let us consider the case of a massless scalar with $l = 0$ in AdS_5 , where the conformal dimension of the dual operator is $\Delta_0 = 4$. The spherical harmonics (17.20) at $l = 0$ correspond to the s-wave, which transforms trivially under the $\text{SO}(6)$ symmetry. Consequently, the dual operators are singlets under the $\text{SU}(4)$ R -symmetry, and they do not involve the scalars X_i of the $\mathcal{N} = 4$ SYM theory. The only operator with these properties is the gauge-invariant glueball operator:

$$\mathcal{O} = \text{Tr} [F_{\mu\nu} F^{\mu\nu}] .$$

Note that the conformal dimension of this operator is $\Delta = 4$ since $\dim[\partial] = \dim[A] = 1$.

For higher KK modes with $l > 0$, the dual operator will transform non-trivially under the $\text{SU}(4)$ R -symmetry, involving the scalar fields X_i . A natural candidate for the dual operator to the l -th KK mode is given by:

$$\mathcal{O}_{i_1, \dots, i_l} = \text{Tr} [X_{(i_1, \dots, i_l)} F_{\mu\nu} F^{\mu\nu}] , \quad (17.23)$$

where $X_{(i_1, \dots, i_l)}$ represents the traceless symmetric product of l scalar fields X_i in the $\mathcal{N} = 4$ SYM theory. It can be observed that the conformal dimension of this operator is $4 + l$, which is consistent with (17.22). In fact, this field/operator correspondence has been extended to cover all fields of ten-dimensional supergravity on $\text{AdS}_5 \times S^5$.

References

- [ABJM08] O. Aharony, O. Bergman, D. L. Jafferis, and J. Maldacena, *$N=6$ superconformal Chern-Simons-matter theories, $M2$ -branes and their gravity duals*, *JHEP* **10** (2008) 091, [arXiv:0806.1218 \[hep-th\]](#).
- [AGM⁺00] O. Aharony, S. S. Gubser, J. M. Maldacena, H. Ooguri, and Y. Oz, *Large N field theories, string theory and gravity*, *Phys. Rept.* **323** (2000) 183–386, [arXiv:hep-th/9905111](#).
- [AGT10] L. F. Alday, D. Gaiotto, and Y. Tachikawa, *Liouville Correlation Functions from Four-dimensional Gauge Theories*, *Lett. Math. Phys.* **91** (2010) 167–197, [arXiv:0906.3219 \[hep-th\]](#).
- [AGVM24] L. Alvarez-Gaume and M. A. Vazquez-Mozo, *Anomalies and the Green-Schwarz Mechanism*, 2024. [arXiv:2211.06467 \[hep-th\]](#).
- [AGW84] L. Alvarez-Gaume and E. Witten, *Gravitational Anomalies*, *Nucl. Phys.* **B234** (1984) 269.
- [AHK98] O. Aharony, A. Hanany, and B. Kol, *Webs of (p,q) five-branes, five-dimensional field theories and grid diagrams*, *JHEP* **01** (1998) 002, [arXiv:hep-th/9710116](#).

