Lectures Notes for MSc in Mathematical and Theoretical Physics
University of Oxford, Hilary Term 2023

String Theory I

Dr Ling Lin[1]
University of Oxford
Mathematical Institute

These lecture notes are designed for a 16-hours introductory course on String Theory. They cover Old Covariant Quantization of the bosonic string, basics of string amplitudes, spacetime background fields and effective actions, and compactifications. The content is largely built on previous courses by Prof. Beem and Prof. de la Ossa. For a more comprehensive introduction to the subject, here are some lecture notes with a similar approach:

- D. Tong, "String Theory" (Cambridge, Part III Maths, `https://www.damtp.cam.ac.uk/user/tong/string.html`),
- T. Weigand, "Introduction to String Theory" (Heidelberg University, `https://www.thphys.uni-heidelberg.de/courses/weigand/Strings11-12.pdf`),
- A. Uranga, "Graduate Course in String Theory" (Universidad Autonoma de Madrid, `https://members.ift.uam-csic.es/auranga/firstpage.html`).

Further reading and references:

- R. Blumenhagen, D. Lüst, S. Theisen, "Basic Concepts of String Theory",
- M. Green, J. Schwarz, E. Witten, "Superstring Theory" (Vol. I),
- J. Polchinski, "String Theory" (Vol. I),
- B. Zwiebach, "A First Course in String Theory",

[1] Please send comments and corrections to `ling.lin@maths.ox.ac.uk`.

Contents

1 Motivation and Overview

To be added.

Some notational conventions:

- throughout the notes, we set $c = \hbar = 1$; this means that [length] = [time] = $[\text{mass}]^{-1} = [\text{energy}]^{-1}$;
- signature convention for Lorentzian spacetime is $(-,+,+,\dots)$; the Minkowski metric tensor is then $\eta_{ij} = \text{diag}(-1,1,\dots)$;
- we use lower-case Greek indices $(\mu,\nu,\dots)$ for spacetime, and Latin indices $(a,b,\dots)$ for the string worldsheet;
- we write $G_{ij}V^iV^j \equiv V\cdot V$ for index contractions with a metric; in case of multiple metrics, the correct one can be inferred from the text;
- partial derivatives $\frac{\partial}{\partial x} \equiv \partial_x$ are denoted with subscripts; if the variable has an index, we write $\frac{\partial}{\partial x^i} \equiv \partial_i$;
- ...

2 Classical Relativistic String

In this chapter, we will study the relativistic propagation of a classical string in a fixed background spacetime $\mathcal{M}$. We will make use of some differential geometry facts that should have been covered in the General Relativity I lecture course in the previous term.

2.1 Classical relativistic particle

To motivate the formalism that describes a string, we first revisit the description for a point-like particle. Classically, a particle of mass m travels along a *worldline* λ in spacetime $\mathcal{M}$ with metric $g_{\mu\nu}$. Denoting the (local) spacetime coordinates by X^μ, $\mu = 0,\dots,D-1$, we can introduce the parametrization, i.e., an *embedding* of the worldline into spacetime,

$$\lambda : \mathbb{R} \to \mathcal{M}, \quad \tau \mapsto X^\mu(\tau)\,. \tag{2.1}$$

Recall from elementary differential geometry (e.g., from GR I), that we can use this embedding to pullback the spacetime worldvolume, defined by the metric $g_{\mu\nu}(X) \equiv g_{\mu\nu}$, to the line element, or volume form ds on the worldline,

$$ds = \sqrt{-g_{\mu\nu} dX^\mu dX^\nu} = \sqrt{-g_{\mu\nu} \left(\frac{dX^\mu}{d\tau} d\tau\right)\left(\frac{dX^\nu}{d\tau} d\tau\right)} = \sqrt{-\frac{dX}{d\tau} \cdot \frac{dX}{d\tau}} d\tau \,. \tag{2.2}$$

In less technical terms, ds is the infinitesimal proper length of the worldline.

In the absence of any external forces, the equations of motion should fix the physical trajectory of the particle to be the one with minimal (extremal) proper length, $\int_\lambda ds$, i.e., the particle moves along a geodesic. In other words, the action (which has to be a dimensionless quantity) for a point particle is

$$S = -m \int_\lambda ds = -m \int d\tau \sqrt{-\frac{dX}{d\tau} \cdot \frac{dX}{d\tau}} \,. \tag{2.3}$$

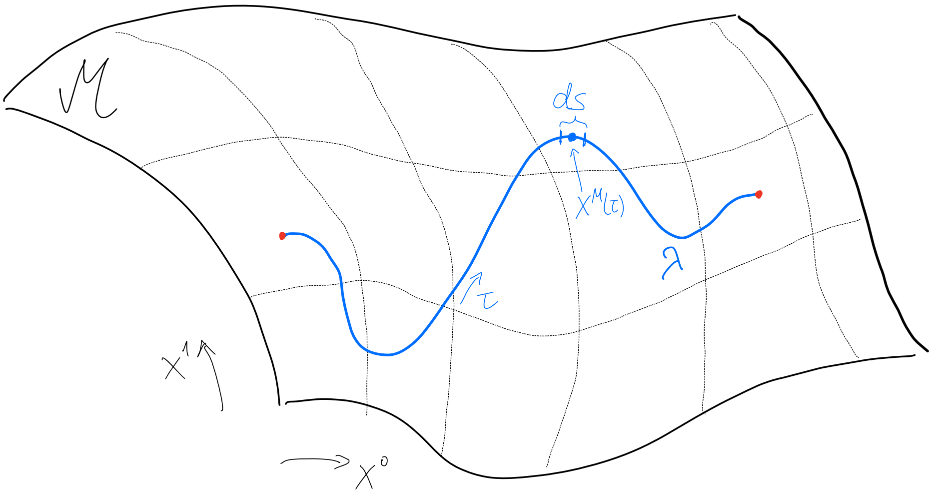


Fig. 1: Point-particle moving in spacetime $\mathcal{M}$ traces out a worldline λ parametrized by τ. The infinitesimal length at any point with spacetime coordinate $X^\mu(\tau)$ is given by (2.2).

Since the indices μ, ν run from 0 to $D-1$, it is tempting to conclude that the system D degrees of freedom. On the other hand, the motion of a particle should be entirely described by its $(D-1)$ spatial coordinates (plus initial conditions). The reason for this discrepancy is an important property of the action (2.3): it is invariant under reparametrization $\tau \longrightarrow \tilde{\tau}(\tau)$. In physics language, this corresponds to a *gauge symmetry*, which captures a redundancy of the description that we chose for the system. For example, we could get rid of this redundancy by using the spacetime time coordinate $X^0 \equiv t$ to parametrize the worldline, for which the action reduces to the more familiar form

$$S = -m \int dt \sqrt{1 - \dot{\vec{X}} \cdot \dot{\vec{X}}}\,, \tag{2.4}$$

with $\dot{\vec{X}} = \frac{d}{dt}\vec{X} = \frac{d}{dt}(X^1, X^2, \ldots, X^{D-1})$.

In theoretical physics, gauge redundancies are oftentimes a good thing to have. For example, they are essential to describe the nuclear and electromagnetic forces within the Standard Model of particle physics. In the present case, having the redundant parameter τ makes manifest the *global* symmetries of the system, namely, the spacetime isometries (i.e., transformations on X^μ which leave invariant the line element (2.2)). If $\mathcal{M}$ is flat, i.e., $g_{\mu\nu} = \eta_{\mu\nu} = \mathrm{diag}[-1, 1, \ldots, 1]$, then (2.3) would be invariant under the Poincaré symmetry on $\mathbb{R}^{1,D-1}$.

One caveat of describing point particles using the action above is that it is strictly speaking not defined for $m = 0$. Instead, one needs to use the action

$$S' = \frac{1}{2} \int d\tau\, e(\tau) \left(\frac{1}{e^2(\tau)}\, g_{\mu\nu} \frac{dX^\mu}{d\tau} \frac{dX^\nu}{d\tau} - m^2 \right), \tag{2.5}$$

where we have introduced a further, seemingly independent, field $e(\tau)$ on the worldline. To maintain invariance under any reparametrization $\tau \longrightarrow \tilde{\tau}(\tau)$, or *local diffeomorphisms* on the worldline, the fields must transform as

$$X^\mu(\tau) \rightarrow \tilde{X}^\mu(\tilde{\tau}(\tau)) = X^\mu(\tau)\,, \quad e(\tau) \rightarrow \tilde{e}(\tilde{\tau}) = \frac{d\tau}{d\tilde{\tau}} e(\tau)\,. \tag{2.6}$$

The equation of motion for e is

$$0 = \frac{\delta S'}{\delta e(\tau)} \quad \Longleftrightarrow \quad \frac{1}{e^2(\tau)} \frac{dX}{d\tau} \cdot \frac{dX}{d\tau} + m^2 = 0\,, \tag{2.7}$$

For $m \neq 0$, this fixes e entirely in terms of X^μ, and, once plugged back into (2.5), shows the equivalence to the first action (2.3).

More generally, we can use the reparametrization invariance to *gauge fix* e. A convenient choice is to set it to constant, e.g.,

$$e(\tau) = \begin{cases} \frac{1}{m}\,, & m \neq 0 \\ 1\,, & m = 0\,. \end{cases} \tag{2.8}$$

Then, the variation of the action (2.5) with respect to X^μ just leads to the same Euler–Lagrange equation, up to the constant factor $1/e$, as those derived from the action Lagrangian $g_{\mu\nu}(dX^\mu/d\tau)(dX^\nu/d\tau)$. From General Relativity I, we know that this is nothing but the geodesic equation,

$$\frac{d^2X^\mu}{d\tau^2} + \Gamma^\mu_{\rho\sigma}\frac{dX^\rho}{d\tau}\frac{dX^\sigma}{d\tau} = 0\,, \tag{2.9}$$

with $\Gamma^\mu_{\rho\sigma}$ the Christoffel symbols associated to the spacetime metric $g_{\mu\nu}$. Moreover, the equation of motion (2.7) becomes the following constraint:

$$\frac{dX}{d\tau}\cdot\frac{dX}{d\tau} = \begin{cases} -1\,, & m \neq 0\,, \\ 0\,, & m = 0\,. \end{cases} \tag{2.10}$$

These tell us that the parametrization we chose above to gauge fix e corresponds to τ being the proper time along the worldline. In fact, for any parametrization, the field $e(\tau)$ is nothing but the worldline metric, $g_{\tau\tau} = e^2(\tau)$, in the parametrization coordinate τ. The equations of motion for e simply states that this metric is not any arbitrary one on a one-dimensional manifold, but rather the pullback of the spacetime metric onto the worldline of the particle. We will revisit this interpretation when we study the propagation of strings.

Last, but not least, there is a more general lesson here: Whenever we have a redundant description (or, gauge symmetry), we may eliminate the superfluous degrees of freedom by some choice of gauge fixing; however, the equations of motions for these degrees of freedom remain important, as they become certain constraints for the gauge-fixed system. In particular, such constraints will play a crucial role if one wishes to further quantize the system.

2.2 Action principles for a string

Since the string is a one-dimensional object, it sweeps out a two-dimensional *worldsheet* Σ, with coordinates $\xi^a \equiv (\xi^0, \xi^1) = (\tau, \sigma)$, as it moves through spacetime. As the nomenclature suggest, we think of σ as the spatial coordinate along the string, while τ is a time coordinate; of course, this is just a parametrization of a geometric object, so the physics should eventually be invariant under any other choice of parameters. It is convenient to pick σ to take values in an interval $[0, l]$ where l can be thought of as the length of the string *as measured in some arbitrary* *units on* Σ (the physical string length is a fixed parameter, see below). In the case of closed strings, σ can be taken to be periodic, $\sigma \equiv \sigma + l$.

The configuration of the string in spacetime is then described by the embedding $X^\mu(\tau, \sigma)$, which specifies the spacetime coordinates X^μ of any given point $\xi^a = (\tau, \sigma)$ on the worldsheet. Unlike a particle, the spatial extension of the string allows for two different topological configurations: an open or a closed string. The latter is characterized by the condition $X^\mu(\tau, \sigma) = X^\mu(\tau, \sigma + l)$. For open strings, there will be boundary conditions associated to the end points, which we will discuss later.

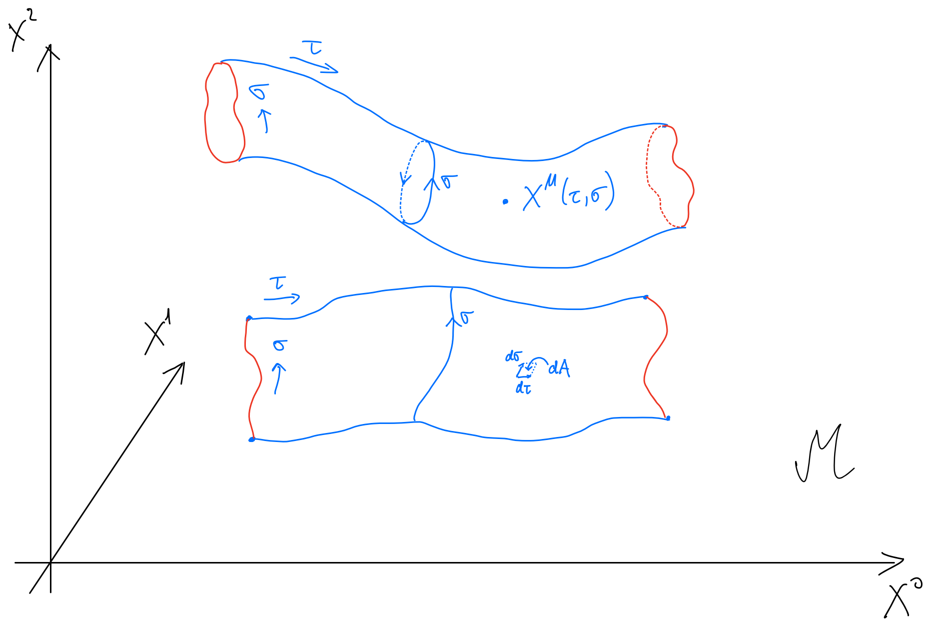


Fig. 2: Strings (closed or open) trace out a worldsheet Σ as they move through spacetime. In the local coordinates $\xi^a = (\tau, \sigma)$, the infinitesimal area element dA is proportional to $d\tau d\sigma$, with the proportionality factor given by $\sqrt{-\det(h)}$, with h the induced metric on Σ.

2.2.1 The Nambu–Goto action

For the point particle, we have seen that its dynamics is captured by an action which is the proper length of its worldline. In (2.3), this was computed as a worldline integral of the line element, or one-dimensional volume form, that is the pullback of the spacetime volume form by the embedding $X^\mu(\tau)$. The natural generalization to a string would then be an action that measures the *area* of the worldsheet Σ. The corresponding "area element", or two-dimensional volume form on Σ is

$$dA = \sqrt{-(\partial_\tau X\cdot\partial_\tau X)(\partial_\sigma X\cdot\partial_\sigma X)+(\partial_\tau X\cdot\partial_\sigma X)^2}\,d\tau d\sigma \equiv \sqrt{-\det(h_{ab})}\,d\tau d\sigma\,, \tag{2.11}$$

where

$$h_{ab} = \frac{\partial X^\mu}{\partial\xi^a}\frac{\partial X^\nu}{\partial\xi^b}g_{\mu\nu}(X(\xi)) \tag{2.12}$$

is the *induced* worldsheet metric, i.e., the pullback of the spacetime metric onto Σ by the embedding $X^\mu(\xi)$. Note that in a general curved spacetime, the metric $g_{\mu\nu}$ are non-trivial functions in the worldsheet coordinates ξ^a through the explicit dependence on the spacetime coordinates $X^\mu(\xi)$.

The natural action for a relativistic string, called the **Nambu–Goto action**, is therefore

$$S_{\rm NG}[X^\mu(\xi)] = -T\int_\Sigma dA = -T\int\sqrt{-\det(h_{ab})}\,d\tau d\sigma\,. \tag{2.13}$$

By dimensional analysis, the coefficient T must have (spacetime) mass dimension 2 since X^μ has mass dimension -1 (the *worldsheet* length dimensions cancel out between the derivatives in h_{ab} and the differentials $d\tau d\sigma$), hence it can be thought of as the *string tension* (energy/mass per unit length). Related quantities are:

$$T = \frac{1}{2\pi\alpha'}\,,\qquad \alpha' :\ \text{Regge-slope}, \tag{2.14}$$

$$\ell_s = 2\pi\sqrt{\alpha'}\,,\qquad \ell_s :\ \text{string length}, \tag{2.15}$$

$$M_s = \tfrac{1}{\sqrt{\alpha'}}\,,\qquad M_s :\ \text{string (mass) scale}. \tag{2.16}$$

The Nambu–Goto action can be viewed as describing a two-dimensional field theory on Σ with degrees of freedom given by $X^\mu(\xi^a)$. Note that X^μ is a *scalar field* on Σ. This action manifestly has the spacetime isometries, e.g., for $g_{\mu\nu}\equiv\eta_{\mu\nu} := \mathrm{diag}(-1,1,\ldots,1)$, the Poincaré transformations,

$$X^\mu(\xi)\to\Lambda^\mu{}_\nu X^\nu(\xi)+b^\mu\,,\quad \Lambda^\mu{}_\nu\in SO(1,D-1)\,,\quad b^\mu\in\mathbb{R}^{1,D-1}\,, \tag{2.17}$$

as an internal, or global symmetry. Moreover, it is also invariant under reparametrization, or local diffeomorphisms, which as a redundancy of the description constitutes a gauge symmetry.

By construction, the solutions to the Euler–Lagrange equations will extremize the worldsheet area. Using the variation $\delta\sqrt{-\det(h)} = \frac{1}{2}\sqrt{-\det(h)}h^{ab}\delta h_{ab}$, one can derive the equations of motion

$$\partial_a\left(\sqrt{-\det(h)}h^{ab}g_{\mu\nu}(X(\xi))\,\partial_b X^\nu\right) = 0\,, \tag{2.18}$$

and study the classical dynamics of a string (see Problem Sheet 1). However, the non-linear nature of these equations (originating from the square root), even in a flat spacetime $g_{\mu\nu} \equiv \eta_{\mu\nu}$, make the quantization procedure more challenging. As the point-particle discussion foreshadows, it turns out that there is another, classically equivalent starting point, that will be easier to quantize.

2.2.2 The Polyakov action

Just as for the point particle, the simplification comes from adding a superfluous field to the system that describes the intrinsic geometry, i.e., the metric γ_{ab} of the worldsheet:

$$S_{\mathrm{P}}[X^\mu(\xi),\gamma_{ab}(\xi)] = -\frac{T}{2}\int_\Sigma d\tau d\sigma\sqrt{-\det(\gamma)}\,\gamma^{ab}\partial_a X^\mu(\xi)\partial_b X^\nu(\xi)g_{\mu\nu}\,. \tag{2.19}$$

This is the **Polyakov action**. Variations of this action with respect to the spacetime coordinates X^μ gives the equations of motion

$$\partial_a\left(\sqrt{-\det(\gamma)}\gamma^{ab}g_{\mu\nu}(X)\partial_b X^\nu\right) = 0\,, \tag{2.20}$$

which structurally is identical to (2.18) that were derived from the Nambu–Goto action. Whereas the worldsheet metric h_{ab} in (2.18) is itself an expression containing X^μ, cf. (2.12), the crucial difference in (2.20) is that γ_{ab} is a priori an independent field. Hence, in a flat spacetime, $g_{\mu\nu} \equiv \eta_{\mu\nu}$, this equation is *linear* in X^μ.