- [Asp96] P. S. Aspinwall, *K3 surfaces and string duality*, Fields, strings and duality. Proceedings, Summer School, Theoretical Advanced Study Institute in Elementary Particle Physics, TASI'96, Boulder, USA, June 2-28, 1996, 1996, pp. 421–540. [arXiv:hep-th/9611137](https://arxiv.org/abs/hep-th/9611137).
- [AST13] O. Aharony, N. Seiberg, and Y. Tachikawa, *Reading between the lines of four-dimensional gauge theories*, *JHEP* **08** (2013) 115, [arXiv:1305.0318 \[hep-th\]](https://arxiv.org/abs/1305.0318).
- [BBS06] K. Becker, M. Becker, and J. H. Schwarz, *String theory and M-theory: A modern introduction*, Cambridge University Press, 2006.
- [BCH73] J. M. Bardeen, B. Carter, and S. W. Hawking, *The Four laws of black hole mechanics*, *Commun. Math. Phys.* **31** (1973) 161–170.
- [Bek72] J. D. Bekenstein, *Black holes and the second law*, *Lett. Nuovo Cim.* **4** (1972) 737–740.
- [Bek73] J. D. Bekenstein, *Black holes and entropy*, *Phys. Rev. D* **7** (1973) 2333–2346.
- [Ber04] N. Berkovits, *Multiloop amplitudes and vanishing theorems using the pure spinor formalism for the superstring*, *JHEP* **09** (2004) 047, [arXiv:hep-th/0406055](https://arxiv.org/abs/hep-th/0406055).
- [Bil08] A. Bilal, *Lectures on Anomalies*, [arXiv:0802.0634 \[hep-th\]](https://arxiv.org/abs/0802.0634).
- [BL07] J. Bagger and N. Lambert, *Modeling Multiple M2's*, *Phys. Rev. D* **75** (2007) 045020, [arXiv:hep-th/0611108](https://arxiv.org/abs/hep-th/0611108).
- [BLT13] R. Blumenhagen, D. Lust, and S. Theisen, *Basic concepts of string theory*, Theoretical and Mathematical Physics, Springer, Heidelberg, Germany, 2013. <http://www.springer.com/physics/theoretical%2C+mathematical+%26+computational+physics/book/978-3-642-29496-9>.
- [BP09] R. Blumenhagen and E. Plauschinn, *Introduction to conformal field theory*, *Lect. Notes Phys.* **779** (2009) 1–256.
- [BPZ84] A. A. Belavin, A. M. Polyakov, and A. B. Zamolodchikov, *Infinite conformal symmetry in two-dimensional quantum field theory*, *Nuclear Physics B* **241** (1984) no. 2 333–380.
- [Car86] J. L. Cardy, *Operator Content of Two-Dimensional Conformally Invariant Theories*, *Nucl. Phys. B* **270** (1986) 186–204.
- [CCCDV12] A. Cappelli, E. Castellani, F. Colomo, and P. Di Vecchia, *The birth of string theory*, Cambridge University Press, 2012.
- [CDLOGP91] P. Candelas, X. C. De La Ossa, P. S. Green, and L. Parkes, *A Pair of Calabi-Yau manifolds as an exactly soluble superconformal theory*, *Nucl. Phys. B* **359** (1991) 21–74. [AMS/IP Stud. Adv. Math.9,31(1998)].
- [Cec23] S. Cecotti, *Introduction to string theory*, Springer Nature, 2023.
- [CJS78] E. Cremmer, B. Julia, and J. Scherk, *Supergravity Theory in Eleven-Dimensions*, *Phys. Lett. B* **76** (1978) 409–412.
- [CM96] C. G. Callan and J. M. Maldacena, *D-brane approach to black hole quantum mechanics*, *Nucl. Phys. B* **472** (1996) 591–610, [arXiv:hep-th/9602043](https://arxiv.org/abs/hep-th/9602043).