The worldsheet metric comes with its own equation of motion. Recalling from GR I that the variation

$$T_{ab} := -\frac{2}{T}\frac{1}{\sqrt{-\det\gamma}}\frac{\delta S_{\mathrm{P}}}{\delta\gamma^{ab}} \tag{2.21}$$

with respect to the metric is called the *energy-momentum* or *stress tensor*, the equation of motion for γ_{ab} are then (could be useful exercise for GR exam to verify!)

$$\begin{aligned} T_{ab} &= \underbrace{\partial_a X^\mu \partial_b X^\nu g_{\mu\nu}}_{\equiv h_{ab}} - \frac{1}{2}\gamma_{ab} \left(\gamma^{bc}\partial_b X^\mu \partial_c X^\nu g_{\mu\nu}\right) = 0\,, \\ \Longrightarrow \quad h_{ab} &= \frac{1}{2}(\gamma^{bc} h_{bc})\,\gamma_{ab}\,. \end{aligned} \tag{2.22}$$

This means that the dynamical worldsheet metric γ_{ab} is *proportional* to the pullback metric $\partial_a X^\mu \partial_b X^\nu g_{\mu\nu} = h_{ab}$. The proportionality factor is itself a function on the worldsheet, which, however, drops out in both the action (2.19) and the equation of motion (2.20) for X^μ if we plug in the solution for γ_{ab}. This shows that, on-shell, the Nambu–Goto and the Polyakov action describe the same dynamics.

Symmetries of the Polyakov action The non-physicality of the prefactor $\gamma_{bc}h^{bc}$ reflects the fact that the 2d field theory on the worldsheet Σ, described by the Polyakov action, has an additional symmetry. Namely, in addition to the global symmetries in form of spacetime isometries (which does not affect γ), the reparametrization (gauge) symmetry $\xi^a \to \tilde{\xi}^a(\xi)$, which acts the worldsheet fields as

$$\text{local diffeomorphisms:} \qquad \begin{cases} X^\mu(\xi) & \to \tilde{X}^\mu(\tilde{\xi}(\xi)) = X^\mu(\xi)\,, \\ \gamma_{ab}(\xi) & \to \tilde{\gamma}_{ab}(\tilde{\xi}(\xi)) = \gamma_{cd}(\xi)\frac{\partial \xi^c}{\partial \tilde{\xi}_a}\frac{\partial \xi^d}{\partial \tilde{\xi}_b}\,, \end{cases} \tag{2.23}$$

we also have *Weyl invariance*, or local scale symmetry, which acts only on the worldsheet metric:

$$\text{Weyl transformation:} \qquad \gamma_{ab}(\xi) \to \tilde{\gamma}_{ab}(\xi) = e^{2\omega(\xi)}\gamma_{ab}(\xi)\,, \tag{2.24}$$

for some function $\omega(\xi)$ on Σ (the factor 2 is introduced for later convenience).

Notice that Weyl invariance is special to a string: for a k-dimensional object with $(k+1)$-dimensional worldvolume, the analogous version of the Polyakov action (2.19) would have $\sqrt{-\det\gamma}\gamma^{ab} \to (e^{(k+1)\omega}\sqrt{-\det\gamma})(e^{-2\omega}\gamma^{ab})$, which is not invariant unless $k = 1$. As we will learn later, Weyl symmetry is an essential ingredient for the quantization procedure. The lack of Weyl symmetry for higher dimensional membranes is in part the reason why our understanding of non-perturbative objects in string and M-theory is still incomplete.

A general consequence, also in higher dimensions, of Weyl invariance is the *tracelessness* of the energy-momentum tensor. To see this, recall that, by definition, the variation of the action with respect to variations of the metric is

$$\delta S = \frac{\delta S}{\delta\gamma_{ab}}\delta\gamma_{ab} \propto \sqrt{-\det\gamma}\, T^{ab}\delta\gamma_{ab}. \tag{2.25}$$

For an infinitesimal Weyl transformation (i.e., infinitesimal $\omega(\xi)$ in (2.24)), we have $\delta\gamma_{ab} = 2\omega\gamma_{ab}$. Weyl invariance then requires

$$0 = \delta S \propto 2\sqrt{-\det\gamma}\,\omega\, T^{ab}\gamma_{ab} \tag{2.26}$$

for *any* ω, implying that $T^{ab}\gamma_{ab} = T^a{}_a = 0$. Note that this does not require any fields to satisfy their equations of motion!

Furthermore, the local diffeomorphism invariance on the worldsheet can be shown to enforce

$$\nabla_a T^{ab} = 0\,, \tag{2.27}$$

where ∇^a is the Levi-Civita connection associated to the worldsheet metric γ_{ab}. Importantly, this holds only *on-shell* for X, i.e., when X^μ satisfy their equations of motion. Intuitively, this is because T_{ab} only measures the response to variations in h_{ab} while $\delta X^\mu = 0$.

As a final comment to this section, one can regard the Polyakov action (2.19) as “matter fields” X^μ coupled to the worldsheet metric, i.e., 2d gravity. What about other terms, that is, other types of interactions, in a typical (classical) theory of gravity? It turns out that most other interactions are forbidden if we wish to preserve Weyl invariance. E.g., a cosmological constant term, $\Lambda\int\sqrt{-\det\gamma}d\tau d\sigma$(which is the analog of the mass term in (2.5) for a particle), or any scalar potential term, $\int\sqrt{-\det\gamma}V(X)d\tau d\sigma$, explicitly break Weyl invariance. One may then wonder what happens with the Einstein–Hilbert term, $\frac{1}{2\pi}\int\sqrt{-\det\gamma}\mathcal{R}(\gamma)d\tau d\sigma$, with $\mathcal{R}(\gamma)$ the Ricci scalar of γ. For a two-dimensional manifold, this integral is actually a topological invariant, called the *Euler characteristic*,[2] so it does not have any affect on the (local) dynamics. This coins the phrase “2d gravity is trivial”, see Problem Sheet 1.

[2] For the worldsheet of open strings, i.e., 2d manifolds with boundaries, one needs to further add a boundary term, see explanation on Problem Sheet 1.

2.3 Classical solutions for the Polyakov string

Next, we will discuss in more detail the string solutions that solve the equations of motions derived from the Polyakov action, for which the large gauge symmetry (which were redundancies!) will be extremely important. For most of the lectures, we will restrict our attention to a flat spacetime background, and we will make use of the notation $g_{\mu\nu}V^\mu V^\nu = \eta_{\mu\nu}V^\mu V^\nu = V \cdot V$ more frequently.

2.3.1 Gauge fixing the Polyakov String

Since the worldsheet has diffeomorphisms and Weyl transformations as gauge symmetries, we can eliminate redundant degrees of freedom by gauge fixing. By first choosing an appropriate reparametrization (i.e., coordinate change), we can bring any metric into *conformal gauge*, i.e.,

$$\gamma_{ab} \to e^{2\omega(\xi)}\eta_{ab} = e^{2\omega(\xi)} \left(\begin{smallmatrix} -1 & 0 \\ 0 & 1 \end{smallmatrix} \right) . \tag{2.28}$$

For an ad-hoc explanation, notice that, since we have two worldsheet coordinate which we can reparametrize individually, we expect to be able to also fix two out of the three independent metric entries. The final degree of freedom can be further fixed, using Weyl rescaling; an obvious gauge choice is then *unit gauge*,

$$e^{2\omega(\xi)}\eta_{ab} \to \eta_{ab} . \tag{2.29}$$

Already in conformal gauge, the Polyakov action simplifies dramatically,

$$S_{\mathrm{P}}^{\text{c.g.}}[X^\mu] = -\frac{T}{2}\int d\tau d\sigma\, \partial_a X \cdot \partial^a X , \tag{2.30}$$

which is a theory of D massless scalar fields X^μ in a $(1+1)$-dimensional Minkowski space Σ. Their equations of motion (2.20) become, if the spacetime metric is flat, just free wave equations on Σ,

$$\partial_a(g_{\mu\nu}\partial^a X^\nu) = \partial_a \partial^a X_\mu = 0 . \tag{2.31}$$

As we learned from the point particle, the metric's equations of motion become constraints after gauge fixing. In conformal gauging, these constraints are

$$T_{\tau\tau} = T_{\sigma\sigma} = \frac{1}{2}(\partial_\tau X \cdot \partial_\tau X + \partial_\sigma X \cdot \partial_\sigma X) \overset{!}{=} 0 , \quad T_{\tau\sigma} = \partial_\tau X \cdot \partial_\sigma X \overset{!}{=} 0 . \tag{2.32}$$

These restricts the solutions to the wave equations. Notice the tracelessness condition, $T^a{}_a = -T_{\tau\tau} + T_{\sigma\sigma} = 0$, holds *irrespective* of these constraints, and reduce the number of inequivalent constraints to two.

2.3.2 Oscillator mode expansions

To solve the equations of motions (2.31) that describe the string's motion in spacetime, it will be useful to perform a coordinate change into the so-called *lightcone coordinates* $\tilde{\xi}^a(\xi) \equiv (\xi^+(\tau,\sigma), \xi^-(\tau,\sigma))$ on the worldsheet Σ:

$$\xi^+ := \tau + \sigma\,, \quad \xi^- := \tau - \sigma\,; \quad \partial_\pm := \frac{\partial}{\partial\xi^\pm} = \tfrac{1}{2}(\partial_\tau \pm \partial_\sigma)\,. \tag{2.33}$$

In these, the gauge fixed worldsheet metric $\gamma_{ab} = \eta_{ab}$ becomes

$$\gamma_{++} = \gamma_{--} = \gamma^{++} = \gamma^{--} = 0\,,\ \gamma_{+-} = \gamma_{-+} = -\frac{1}{2}\,,\ \gamma^{+-} = \gamma^{-+} = -2\,, \tag{2.34}$$

for which the volume form (or integration measure) on Σ reads

$$d\tau d\sigma = d\xi^+ d\xi^- \ \det\frac{\partial(\tau,\sigma)}{\partial(\xi^+,\xi^-)} = \frac{1}{2} d\xi^+ d\xi^-\,. \tag{2.35}$$

The equations of motion (2.31), which (recall we have fixed the spacetime metric to be flat) in lightcone coordinates read

$$\partial_+\partial_- X^\mu = 0\,. \tag{2.36}$$

Written in this way, it is easy to spot the general solutions:

$$X^\mu(\xi^+,\xi^-) = X^\mu_L(\xi^+) + X^\mu_R(\xi^-)\,, \tag{2.37}$$

where the summands describe waves moving in *opposite directions* along the string. Conventionally, the part that depends on ξ^+ (ξ^-) is called left-(right-)moving.

We have not yet implemented to constraints. To do so, it is useful to expand the waves into their Fourier modes. Here, the topology of the string makes a difference.

Closed string expansion For a closed string, recall that we had the periodic boundary condition

$$X^\mu(\tau, \sigma = 0) = X^\mu(\tau, \sigma = l)\,. \tag{2.38}$$

The Fourier expansion for such a function into left- and right-moving parts is

$$\begin{aligned} X_L^\mu(\xi^+) &= \frac{1}{2}x^\mu + \frac{\pi\alpha'}{l}p^\mu \xi^+ + i\sqrt{\frac{\alpha'}{2}} \sum_{n\in\mathbb{Z},\ n\neq 0} \frac{1}{n}\tilde{\alpha}_n^\mu \exp(-\tfrac{2\pi i}{l} n\, \xi^+)\,, \\ X_R^\mu(\xi^-) &= \frac{1}{2}x^\mu + \frac{\pi\alpha'}{l}p^\mu \xi^- + i\sqrt{\frac{\alpha'}{2}} \sum_{n\in\mathbb{Z},\ n\neq 0} \frac{1}{n}\alpha_n^\mu \exp(-\tfrac{2\pi i}{l} n\, \xi^-)\,, \end{aligned} \tag{2.39}$$

where

- the coefficients $\alpha_n^\mu, \tilde{\alpha}_n^\mu$ (not to be confused to be in any relation with the string tension α'!) are independent modes that, by convention, correspond to positive frequency modes for $n < 0$;
- the periodic boundary condition is satisfied by having the same coefficient for the left-/right-moving zero mode,

$$\tilde{\alpha}_0^\mu = \alpha_0^\mu = \sqrt{\frac{\alpha'}{2}} p^\mu\,, \tag{2.40}$$

- X^μ being a real-valued field, $X^\mu = (X^\mu)^*$, implies

$$x^\mu = (x^\mu)^*\,, \quad p^\mu = (p^\mu)^*\,, \quad (\alpha_n^\mu)^* = \alpha_{-n}^\mu\,, \quad (\tilde{\alpha}_n^\mu)^* = \tilde{\alpha}_{-n}^\mu\,. \tag{2.41}$$

For later convenience, it is helpful to display the derivatives of the coordinate fields:

$$\begin{aligned} \partial_+ X^\mu &= \partial_\tau X_L^\mu = \frac{2\pi}{l}\sqrt{\frac{\alpha'}{2}} \sum_{n\in\mathbb{Z}} \tilde{\alpha}_n^\mu \exp(-\tfrac{2\pi i}{l} n(\tau+\sigma))\,, \\ \partial_- X^\mu &= \partial_\tau X_R^\mu = \frac{2\pi}{l}\sqrt{\frac{\alpha'}{2}} \sum_{n\in\mathbb{Z}} \alpha_n^\mu \exp(-\tfrac{2\pi i}{l} n(\tau-\sigma))\,. \end{aligned} \tag{2.42}$$

Open string expansion While the equation of motion (2.36) remains the same for open strings, there is an additional condition on the motion of the endpoints

$X^\mu(\tau, \sigma = 0)$ and $X^\mu(\tau, \sigma = l)$. By doing the variation of the gauge fixed action (2.30) including the boundary contributions more carefully, we find the condition

$$\partial_\sigma X^\mu \delta X_\mu = 0 \quad \text{at} \quad \sigma = 0, l\,, \tag{2.43}$$

which of course should be familiar from a classical mechanics treatment of vibrating strings: The two possibilities for each endpoint are Neumann (N) boundary conditions, $\partial_\sigma X^\mu|_{\sigma=\text{bdry}} = 0$, or Dirichlet (D) boundary conditions, $\delta X^\mu|_{\sigma=\text{bdry}} = 0$. Physically, Neumann endpoints can move freely, provided there is no momentum flowing off the string; Dirichlet endpoints are fixed and cannot move.

Combining these for the two endpoints, an open string has the following possible mode expansions:

1. Neumann boundary conditions at both ends (NN):

$$X^\mu(\tau,\sigma) = x^\mu + \frac{2\pi\alpha'}{l} p^\mu \tau + i\sqrt{2\alpha'} \sum_{n\in\mathbb{Z},\, n\neq 0} \tfrac{1}{n}\alpha_n^\mu \exp(-\tfrac{i\pi n}{l}\tau)\, \cos(\tfrac{n\pi\sigma}{l})\,, \tag{2.44}$$

where the left- and right-moving parts are no longer independent, $\tilde{\alpha}_n = \alpha_n$, due to the boundary conditions. Again, reality imposes $(\alpha_n^\mu)^* = \alpha_{-n}^\mu$. Defining $\alpha_0^\mu := \sqrt{2\alpha'} p^\mu$, we find

$$\partial_\pm X^\mu = \frac{1}{2}(\partial_\tau X^\mu \pm \partial_\sigma X^\mu) = \frac{\pi}{l}\sqrt{\frac{\alpha'}{2}} \sum_{n\in\mathbb{Z}} \alpha_n^\mu \exp(-\tfrac{i\pi n}{l}(\tau \pm \sigma))\,. \tag{2.45}$$

2. Dirichlet boundary conditions at both ends (DD): fixing endpoints means $\delta X^\mu(\sigma = 0, l) = \partial_\tau X^\mu(\sigma = 0, l)\delta\tau \overset{!}{=} 0$, so $\partial_\tau X^\mu|_{\sigma=\text{bdry}} = 0$. Defining $x_0^\mu := X^\mu(\tau, \sigma = 0)$, $x_1^\mu := X^\mu(\tau, \sigma = l)$, and $\alpha_0^\mu = \frac{1}{\sqrt{2\alpha'}\pi}(x_1^\mu - x_0^\mu)$, we have

$$\begin{aligned} X^\mu(\tau,\sigma) &= x_0^\mu + \frac{x_1^\mu - x_0^\mu}{l}\sigma + \sqrt{2\alpha'} \sum_{n\in\mathbb{Z},\, n\neq 0} \tfrac{1}{n}\alpha_n^\mu \exp(-\tfrac{i\pi n}{l}\tau)\sin(\tfrac{n\pi\sigma}{l})\,, \\ \partial_\pm X^\mu &= \pm\frac{\pi}{l}\sqrt{\frac{\alpha'}{2}} \sum_{n\in\mathbb{Z}} \alpha_n^\mu \exp(-\tfrac{i\pi n}{l}(\tau \pm \sigma)) \end{aligned} \tag{2.46}$$

with reality condition $(\alpha_n^\mu)^* = \alpha_{-n}^\mu$ as before.

3. Mixed boundary conditions (ND): fixing (N) at $\sigma = 0$ and (D) are $\sigma = l$ with $X^\mu(\sigma = l) = x^\mu$ (or the way around):

$$
\begin{aligned}
X^\mu(\tau,\sigma) &= x^\mu + i\sqrt{2\alpha'} \sum_{n\in\mathbb{Z}+\frac{1}{2}} \frac{1}{n}\alpha_n^\mu \exp(-\tfrac{i\pi n}{l}\tau)\cos(\tfrac{n\pi\sigma}{l}) \,, \\
\partial_\pm X^\mu &= \frac{\pi}{l}\sqrt{\frac{\alpha'}{2}} \sum_{n\in\mathbb{Z}+\frac{1}{2}} \alpha_n^\mu \exp(-\tfrac{i\pi n}{l}(\tau \pm \sigma)) \,.
\end{aligned}
\tag{2.47}
$$

Notice that the Fourier coefficients, which must still satisfy $(\alpha_n^\mu)^* = \alpha_{-n}^\mu$, are labelled by half-integers here.

Note that one can choose separate boundary conditions for each of the D scalar fields X^μ individually. This means that the ends of a string can move freely in $(D-p)$ spacetime directions, but fixed in p others, as if they are attached to a p-dimensional object that extends in these directions coordinates. It turns out that these objects, which are called ***p*-brane**, are needed for internal consistency of string theory at a non-perturbative level. We will come back to branes at a later point.

Poincaré charges To get some more physical intuition for the mode expansion, let us consider the Poincaré symmetry charges of a moving string. In the worldsheet description, these are the Noether charges associated to the global symmetry (2.17) for the fields X^μ. Applying Noether's theorem on their infinitesimal versions,

$$\delta X^\mu = \epsilon^\mu \,, \qquad \text{(spacetime translations)} \tag{2.48}$$

$$\delta X^\mu = \epsilon^{\mu\nu} X_\nu \,, \qquad \text{(Lorentz transformations)} \tag{2.49}$$

we find the currents (note that these are vectors on Σ!)

$$q^\mu{}_a(\xi) = -T\partial_a X^\mu \,, \qquad \text{(spacetime translations)} \tag{2.50}$$

$$J^{\mu\nu}{}_a(\xi) = -T(X^\mu\partial_a X^\nu - X^\nu\partial_a X^\mu) = X^\mu q_a^\nu - X^\nu q_a^\mu \,, \qquad \text{(Lorentz transformations)} \tag{2.51}$$

which are conserved, i.e., $\partial_a(q^\mu)^a = \partial_a(J^{\mu\nu})^a = 0$.

The Noether charges are then the sptial integral over the temporal component of each current. Plugging in the above mode expansions, we find for spacetime translations the Noether charge

$$\int_0^l d\sigma\,(q^\mu)^\tau = p^\mu\,, \tag{2.52}$$

for the closed string and the open string with (NN) boundary conditions, and 0 for open strings with Dirichlet boundary conditions. Clearly, the mode expansion parameter p^μ has the interpretation of the center-of-mass momentum of the string.