- [CT89] C. G. Callan, Jr. and L. Thorlacius, *SIGMA MODELS AND STRING THEORY*, Theoretical Advanced Study Institute in Elementary Particle Physics: Particles, Strings and Supernovae (TASI 88) Providence, Rhode Island, June 5-July 1, 1988, pp. 795–878.
- [Dab95] A. Dabholkar, *Ten-dimensional heterotic string as a soliton*, *Phys. Lett. B* **357** (1995) 307–312, [arXiv:hep-th/9506160](https://arxiv.org/abs/hep-th/9506160).
- [Dab97] ———, *Lectures on orientifolds and duality*, ICTP Summer School in High-Energy Physics and Cosmology, 6 1997. [arXiv:hep-th/9804208](https://arxiv.org/abs/hep-th/9804208).
- [DF02] E. D’Hoker and D. Z. Freedman, *Supersymmetric gauge theories and the AdS/CFT correspondence*, Theoretical Advanced Study Institute in Elementary Particle Physics (TASI 2001): Strings, Branes and EXTRA Dimensions, 1 2002, pp. 3–158. [arXiv:hep-th/0201253](https://arxiv.org/abs/hep-th/0201253).
- [Dij97] R. Dijkgraaf, *Les Houches lectures on fields, strings and duality*, Quantum symmetries. Proceedings, NATO Advanced Study Institute, 64th Session, Les Houches, France, August 1-September 8, 1995, 1997, pp. 3–147. [arXiv:hep-th/9703136](https://arxiv.org/abs/hep-th/9703136). http://inspirehep.net/record/441345/files/arXiv:hep-th_9703136.pdf.
- [Din07] M. Dine, *Supersymmetry and string theory: Beyond the standard model*, Cambridge University Press, 2007.
- [DLP89] J. Dai, R. G. Leigh, and J. Polchinski, *New Connections Between String Theories*, *Mod. Phys. Lett. A* **4** (1989) 2073–2083.
- [DM96a] K. Dasgupta and S. Mukhi, *Orbifolds of M theory*, *Nucl. Phys. B* **465** (1996) 399–412, [arXiv:hep-th/9512196](https://arxiv.org/abs/hep-th/9512196).
- [DM96b] M. R. Douglas and G. W. Moore, *D-branes, quivers, and ALE instantons*, [arXiv:hep-th/9603167](https://arxiv.org/abs/hep-th/9603167).
- [DMW02] J. R. David, G. Mandal, and S. R. Wadia, *Microscopic formulation of black holes in string theory*, *Phys. Rept.* **369** (2002) 549–686, [arXiv:hep-th/0203048](https://arxiv.org/abs/hep-th/0203048).
- [DN12] A. Dabholkar and S. Nampuri, *Quantum black holes*, *Lect. Notes Phys.* **851** (2012) 165–232, [arXiv:1208.4814 \[hep-th\]](https://arxiv.org/abs/1208.4814).
- [DP96] A. Dabholkar and J. Park, *Strings on orientifolds*, *Nucl. Phys. B* **477** (1996) 701–714, [arXiv:hep-th/9604178](https://arxiv.org/abs/hep-th/9604178).
- [Duf95] M. J. Duff, *Strong / weak coupling duality from the dual string*, *Nucl. Phys. B* **442** (1995) 47–63, [arXiv:hep-th/9501030](https://arxiv.org/abs/hep-th/9501030).
- [FL98] S. Forste and J. Louis, *Duality in string theory*, *Nucl. Phys. Proc. Suppl.* **61A** (1998) 3–22, [arXiv:hep-th/9612192](https://arxiv.org/abs/hep-th/9612192).
- [FMS12] P. Francesco, P. Mathieu, and D. Sénéchal, *Conformal field theory*, Springer Science & Business Media, 2012.
- [Fre23] D. S. Freed, *What is an anomaly?*, [arXiv:2307.08147 \[hep-th\]](https://arxiv.org/abs/2307.08147).
- [FVP12] D. Z. Freedman and A. Van Proeyen, *Supergravity*, Cambridge university press, 2012.
- [Gai12] D. Gaiotto, *N=2 dualities*, *JHEP* **08** (2012) 034, [arXiv:0904.2715 \[hep-th\]](https://arxiv.org/abs/0904.2715).

- [GHMR85a] D. J. Gross, J. A. Harvey, E. J. Martinec, and R. Rohm, *Heterotic String Theory. 1. The Free Heterotic String*, *Nucl. Phys.* **B256** (1985) 253.
- [GHMR85b] ———, *The Heterotic String*, *Phys. Rev. Lett.* **54** (1985) 502–505.
- [GHMR86] ———, *Heterotic String Theory. 2. The Interacting Heterotic String*, *Nucl. Phys.* **B267** (1986) 75–124.
- [Gin87] P. H. Ginsparg, *Comment on Toroidal Compactification of Heterotic Superstrings*, *Phys. Rev.* **D35** (1987) 648.
- [Gin88] ———, *APPLIED CONFORMAL FIELD THEORY*, Les Houches Summer School in Theoretical Physics: Fields, Strings, Critical Phenomena Les Houches, France, June 28-August 5, 1988, 1988, pp. 1–168. arXiv:hep-th/9108028. http://inspirehep.net/record/265020/files/arXiv:hep-th_9108028.pdf.
- [GK99] A. Giveon and D. Kutasov, *Brane Dynamics and Gauge Theory*, *Rev. Mod. Phys.* **71** (1999) 983–1084, arXiv:hep-th/9802067.
- [GKP98] S. S. Gubser, I. R. Klebanov, and A. M. Polyakov, *Gauge theory correlators from noncritical string theory*, *Phys. Lett.* **B428** (1998) 105–114, arXiv:hep-th/9802109.
- [GNO77] P. Goddard, J. Nuyts, and D. I. Olive, *Gauge Theories and Magnetic Charge*, *Nucl. Phys. B* **125** (1977) 1–28.
- [GP90] B. R. Greene and M. R. Plesser, *Duality in Calabi-Yau Moduli Space*, *Nucl. Phys. B* **338** (1990) 15–37.
- [Gre99] B. R. Greene, *The elegant universe: Superstrings, hidden dimensions, and the quest of the ultimate theory*, American Association of Physics Teachers, 1999.
- [GS84] M. B. Green and J. H. Schwarz, *Anomaly Cancellation in Supersymmetric D=10 Gauge Theory and Superstring Theory*, *Phys. Lett.* **149B** (1984) 117–122.
- [GS85] ———, *The Hexagon Gauge Anomaly in Type I Superstring Theory*, *Nucl. Phys. B* **255** (1985) 93–114.
- [GSW87] M. B. Green, J. H. Schwarz, and E. Witten, *Superstring Theory Vol. 1,2*, Cambridge University Press, 1987. <http://www.cambridge.org/mw/academic/subjects/physics/theoretical-physics-and-mathematical-physics/superstring-theory-25th-anniversary-edition-volume-1?format=AR>.
- [Har05] J. A. Harvey, *TASI 2003 lectures on anomalies*, 9 2005. arXiv:hep-th/0509097.
- [Haw71] S. W. Hawking, *Gravitational radiation from colliding black holes*, *Phys. Rev. Lett.* **26** (1971) 1344–1346.
- [Haw72] ———, *Black holes in general relativity*, *Commun. Math. Phys.* **25** (1972) 152–166.
- [Haw75] ———, *Particle Creation by Black Holes*, *Commun. Math. Phys.* **43** (1975) 199–220. [,167(1975)].
- [Haw76] ———, *Breakdown of Predictability in Gravitational Collapse*, *Phys. Rev.* **D14** (1976) 2460–2473.