The Noether charges for Lorentz transformations are the angular momenta

$$M^{\mu\nu} = \int_0^l d\sigma (J^{\mu\nu})^\tau = \begin{cases} \ell^{\mu\nu} + E^{\mu\nu} + \tilde{E}^{\mu\nu} & \text{(closed string)}, \\ \ell^{\mu\nu} + E^{\mu\nu} & \text{(open string)}, \end{cases} \tag{2.53}$$

where

$$\ell^{\mu\nu} = x^\mu p^\nu - x^\nu p^\mu \tag{2.54}$$

is the center-of-mass contribution, and

$$E^{\mu\nu} = \sum_{n\neq 0} -\frac{i}{n}(\alpha^\mu_{-n}\alpha^\nu_n - \alpha^\nu_{-n}\alpha^\mu_n)\,, \quad \tilde{E}^{\mu\nu} = \sum_{n\neq 0} -\frac{i}{n}(\tilde{\alpha}^\mu_{-n}\tilde{\alpha}^\nu_n - \tilde{\alpha}^\nu_{-n}\tilde{\alpha}^\mu_n)\,, \tag{2.55}$$

are the contributions from the left- and right-moving waves on the string.

2.3.3 Imposing the constraints

Recall that the energy-momentum tensor $T_{ab} = \partial_a X \cdot \partial_b X - \frac{1}{2}\eta_{ab}\partial_c X \cdot \partial^c X$ (in conformal gauge) satisfies certain relations due to Weyl and diffeomorphism invariance. In lightcone coordinates, these are

$$\begin{aligned} T^a{}_a &= T_{+-} + T_{-+} = 0 \;\Rightarrow\; T_{+-} = T_{-+} = 0\,, \\ \nabla^a T_{ab} &= 0 \;\Rightarrow\; \partial_- T_{++} = \partial_+ T_{--} = 0\,. \end{aligned} \tag{2.56}$$

The second conditions implies that $T_{++} = -\partial_+ X \cdot \partial_+ X$ depends only on the coordinate ξ^+, and $T_{--} = -\partial_- X \cdot \partial_- X$ only on ξ^-. This is the 2d Lorentzian analog to (anti-)holomorphicity, and implies the existence of *infinite many* conserved charges.

Closed strings To construct these conserved charges for the closed string, consider, for any function $f(\xi^-)$,

$$\begin{aligned} Q_f &:= \int d\sigma f(\xi^-)T_{--}(\xi^-) \\ \Longrightarrow \quad \partial_\tau Q_f &= \int d\sigma (2\partial_+ - \partial_\sigma) f(\xi^-)T_{--}(\xi^-) \\ &= -\int d\sigma\, \partial_\sigma (f(\xi^-)T_{--}(\xi^-)) \\ &= \Big[f(\xi^-)T_{--}(\xi^-)\big|_{\sigma=l} - f(\xi^-)T_{--}(\xi^-)\big|_{\sigma=0} \Big]\Big|_{\text{fixed } \tau} . \end{aligned} \tag{2.57}$$

With periodic boundary conditions for X (and therefore also for T_{ab}), $\partial_\tau Q_f$ vanishes if $f(\xi^-)$ is also periodic in σ. A complete set of such functions is $f_m(\xi^-) = \exp(\frac{2\pi i}{l} m\xi^-)$. A straightforward computation, using the mode expansion (2.39), then yields the conserved charges

$$\begin{aligned} L_m :=& \frac{T}{2l}\int_0^l d\sigma e^{\frac{2\pi i}{l} m\xi^-} T_{--}(\xi^-) = \frac{T}{2l}\int_0^l d\sigma e^{\frac{2\pi i}{l} m\xi^-}\, \partial_- X_R \cdot \partial_- X_R \\ =& \frac{1}{2}\sum_{n\in\mathbb{Z}} \alpha_{m-n}\cdot\alpha_n \quad (\text{with } \alpha_0^\mu = \sqrt{\tfrac{\alpha'}{2}} p^\mu) . \end{aligned} \tag{2.58}$$

Similar considerations with T_{++} and $f_m(\xi^+) = \exp(\frac{2\pi i}{l} m\xi^+)$ lead to conserved charges

$$\tilde{L}_m := \frac{T}{2l}\int_0^l d\sigma e^{\frac{2\pi i}{l} m\xi^+} T_{++}(\xi^+) = \frac{1}{2}\sum_{n\in\mathbb{Z}} \tilde{\alpha}_{m-n}\cdot\tilde{\alpha}_n \quad (\text{with } \tilde{\alpha}_0^\mu = \sqrt{\tfrac{\alpha'}{2}} p^\mu) . \tag{2.59}$$

The L_m's and $\tilde{L}_m$'s are just Fourier coefficients of T_{--} and T_{++}, respectively. So the constraints (2.32),

$$T_{++} \overset{!}{=} 0 , \quad T_{--} \overset{!}{=} 0 , \tag{2.60}$$

are equivalent to $L_m = 0$ and $\tilde{L}_m = 0$ for all m. Amongst these infinitely many quadratic constraints, those for $m = 0$ are particularly interesting. From

$$\begin{aligned} L_0 &= \frac{1}{2}\sum_{n\in\mathbb{Z}} \alpha_{-n}\cdot\alpha_n = \frac{1}{2}\alpha_0^2 + \sum_{n=1}^\infty \alpha_{-n}\cdot\alpha_n = \frac{\alpha'}{4}p^2 + \sum_{n=1}^\infty \alpha_{-n}\cdot\alpha_n = 0 , \\ \tilde{L}_0 &= \frac{1}{2}\sum_{n\in\mathbb{Z}} \tilde{\alpha}_{-n}\cdot\tilde{\alpha}_n = \frac{1}{2}\tilde{\alpha}_0^2 + \sum_{n=1}^\infty \tilde{\alpha}_{-n}\cdot\tilde{\alpha}_n = \frac{\alpha'}{4}p^2 + \sum_{n=1}^\infty \tilde{\alpha}_{-n}\cdot\tilde{\alpha}_n = 0 , \end{aligned} \tag{2.61}$$

we obtain the *level matching condition*,

$$\sum_{n=1}^{\infty} \alpha_{-n} \cdot \alpha_n = \frac{\alpha'}{4} p^2 = \sum_{n=1}^{\infty} \tilde{\alpha}_{-n} \cdot \tilde{\alpha}_n \,, \tag{2.62}$$

which relates the left- and right-movers. Moreover, since $\alpha_0^\mu \propto p^\mu$ is the spacetime center-of-mass momentum of the string, the L_0 and $\tilde{L}_0$ constraints also give the *mass shell condition*,

$$M^2 = -p^2 = \frac{2}{\alpha'} \sum_{n=1}^{\infty} (\alpha_{-n} \cdot \alpha_n + \tilde{\alpha}_{-n} \cdot \tilde{\alpha}_n) \,, \tag{2.63}$$

which quantifies the contributions of the oscillator modes to the effective mass of the string in spacetime.

Open strings For open strings, the boundary conditions relate the left- and right-moving waves, and, by extension, also the different components of the energy-momentum tensor. For simplicity, let us focus here just on open strings with (NN) boundaries, i.e., $\partial_\sigma X^\mu = 0$ for both $\sigma = 0$ and $\sigma = l$. At the boundaries, we therefore have $\partial_+ X = \partial_- X = \partial_\tau X$, so $T_{--}|_{\sigma=\text{bdry}} = T_{++}|_{\sigma=\text{bdry}}$. Analogous to (2.57), we can then show that the conserved charges are

$$Q_f = \int d\sigma \left(f(\xi^-) T_{--}(\xi^-) + f(\xi^+) T_{++}(\xi^+) \right) \,, \tag{2.64}$$

if $f(x - l) = f(x + l)$ is a $2l$-periodic function. A complete set of charges for the (NN) open string, with mode expansion (2.44), is then given by

$$L_m = \frac{T}{2l} \int_0^l d\sigma \left(e^{\frac{\pi i}{l} m \xi^-} T_{--} + e^{\frac{\pi i}{l} m \xi^+} T_{++} \right) = \frac{1}{2} \sum_{n \in \mathbb{Z}} \alpha_{m-n} \cdot \alpha_n \,, \tag{2.65}$$

which for $m = 0$ gives the open-string mass shell condition

$$M^2 = -p^2 = \frac{1}{\alpha'} \sum_{n=1}^{\infty} \alpha_{-n} \cdot \alpha_n \,. \tag{2.66}$$

2.3.4 Poisson brackets and conformal symmetry

So far, the discussion was in the Lagrangian formalism. There, we have explicitly constructed the solutions $(X^\mu, \partial_a X^\mu)$ of the equation of motions in terms of their mode expansion parameters $(x^\mu, p^\mu, \alpha_n^\mu, \tilde{\alpha}_n^\mu)$, subject to the constraints $L_m = \tilde{L}_m = 0$ for all $m \in \mathbb{Z}$. For the usual quantization procedure, we need to pass to the Hamiltonian formalism. Specifically, we need to compute the *Poisson bracket* relations for classical observables, and then promote them to commutator relations of quantum operators.