- [HK00] A. Hanany and B. Kol, *On orientifolds, discrete torsion, branes and M theory*, JHEP **06** (2000) 013, arXiv:hep-th/0003025.
- [Hos15] K. Hosomichi, *Lecture notes on string theory*, his webpage (2015) .
- [HR19] J. J. Heckman and T. Rudelius, *Top Down Approach to 6D SCFTs*, J. Phys. A **52** (2019)no. 9 093001, arXiv:1805.06467 [hep-th].
- [HT95] C. M. Hull and P. K. Townsend, *Unity of superstring dualities*, Nucl. Phys. B **438** (1995) 109–137, arXiv:hep-th/9410167.
- [Hul95] C. M. Hull, *String-string duality in ten-dimensions*, Phys. Lett. B **357** (1995) 545–551, arXiv:hep-th/9506194.
- [HW96a] P. Horava and E. Witten, *Eleven-dimensional supergravity on a manifold with boundary*, Nucl. Phys. B **475** (1996) 94–114, arXiv:hep-th/9603142.
- [HW96b] ———, *Heterotic and type I string dynamics from eleven-dimensions*, Nucl. Phys. B **460** (1996) 506–524, arXiv:hep-th/9510209.
- [HW97] A. Hanany and E. Witten, *Type IIB superstrings, BPS monopoles, and three-dimensional gauge dynamics*, Nucl. Phys. B **492** (1997) 152–190, arXiv:hep-th/9611230.
- [HZ99] A. Hanany and A. Zaffaroni, *Issues on orientifolds: On the brane construction of gauge theories with $SO(2n)$ global symmetry*, JHEP **07** (1999) 009, arXiv:hep-th/9903242.
- [IS96] K. A. Intriligator and N. Seiberg, *Mirror symmetry in three-dimensional gauge theories*, Phys. Lett. B **387** (1996) 513–519, arXiv:hep-th/9607207.
- [IU12] L. E. Ibanez and A. M. Uranga, *String theory and particle physics: An introduction to string phenomenology*, Cambridge University Press, 2012. http://www.cambridge.org/de/knowledge/isbn/item6563092/?site_locale=de_DE.
- [Joh00] C. V. Johnson, *D-brane primer*, Theoretical Advanced Study Institute in Elementary Particle Physics (TASI 99): Strings, Branes, and Gravity, 7 2000, pp. 129–350. arXiv:hep-th/0007170.
- [Joh02] C. V. Johnson, *D-branes*, Cambridge university press, 2002.
- [Joh06] ———, *D-branes*, Cambridge university press, 2006.
- [Kir08] A. Kirillov, *An introduction to Lie groups and Lie algebras*, vol. 113, Cambridge University Press, 2008.
- [Kir19] E. Kiritsis, *String theory in a nutshell*, vol. 21, Princeton University Press, 2019.
- [KN69a] Z. Koba and H. B. Nielsen, *Reaction amplitude for n mesons: A Generalization of the Veneziano-Bardakci-Ruegg-Virasora model*, Nucl. Phys. B **10** (1969) 633–655.
- [KN69b] Z. Koba and H. B. Nielsen, *Manifestly crossing invariant parametrization of n meson amplitude*, Nucl. Phys. B **12** (1969) 517–536.
- [Lei89] R. G. Leigh, *Dirac-Born-Infeld Action from Dirichlet Sigma Model*, Mod. Phys. Lett. A **4** (1989) 2767.