In Hamiltonian formalism, the solution space is parametrized by the fields X^μ and their conjugate momenta. In conformal gauge, the Polyakov action (2.30) with Lagrangian $\mathcal{L} = \frac{T}{2}(\partial_\tau X \cdot \partial_\tau X - \partial_\sigma X \cdot \partial_\sigma X)$, defines the conjugate momenta,

$$\Pi^\mu = \frac{\delta \mathcal{L}}{\delta(\partial_\tau X_\mu)} = T \partial_\tau X^\mu \,. \tag{2.67}$$

The *Hamiltonian* is then

$$\begin{aligned} H &= \int_0^l d\sigma (\partial_\tau X^\mu \Pi_\mu - \mathcal{L}) = \frac{T}{2} \int_0^l d\sigma [(\partial_\tau X)^2 + (\partial_\sigma X)^2] \\ &= T \int_0^l d\sigma [(\partial_+ X)^2 + (\partial_- X)^2] = \begin{cases} \frac{2\pi}{l}(L_0 + \tilde{L}_0)\,, & \text{closed strings,} \\ \frac{\pi}{l} L_0\,, & \text{open strings.} \end{cases} \end{aligned} \tag{2.68}$$

A general observable is then a functional F in $(X^\mu(\tau,\sigma), \Pi^\mu(\tau,\sigma))$, which are coordinates on the *phase space*. The Poisson bracket is a symplectic pairing $\{F, G\}_{\text{PB}} = -\{G, F\}_{\text{PB}}$, defined by

$$\{F, G\}_{\text{PB}}(\tau, \sigma, \sigma') = \int d\tilde{\sigma} \left(\frac{\delta F(\tau,\sigma)}{\delta \Pi_\mu(\tau,\tilde{\sigma})} \frac{\delta G(\tau,\sigma')}{\delta X^\mu(\tau,\tilde{\sigma})} - \frac{\delta G(\tau,\sigma')}{\delta \Pi_\mu(\tau,\tilde{\sigma})} \frac{\delta F(\tau,\sigma)}{\delta X^\mu(\tau,\tilde{\sigma})} \right) . \tag{2.69}$$

The canonical equal time Poisson brackets are

$$\{X, X\}_{\text{PB}} = \{\Pi, \Pi\}_{\text{PB}} = 0\,, \quad \{\Pi^\mu(\tau,\sigma), X^\nu(\tau,\sigma')\}_{\text{PB}} = \eta^{\mu\nu} \delta(\sigma - \sigma')\,. \tag{2.70}$$

From these, we can compute the Poisson brackets of the oscillator modes from Fourier expansion. One finds the non-zero results

$$\{\alpha_m^\mu, \alpha_n^\nu\}_{\text{PB}} = im\delta_{m+n,0}\, \eta^{\mu\nu} = \{\tilde{\alpha}_m^\mu, \tilde{\alpha}_n^\nu\}_{\text{PB}}\,, \quad \{p^\mu, x^\nu\}_{\text{PB}} = \eta^{\mu\nu}\,, \tag{2.71}$$

where, for open strings, there is only one (independent) set of oscillators.

We can use these to further compute the Poisson brackets for the Fourier modes $(L_m, \tilde{L}_m)$ of the energy momentum tensor. Explicit computation (see problem sheet) shows

$$\{L_m, X^\mu\}_{\mathrm{PB}} = -\frac{l}{2\pi} e^{\frac{2\pi i}{l} m\xi^-} \partial_- X^\mu \,, \quad \{\tilde{L}_m, X^\mu\}_{\mathrm{PB}} = -\frac{l}{2\pi} e^{\frac{2\pi i}{l} m\xi^+} \partial_+ X^\mu \,, \quad (2.72)$$

and the important relation

$$\{L_m, L_n\}_{\mathrm{PB}} = i(m-n)L_{m+n} \,, \quad \{\tilde{L}_m, \tilde{L}_n\}_{\mathrm{PB}} = i(m-n)\tilde{L}_{m+n} \,. \quad (2.73)$$

These relations define a Lie algebra called the **Witt algebra**, which are the generators for *conformal transformations* in 2d.

These transformations are diffeomorphisms of a Riemannian/Lorentzian manifold $\mathcal{M}$ that preserves the metric up to rescaling,

$$g(x) \to \tilde{g}(\tilde{x}) = e^{2\omega(\tilde{x})} g(\tilde{x}) \,. \quad (2.74)$$

What are the generators for such transformations on the worldsheet? Consider a general infinitesimal diffeomorphism, which acts via

$$\xi^a \to \xi^a + \epsilon^a(\xi) \,, \quad \gamma_{ab} \to \gamma_{ab} + \nabla_a \epsilon_b + \nabla_b \epsilon_a . \quad (2.75)$$

For this to be a conformal transformation, ϵ must satisfy the *conformal Killing equation* $\nabla_a \epsilon_b + \nabla_b \epsilon_a = (\nabla_c \epsilon^c)\gamma_{ab}$. In unit gauge ($\gamma_{ab} = \eta_{ab}$ and $\nabla_a = \partial_a$) and lightcone coordinates ($\eta_{+-} = \eta_{-+} = -\frac{1}{2}, \eta^{+-} = \eta^{-+} = -2$, all other entries zero), this equation becomes quite simple:

$$\begin{aligned}
(++): &\quad \partial_+ \epsilon_+ = 0 \;\Rightarrow\; \partial_+ \epsilon^- = 0 \;\Rightarrow\; \epsilon^- = \epsilon^-(\xi^-) \,, \\
(--): &\quad \partial_- \epsilon_- = 0 \;\Rightarrow\; \partial_- \epsilon^+ = 0 \;\Rightarrow\; \epsilon^+ = \epsilon^+(\xi^+) \,, \\
(+-): &\quad \partial_+ \epsilon_- + \partial_- \epsilon_+ = -\frac{1}{2}(\partial^a \epsilon_a) = -\frac{1}{2}(-2\partial_- \epsilon_+ - 2\partial_+ \epsilon_-) \quad \text{(trivially true)} \,.
\end{aligned} \quad (2.76)$$

This means that the conformal transformations are generated infinitesimally by "(anti-)holomorphic" vector fields $\epsilon^\pm(\xi^\pm)$, for which we can, as before, pick a complete set in terms of $e^{\frac{2\pi i}{l} n\xi^\pm}$.

When we represent (tangent) vector fields on a manifold as differential operators (i.e., we use a local basis $e_a = \partial_a$ for the tangent space, see GR I), this gives rise to a complete set of operators

$$V_n = -\frac{l}{2\pi} e^{\frac{2\pi i}{l} n\xi^-} \partial_- \,, \quad \tilde{V}_n = -\frac{l}{2\pi} e^{\frac{2\pi i}{l} n\xi^+} \partial_+ \,, \quad (2.77)$$

which have the *commutator relations* (see problem sheet)

$$[V_n, V_m] = i(n-m)V_{n+m}\,, \tag{2.78}$$

and similarly for $\tilde{V}_{n,m}$'s. As claimed, these are precisely the defining relations of the Witt algebra.

What this tells us is that the modes L_m and $\tilde{L}_m$'s generate conformal transformations on the phase space.[3] This means in particular that, even after fixing to unit gauge, the worldsheet theory still has residual gauge symmetries, namely *conformal symmetry*: The constraints $L_m = 0$ and $\tilde{L}_m = 0$ can be understood as requiring all observables to not vary under conformal transformations.

The appearance of conformal symmetry suggests a proper treatment using *conformal field theory* (CFT). We will not be able to give a full introduction to this topic, and instead simply "import" some tools as we move along; it is highly recommended that you take the course in Trinity term (and/or consult the references).

However, we remark that the worldsheet theory is slightly different than an ordinary CFT, in that here, the conformal symmetry is a *gauge* symmetry. One could in principle imagine further gauge-fixing the system; however, it turns out there is no way to do so without losing a manifestly spacetime-covariant description. One such gauge-fixing procedure is *lightcone gauge*, which we will revisit briefly in the next section.

To close our discussion of the classical string, we point out that conformal symmetry in $(1+1)$ dimensions is special in that the algebra has, as seen above, infinitely many generators. This enormous symmetry is what makes the quantization of the classical description possible in the first place. In $d > 2$ dimensions (with Lorentzian signature), the conformal algebra is the finite algebra $\mathfrak{so}(2, d)$. Extrapolating to $d = 2$, $\mathfrak{so}(2,2) \cong \mathfrak{sl}(2,\mathbb{R}) \oplus \widetilde{\mathfrak{sl}(2,\mathbb{R})}$ is only the "global part",

[3] More precisely, in the Hamiltonian picture, it is a standard feature that symmetry transformations parametrized by ϵ are always generated by the conserved charges Q_ϵ via the Poisson bracket: $\delta_\epsilon F = \{Q_\epsilon, F\}_{\text{PB}}$ where δF is the infinitesimal change of an observable F under this transformation. For transformations that arise from diffeomorphisms on spacetime (parametrized by a vector field ϵ), there is an analogous concept called the *Lie derivative*, which encodes infinitesimal transformations $\delta F = \mathcal{L}_\epsilon F$, or *flows*, of any differential form F. From GR I, you might recall that the Lie bracket is intimately tied to commutators of vector fields; in particular, they satisfy $[\mathcal{L}_\epsilon, \mathcal{L}_{\epsilon'}](F) := (\mathcal{L}_\epsilon \mathcal{L}_{\epsilon'} - \mathcal{L}_{\epsilon'} \mathcal{L}_\epsilon)(F) = \mathcal{L}_{[\epsilon,\epsilon']}(F)$. What we have shown above is that the flows generated by the vector fields V_n are structurally identical to the phase space transformations generated by the L_n's, which further implies $\{Q_\epsilon, Q_{\epsilon'}\}_{\text{PB}} = Q_{[\epsilon,\epsilon']}$.

generated by $\{V_{0,\pm1}, \tilde{V}_{0,\pm1}\}$ of the full conformal algebra $\mathrm{Witt} \oplus \widetilde{\mathrm{Witt}}$. Consequently, CFTs in higher dimensions are much more complicated to study.

3 The Quantum String

We now move to the quantum string. More specifically, we will quantize the $(1+1)$d field theory on the worldsheet starting from the Polyakov action. As oftentimes with "ordinary" field theories, there are generally new consistency conditions that arise that go beyond the classical constraints. In the case of the bosonic string, one of the main quantum consistency conditions is *criticality*: the number of scalar fields X^μ — which is the spacetime dimension D — must be 26!

There are two quantization methods for the string, starting from the classical Hamiltonian picture.

- In **old covariant quantization (OCQ)**, which is what we will discuss in this course, we promote the mode expansion coefficients to quantum operators, and then impose the constraints at the level of the states. This manifestly preserves spacetime covariance, but the quantum theory is only *unitary* in the critical dimension $D = 26$.
- In **lightcone quantization (LCQ)**, the constraints are implemented before quantizing, leading to a manifestly unitary theory. However, spacetime covariance is only restored in $D = 26$.

For completeness, note that there is also a third quantization method, starting from the Lagrangian formulation, namely the **path-integral quantization**. This manifestly covariant procedure is sometimes also called the *modern* covariant quantization of the string. Again, the critical dimension arises as a consistency condition, but the meaning of the phrase "the BRST algebra needs to be closed" will be explained in the "Advanced QFT" course.