- [Mal96] J. M. Maldacena, *Black holes in string theory*, Ph.D. thesis, Princeton U., 1996. [arXiv:hep-th/9607235](https://arxiv.org/abs/hep-th/9607235). <http://wwwlib.umi.com/dissertations/fullcit?p9627605>.
- [Mal99] _____, *The Large N limit of superconformal field theories and supergravity*, *Int. J. Theor. Phys.* **38** (1999) 1113–1133, [arXiv:hep-th/9711200](https://arxiv.org/abs/hep-th/9711200). [Adv. Theor. Math. Phys. 2, 231 (1998)].
- [Mal03] _____, *TASI 2003 lectures on AdS / CFT*, Theoretical Advanced Study Institute in Elementary Particle Physics (TASI 2003): Recent Trends in String Theory, 9 2003, pp. 155–203. [arXiv:hep-th/0309246](https://arxiv.org/abs/hep-th/0309246).
- [MO77] C. Montonen and D. I. Olive, *Magnetic Monopoles as Gauge Particles?*, *Phys. Lett. B* **72** (1977) 117–120.
- [MV96a] D. R. Morrison and C. Vafa, *Compactifications of F theory on Calabi-Yau threefolds. 1*, *Nucl. Phys. B* **473** (1996) 74–92, [arXiv:hep-th/9602114](https://arxiv.org/abs/hep-th/9602114).
- [MV96b] _____, *Compactifications of F theory on Calabi-Yau threefolds. 2.*, *Nucl. Phys. B* **476** (1996) 437–469, [arXiv:hep-th/9603161](https://arxiv.org/abs/hep-th/9603161).
- [Mye99] R. C. Myers, *Dielectric branes*, *JHEP* **12** (1999) 022, [arXiv:hep-th/9910053](https://arxiv.org/abs/hep-th/9910053).
- [Nah78] W. Nahm, *Supersymmetries and their Representations*, *Nucl. Phys. B* **135** (1978) 149.
- [Nar86] K. S. Narain, *New Heterotic String Theories in Uncompactified Dimensions < 10*, *Phys. Lett.* **169B** (1986) 41–46.
- [Năs15] H. Năstase, *Introduction to the ads/cft correspondence*, Cambridge University Press, 2015.
- [Nel87] P. C. Nelson, *Lectures on Strings and Moduli Space*, *Phys. Rept.* **149** (1987) 337.
- [NSW87] K. S. Narain, M. H. Sarmadi, and E. Witten, *A Note on Toroidal Compactification of Heterotic String Theory*, *Nucl. Phys. B* **279** (1987) 369–379.
- [Oog15] H. Ooguri, *Introduction to superstring theory*, Book for the general public in Chinese (2015).
- [OY96] H. Ooguri and Z. Yin, *TASI lectures on perturbative string theories*, Fields, strings and duality. Proceedings, Summer School, Theoretical Advanced Study Institute in Elementary Particle Physics, TASI'96, Boulder, USA, June 2-28, 1996, 1996, pp. 5–81. [arXiv:hep-th/9612254](https://arxiv.org/abs/hep-th/9612254).
- [PC88] J. Polchinski and Y. Cai, *Consistency of Open Superstring Theories*, *Nucl. Phys. B* **296** (1988) 91–128.
- [Pol95] J. Polchinski, *Dirichlet Branes and Ramond-Ramond charges*, *Phys. Rev. Lett.* **75** (1995) 4724–4727, [arXiv:hep-th/9510017](https://arxiv.org/abs/hep-th/9510017).
- [Pol96a] _____, *String duality: A Colloquium*, *Rev. Mod. Phys.* **68** (1996) 1245–1258, [arXiv:hep-th/9607050](https://arxiv.org/abs/hep-th/9607050).
- [Pol96b] _____, *Tasi lectures on D-branes*, Fields, strings and duality. Proceedings, Summer School, Theoretical Advanced Study Institute in Elementary Particle Physics, TASI'96, Boulder, USA, June 2-28, 1996, 1996, pp. 293–356. [arXiv:hep-th/9611050](https://arxiv.org/abs/hep-th/9611050). http://inspirehep.net/record/425724/files/arXiv:hep-th_9611050.pdf.

- [Pol98] J. Polchinski, *String theory. Vol. 1, 2*, Cambridge University Press, 1998.
- [Pol06] J. Polchinski, *Open heterotic strings*, *JHEP* **09** (2006) 082, arXiv:hep-th/0510033.
- [Pol17] _____, *Memories of a Theoretical Physicist*, arXiv:1708.09093 [physics.hist-ph].
- [PW96] J. Polchinski and E. Witten, *Evidence for heterotic - type I string duality*, *Nucl. Phys.* **B460** (1996) 525–540, arXiv:hep-th/9510169.
- [Qua15] J. D. Qualls, *Lectures on Conformal Field Theory*, arXiv:1511.04074 [hep-th].
- [Ram15] A. V. Ramallo, *Introduction to the AdS/CFT correspondence*, Springer Proc. Phys. **161** (2015) 411–474, arXiv:1310.4319 [hep-th].
- [Rom86] L. J. Romans, *Massive N=2a Supergravity in Ten-Dimensions*, *Phys. Lett. B* **169** (1986) 374.
- [RT06] S. Ryu and T. Takayanagi, *Holographic derivation of entanglement entropy from AdS/CFT*, *Phys. Rev. Lett.* **96** (2006) 181602, arXiv:hep-th/0603001.
- [Sch97] J. H. Schwarz, *Lectures on superstring and M theory dualities: Given at ICTP Spring School and at TASI Summer School*, *Nucl. Phys. Proc. Suppl.* **55B** (1997) 1–32, arXiv:hep-th/9607201.
- [Sch00] _____, *Interview with John H. Schwarz*, Caltech Oral Histories (2000). http://oralhistories.library.caltech.edu/116/1/Schwarz_OHO.pdf.
- [Sch11] _____, *The Early History of String Theory and Supersymmetry*, *Pontif. Acad. Sci. Scr. Varia* **119** (2011) 69–81, arXiv:1201.0981 [hep-th].
- [Sen94] A. Sen, *Strong - weak coupling duality in four-dimensional string theory*, *Int. J. Mod. Phys. A* **9** (1994) 3707–3750, arXiv:hep-th/9402002.
- [Sen96] _____, *Orbifolds of M theory and string theory*, *Mod. Phys. Lett. A* **11** (1996) 1339–1348, arXiv:hep-th/9603113.
- [Sen97] _____, *Unification of string dualities*, *Nucl. Phys. Proc. Suppl.* **58** (1997) 5–19, arXiv:hep-th/9609176.
- [Sen98a] _____, *An Introduction to nonperturbative string theory*, Duality and supersymmetric theories. Proceedings, Easter School, Newton Institute, Euroconference, Cambridge, UK, April 7-18, 1997, 1998, pp. 297–413. arXiv:hep-th/9802051.
- [Sen98b] _____, *SO(32) spinors of type I and other solitons on brane - anti-brane pair*, *JHEP* **09** (1998) 023, arXiv:hep-th/9808141.
- [Sha70] J. A. Shapiro, *Electrostatic analog for the virasoro model*, *Phys. Lett. B* **33** (1970) 361–362.
- [SV96] A. Strominger and C. Vafa, *Microscopic origin of the Bekenstein-Hawking entropy*, *Phys. Lett. B* **379** (1996) 99–104, arXiv:hep-th/9601029.
- [tH74] G. ’t Hooft, *A Planar Diagram Theory for Strong Interactions*, *Nucl. Phys.* **B72** (1974) 461.
- [tH80] G. ’t Hooft, *Naturalness, chiral symmetry, and spontaneous chiral symmetry breaking*, *NATO Sci. Ser. B* **59** (1980) 135–157.