3.1 Old covariant quantization

3.1.1 Canonical quantization

The canonical quantization procedure promotes the fields X^μ and Π^μ, which were coordinates on phase space in the classical Hamiltonian picture, to operators (which we will, by laziness, denote with the same symbols) acting on the Hilbert space of states. Their classical Poisson bracket relation is now promoted to an *equal time commutation relation*, which reads

$$[X^\mu(\tau,\sigma),\Pi^\nu(\tau,\sigma')] = -[\Pi^\nu(\tau,\sigma'),X^\mu(\tau,\sigma)] = i\eta^{\mu\nu}\delta(\sigma-\sigma')\,. \tag{3.1}$$

Accordingly, also the coefficients of the mode expansions become operators, and complex conjugation becomes taking the Hermitian-conjugate operator:

$$\begin{aligned}
&[x^\mu,p^\nu] = i\eta^{\mu\nu}\,, && x^\mu = (x^\mu)^\dagger,\;\; p^\mu = (p^\mu)^\dagger,\\
&[\alpha^\mu_m,\alpha^\nu_n] = m\,\delta_{m+n,0}\,\eta^{\mu\nu}\,, && \alpha^\mu_{-n} = (\alpha^\mu_n)^\dagger,\\
&[\tilde\alpha^\mu_m,\tilde\alpha^\nu_n] = m\,\delta_{m+n,0}\,\eta^{\mu\nu}\,, && \tilde\alpha^\mu_{-n} = (\tilde\alpha^\mu_n)^\dagger.
\end{aligned} \tag{3.2}$$

This forms an infinite set of harmonic oscillator (after rescaling $a_n = \alpha_n/\sqrt{n}$) plus a standard “Heisenberg pair” (x,p). This allows us to construct the Hilbert space in a straightforward manner.

First, we look at the oscillators, and identify $(\alpha^\mu_{-n},\alpha^\mu_n)_{n>0}$ (same for the tilded ones) as raising and lowering operators. With the oscillator vacuum state $|0\rangle_o$ defined by

$$\alpha^\mu_n|0\rangle_o = \tilde\alpha^\mu_n|0\rangle_o = 0 \quad \text{for } n>0\,, \tag{3.3}$$

we then build the “oscillator Fock spaces” using the raising operators,

$$\begin{aligned}
\mathcal{H}^{\text{Fock}}_{\text{open}} &= \operatorname{span}_{\mathbb{C}}\left\{\prod_{\mu=0}^{D-1}\prod_{n=1}^{\infty}\left(\alpha^\mu_{-n}\right)^{N^\mu_n}|0\rangle_o \,\middle|\, N^\mu_n \geq 0,\;\text{finitely many} \neq 0\right\}\,,\\
\mathcal{H}^{\text{Fock}}_{\text{closed}} &= \operatorname{span}_{\mathbb{C}}\left\{\prod_{\mu,\nu=0}^{D-1}\prod_{n=1}^{\infty}\prod_{m=1}^{\infty}\left(\alpha^\mu_{-n}\right)^{N^\mu_n}\left(\tilde\alpha^\nu_{-m}\right)^{\tilde N^\nu_j}|0\rangle_o \,\middle|\, N^\mu_n,\tilde N^\nu_m \geq 0,\;\text{finitely many} \neq 0\right\}\\
&\cong \tilde{\mathcal{H}}^{\text{Fock}}_{\text{left}} \otimes \mathcal{H}^{\text{Fock}}_{\text{right}} \cong \mathcal{H}^{\text{Fock}}_{\text{open}} \otimes \mathcal{H}^{\text{Fock}}_{\text{open}}\,.
\end{aligned} \tag{3.4}$$

As for the harmonic oscillator, there are “counting operators” $N := \sum_{k>0}\alpha_{-k}\cdot\alpha_k$ and $\tilde N = \sum_{k>0}\tilde\alpha_{-k}\cdot\tilde\alpha_k$, which satisfy

$$[N,\alpha^\mu_n] = -n\,\alpha^\mu_n\,, \quad [N,(\alpha^\mu_n)^\dagger] = [N,\alpha^\mu_{-n}] = n\,\alpha^\mu_{-n}\,, \tag{3.5}$$

(analogously for $(\tilde{N}, \tilde{\alpha}_n^\mu)$) and which measure the oscillation "quanta",

$$N\left(\prod_{\mu=0}^{D-1}\prod_{n=1}^{\infty}\left(\alpha_{-n}^{\mu}\right)^{N_n^\mu}|0\rangle_o\right) = \left(\sum_{\nu=0}^{D-1}\sum_{m=1}^{\infty} m\, N_m^\nu\right)\left(\prod_{\mu=0}^{D-1}\prod_{n=1}^{\infty}\left(\alpha_{-n^\mu}^{\mu}\right)^{N_n^\mu}|0\rangle_o\right). \quad (3.6)$$

This allows us to organize oscillator states into *levels* (eigenstates under N and $\tilde{N}$). Focusing on left- or right-movers only, or equivalently, the open string states, we have basis vectors

$$\begin{aligned} N = 0 :&\quad |0\rangle_o\,; \\ N = 1 :&\quad \alpha_{-1}^{\mu}|0\rangle_o\,; \\ N = 2 :&\quad \alpha_{-2}^{\mu}|0\rangle_o\,,\quad \alpha_{-1}^{\mu}\alpha_{-1}^{\nu}|0\rangle_o\,; \\ N = 3 :&\quad \alpha_{-3}^{\mu}|0\rangle_o\,,\quad \alpha_{-2}^{\mu}\alpha_{-1}^{\nu}|0\rangle_o\,,\quad \alpha_{-1}^{\mu}\alpha_{-1}^{\nu}\alpha_{-1}^{\rho}|0\rangle_o\,; \\ &\quad \vdots \end{aligned} \quad (3.7)$$

Since we also have the zero modes (x^μ, p^ν) (which commute with the oscillators), the oscillator modes can be paired with the standard representations ("wave functions") for the Hilbert space $\mathcal{H}_{\text{zero-modes}} \cong L^2(\mathbb{R}^{1,D-1})$ of D Heisenberg pairs. We use the momentum space representation, in which the plane wave states $|k\rangle$, $k \in \mathbb{R}^{1,D-1}$ are eigenvectors of the momentum operator, $p^\mu|k\rangle = k^\mu|k\rangle$, and $\langle k'|k\rangle = \delta^{(D)}(k' - k)$. The full Hilbert space is then

$$\mathcal{H}_{\text{open}} = \mathcal{H}_{\text{open}}^{\text{Fock}} \otimes L^2(\mathbb{R}^{1,D-1})\,,\quad \mathcal{H}_{\text{closed}} = \tilde{\mathcal{H}}_{\text{left}}^{\text{Fock}} \otimes \mathcal{H}_{\text{right}}^{\text{Fock}} \otimes L^2(\mathbb{R}^{1,D-1})\,. \quad (3.8)$$

A basis for these spaces can be built starting from the "vacua" $|0; k\rangle$ with $\alpha_n^\mu|0; k\rangle = \tilde{\alpha}_n^\mu|0; k\rangle = 0$ for $n > 0$, and acting on it with the raising operators of the oscillators.

Within this Hilbert space, the states are labelled by spacetime momenta k and spacetime tensor indices $(\mu, \nu, ...)$, and so manifestly exhibit spacetime covariance (i.e., they naturally furnish representations under spacetime Poincaré symmetry). However, this desirable property also introduces (at face value) some severe problems: the existence of *negative-norm*, or ghost states. An example of such states are level-one states of the form $\alpha_{-1}^0|0; k\rangle$, for which we have

$$\langle 0; k|\alpha_+^0\, \alpha_-^0|0; k'\rangle = \langle 0; k|[\alpha_+^0, \alpha_-^0]|0; k'\rangle = \eta^{00}\delta^{(D)}(k - k') = -\delta^{(D)}(k - k')\,. \quad (3.9)$$

The existence of physical ghost states would lead to a violation of unitarity, which in general is unacceptable in quantum theories. Therefore, we must ensure their absence. Luckily, we still have the constraints $T_{ab} = 0$ at our disposal. As we see now, these will distinguish *physical* states with positive norms from ghost states.

3.1.2 Implementing the constraints

Recall that the classical constraints were $L_m = \tilde{L}_m = 0$. Once we quantize, the constraints are now implemented as operator equations. However, as in the *Gupta–Bleuler quantization* of QED (where the classical constraint $\partial_\mu A^\mu = 0$ was to implement Lorenz gauge), we only require the vanishing of the "matrix elements" $\langle\varphi|L_m|\varphi'\rangle$ for *physical* states $|\varphi\rangle$, $|\varphi'\rangle$. Because $L_{-m} = (L_m)^\dagger$ and $\tilde{L}_{-m} = (\tilde{L}_m)^\dagger$, it suffices to impose

$$\forall m \geq 0 : \; L_m|\varphi\rangle = 0 \;\Leftrightarrow\; |\varphi\rangle \text{ physical state}\,. \tag{3.10}$$

For $m \neq 0$, we can, without hesitation, identify the classical expressions,

$$L_m = \frac{1}{2}\sum_{n\in\mathbb{Z}} \alpha_{m-n}\cdot\alpha_n\,, \quad \tilde{L}_m = \frac{1}{2}\sum_{n\in\mathbb{Z}} \tilde{\alpha}_{m-n}\cdot\tilde{\alpha}_n\,, \tag{3.11}$$

as good quantum operators, because α_{m-n} and α_n commute for $m \neq 0$.