- [Tom22] A. Tomasiello, *Geometry of string theory compactifications*, Cambridge University Press, 2022.
- [Ton09] D. Tong, *String Theory*, arXiv:0908.0333 [hep-th].
- [Tow96] P. K. Townsend, *Four lectures on M theory*, High energy physics and cosmology. Proceedings, Summer School, Trieste, Italy, June 10-July 26, 1996, 1996, pp. 385–438. arXiv:hep-th/9612121.
- [Tow97] ———, *Black holes: Lecture notes*, arXiv:gr-qc/9707012.
- [Tse97] A. A. Tseytlin, *On nonAbelian generalization of Born-Infeld action in string theory*, Nucl. Phys. B **501** (1997) 41–52, arXiv:hep-th/9701125.
- [Tse99] ———, *Born-Infeld action, supersymmetry and string theory*, arXiv:hep-th/9908105.
- [TY19] Y. Tachikawa and K. Yonekura, *Why are fractional charges of orientifolds compatible with Dirac quantization?*, SciPost Phys. **7** (2019) no. 5 058, arXiv:1805.02772 [hep-th].
- [Ura05] A. M. Uranga, *Introduction to String Theory*, his webpage (2005) .
- [Vaf96] C. Vafa, *Evidence for F theory*, Nucl. Phys. B **469** (1996) 403–418, arXiv:hep-th/9602022.
- [Vaf97] ———, *Lectures on strings and dualities*, High energy physics and cosmology. Proceedings, Summer School, Trieste, Italy, June 10-July 26, 1996, 1997, pp. 66–119. arXiv:hep-th/9702201.
- [Ven68] G. Veneziano, *Construction of a crossing - symmetric, Regge behaved amplitude for linearly rising trajectories*, Nuovo Cim. **A57** (1968) 190–197.
- [Vir69] M. A. Virasoro, *Alternative constructions of crossing-symmetric amplitudes with regge behavior*, Phys. Rev. **177** (1969) 2309–2311.
- [Wei15] T. Weigand, *Introduction to string theory*, his webpage (2015) .
- [Wei21] G. Weinstein, *Demons in Black Hole Thermodynamics: Bekenstein and Hawking*, arXiv:2102.11209 [physics.hist-ph].
- [Wit94] E. Witten, *Monopoles and four manifolds*, Math. Res. Lett. **1** (1994) 769–796, arXiv:hep-th/9411102.
- [Wit95a] ———, *Some comments on string dynamics*, STRINGS 95: Future Perspectives in String Theory, 7 1995, pp. 501–523. arXiv:hep-th/9507121.
- [Wit95b] ———, *String theory dynamics in various dimensions*, Nucl. Phys. **B443** (1995) 85–126, arXiv:hep-th/9503124.
- [Wit96] ———, *Five-branes and M theory on an orbifold*, Nucl. Phys. **B463** (1996) 383–397, arXiv:hep-th/9512219.
- [Wit97] ———, *Solutions of four-dimensional field theories via M theory*, Nucl. Phys. B **500** (1997) 3–42, arXiv:hep-th/9703166.
- [Wit98a] ———, *Anti-de Sitter space and holography*, Adv. Theor. Math. Phys. **2** (1998) 253–291, arXiv:hep-th/9802150.
- [Wit98b] ———, *Baryons and branes in anti-de Sitter space*, JHEP **07** (1998) 006, arXiv:hep-th/9805112.
- [Wit98c] ———, *D-branes and K theory*, JHEP **12** (1998) 019, arXiv:hep-th/9810188.

- [Wit18] _____, *What every physicist should know about string theory*, Foundations of mathematics and physics one century after Hilbert, Springer, 2018, pp. 197–210. <https://physicstoday.scitation.org/doi/10.1063/PT.3.2980>.
- [Wra11] K. Wray, *An introduction to string theory*, his webpage (2011) .
- [Zwi06] B. Zwiebach, *A first course in string theory*, Cambridge University Press, 2006. <http://www.cambridge.org/uk/catalogue/catalogue.asp?isbn=0521831431>.

Satoshi Nawata, Department of Physics and Center for Field Theory and Particle Physics, Fudan University, Songhu Road 2005, 200438 Shanghai, China, snawata@gmail.com

Runkai Tao, Department of Physics and Astronomy, Rutgers University, 126 Frelinghuysen Road, Piscataway NJ 08854, USA, runkaitao@gmail.com

Daisuke Yokoyama, Department of Physics, Meiji University, Kanagawa 214-8571, Japan, ddyokoyama@meiji.ac.jp