However, for $m = 0$, the classical expressions,

$$L_0 = \frac{1}{2}\sum_{n\in\mathbb{Z}} \alpha_{-n}\cdot\alpha_n\,, \quad \tilde{L}_0 = \frac{1}{2}\sum_{n\in\mathbb{Z}} \tilde{\alpha}_{-n}\cdot\tilde{\alpha}_n\,, \tag{3.12}$$

suffer from an *ordering ambiguity*. That is, depending on how we order the raising/lowering operators for each n, the action of L_0 and $\tilde{L}_0$ on one state can differ by a c-number. Said differently, if we *define* the quantum operators L_0 and $\tilde{L}_0$ in *normal ordering*,

$$\begin{aligned} L_0 &:= \frac{1}{2}\alpha_0\cdot\alpha_0 + \sum_{n=1}^{\infty}\alpha_{-n}\cdot\alpha_n = \frac{1}{2}\alpha_0\cdot\alpha_0 + N\,, \\ \tilde{L}_0 &:= \frac{1}{2}\tilde{\alpha}_0\cdot\tilde{\alpha}_0 + \sum_{n=1}^{\infty}\tilde{\alpha}_{-n}\cdot\tilde{\alpha}_n = \frac{1}{2}\tilde{\alpha}_0\cdot\tilde{\alpha}_0 + \tilde{N}\,, \end{aligned} \tag{3.13}$$

then we can only formulate the constraint for $L_0/\tilde{L}_0$ as

$$(L_0 - a)|\varphi\rangle = 0\,, \quad (\tilde{L}_0 - \tilde{a})|\varphi\rangle = 0\,, \quad \forall \text{ physical } |\varphi\rangle\,, \tag{3.14}$$

where $(a, \tilde{a})$ are some constants that we cannot specify just yet. For now, we can, with some more advanced arguments, relate the difference $a - \tilde{a}$ to a kind of *gravitational anomaly* on the worldsheet, i.e., a violation of diffeomorphism invariance. Its absence therefore sets $a = \tilde{a}$.

The Virasoro algebra The ordering issue will also modify the quantum commutators $[L_m, L_n]$, so that their expression deviates from the naive quantum version of the classical Poisson bracket relations,

$$[L_m, L_n]_{\text{Witt}} = (m-n)L_{m+n}\,. \tag{3.15}$$

Namely, whenever $m + n = 0$, the right-hand side would be proportional to L_0, which, as discussed above, suffers from an ordering ambiguity. On the problem sheet, you will show that if, as above, we *define* L_0 to be normal ordered (3.13), then the commutator must be

$$[L_m, L_n] = (m-n)L_{m+n} + \frac{D}{12}(m^3 - m)\delta_{m+n,0}\,. \tag{3.16}$$

This is the defining commutation relation of the **Virasoro algebra** with *central charge* $c = D$ (number of spacetime dimensions), and we will refer to the quantum operators L_m (and $\tilde{L}_m$) as the *Virasoro generators*. The Virasoro algebra is a so-called *central extension* of the Witt algebra by $\mathbb{C} \ni c$:

$$\begin{aligned} &[x,y]_{\text{Vir.}} := [x,y]_{\text{Witt}} + c\,p(x,y)\,, \quad \forall\, x,y \in \text{Witt}\,, \\ &[x,a]_{\text{Vir.}} := 0\,, \quad [a,a]_{\text{Vir}} := 0\,, \quad \forall\, x \in \text{Witt}, \ \forall\, a \in \mathbb{C}\,. \end{aligned} \tag{3.17}$$

That is, we extend the original algebra (in this case Witt) by an element c (and any complex multiple of it) that commutes with every element (hence, is “central”), but that can appear in the commutator of other elements. The map $p(x,y) \in \mathbb{C}$ must be bilinear and anti-symmetric.

The central charge can be understood as a quantum anomaly of conformal symmetry; we will come back to this later. Notice how the central charge vanish for the subalgebra generated by $\{L_{0,\pm1}\}$. This means that the “global” conformal transformations, $\mathfrak{sl}(2,\mathbb{R})$, are non-anomalous.

3.2 Critical dimension as quantum consistency

The consistency of the quantized worldsheet theory hinges on consistently defining the subspace $\mathcal{H}_{\text{phys}} \subset \mathcal{H}$ of physical states by the Virasoro constraints,

$$|\phi\rangle \in \mathcal{H}_{\text{phys}} \iff \begin{cases} \forall m > 0 :\ \tilde{L}_m|\phi\rangle = 0\,, \\ (L_0 - a)|\phi\rangle = (\tilde{L}_0 - a)|\phi\rangle = 0\,. \end{cases} \tag{3.18}$$

Given the basis vectors (3.7), we can try to implement these constraints level by level, which turns the constraints into a set of linear equations. The non-trivial consistency condition here is that the solutions of these equations, being physical states, must not have negative norms. As we will see now, these arguments pose restrictions to the normal ordering constant a and spacetime dimension D.

3.2.1 Ground state masses

The normal ordering constant has a very concrete physical meaning. Recall that, classically, the L_0 constraints gave rise to the mass shell conditions (2.63) and (2.66). In the quantum theory, the ordering ambiguity enters these as (with $\alpha_0^2 \propto p^2 \equiv -M^2$)

$$\begin{aligned} &\text{closed string :} \quad \begin{cases} (L_0 + \tilde{L}_0 - 2a)|\phi; k\rangle = 0 \;\Leftrightarrow\; \alpha' M^2 = 2N + 2\tilde{N} - 4a \\ (L_0 - \tilde{L}_0)|\phi; k\rangle = 0 \;\Rightarrow\; (N - \tilde{N})|\phi; k\rangle = 0 \end{cases} \\ &\text{open string :} \quad (L_0 - a)|\phi; k\rangle = 0 \;\Rightarrow \alpha' M^2 = N - a \end{aligned} \tag{3.19}$$

for $|\phi; k\rangle$ that are *physical* eigenstates of the operators N (and $\tilde{N}$).

At level zero, physical states $|0; k\rangle$ therefore have spacetime rest mass $M^2 = -k^2$ given by the normal ordering constant:

$$\text{open:}\;\; \alpha' M^2 = -a\,, \quad \text{closed:}\;\; \alpha' M^2 = -4a\,. \tag{3.20}$$

So, depending on the sign of a, the ground state is

- massive for $a < 0$,
- massless for $a = 0$,
- tachyonic for $a > 0$.

A tachyonic ground state is unusual, but a priori not inconsistent. It signals a quantum instability, i.e., the ground state is not the true vacuum of the theory. For the bosonic string, which eventually requires $a = 1 > 0$, it is believed that there is a true vacuum state which is non-perturbative. In superstring theory, the additional fermionic fields on the worldsheet will “lift” the tachyonic mode, and make the perturbative oscillator vacua stable ground states.

Note that at level zero, because of $[N, L_m] = -mL_m$, the Virasoro constraint $L_m|0;k\rangle = 0$ is trivially satisfied for $m > 0$, because such a state would have *negative* level. Alternatively, there is no ambiguity in bringing $L_{m>0} = \frac{1}{2}\sum_n \alpha_{n-m}\cdot\alpha_m$ into normal ordering, which by definition level zero states.

3.2.2 Normal order constant from constraints at level one

Next, we inspect states at level one; recall from earlier that this is where we encountered negative-norm states. Since the closed string Hilbert space is, in some sense, just two copies of the (NN) open string Hilbert space, we will focus on the latter for now.

A general level-one state takes the form

$$|\zeta;k\rangle := \eta_{\mu\nu}\zeta^\mu\alpha^\nu_{-1}|0;k\rangle\,, \tag{3.21}$$

where $\zeta \in \mathbb{R}^{1,D-1}$ labels the spacetime *polarization*. Their norm is

$$\langle\zeta;k|\zeta';k'\rangle = \langle 0;k|(\zeta\cdot\alpha_1)(\zeta'\cdot\alpha_{-1})|0;k'\rangle = \langle 0;k|\zeta\cdot\zeta'|0;k'\rangle = (\zeta\cdot\zeta')\delta(k-k')\,. \tag{3.22}$$

As before, we see from $[N, L_m] = -mL_m$ that the Virasoro constraints $L_m|\zeta;k\rangle$ are automatically satisfied for $m > 1$. The L_0 and L_1 constraints are then

$$\begin{aligned}
&(L_0 - a)|\zeta;k\rangle = 0 \;\Leftrightarrow\; M^2 = -k^2 = \tfrac{1}{\alpha'}(1-a)\,,\\
&L_1|\zeta;k\rangle = \eta_{\mu\nu}\zeta^\mu L_1\alpha^\nu_{-1}|0;k\rangle = \eta_{\mu\nu}\zeta^\mu[L_1,\alpha^\nu_{-1}]|0;k\rangle = \eta_{\mu\nu}\zeta^\mu\alpha^\nu_0|0;k\rangle\\
&\qquad\qquad = \eta_{\mu\nu}\zeta^\mu\sqrt{2\alpha'}k^\nu|0;k\rangle = 0 \;\Leftrightarrow\; \zeta\cdot k = 0\,,
\end{aligned} \tag{3.23}$$

where we have used

$$\begin{aligned}
[L_m,\alpha^\mu_n] &= \tfrac{1}{2}\sum_k[\alpha_{m-k}\cdot\alpha_k,\alpha^\mu_n] = \tfrac{1}{2}\sum_k\eta_{\rho\nu}\left(\alpha^\rho_{m-k}\,[\alpha^\nu_k,\alpha^\mu_n] + [\alpha^\rho_{m-k},\alpha^\mu_n]\alpha^\nu_k\right)\\
&= \tfrac{1}{2}\sum_k\eta_{\rho\nu}\left(\alpha^\rho_{m-k}\,k\,\delta_{k+n,0}\,\eta^{\nu\mu} + (m-k)\delta_{m+n-k,0}\,\eta^{\rho\mu}\alpha^\nu_k\right)\\
&= -n\,\alpha^\mu_{m+n}\,.
\end{aligned} \tag{3.24}$$

So, the full set of consistency condition for level-one physical states are

$$\begin{aligned} \zeta^2 \geq 0 \quad &\Leftrightarrow \quad \zeta \text{ light- or space-like (to avoid negative-norm)}, \\ \zeta \cdot k = 0 \quad &\Leftrightarrow \quad \text{transverse polarization } (L_1 \text{ constraint}), \\ \alpha' k^2 = a - 1 \quad & \quad\quad\;\; (L_0 \text{ constraint}) \end{aligned} \tag{3.25}$$

Depending on the value of a, there are three physically inequivalent situations.

a) For $a > 1$, the L_0 constraint tells us that k^μ is space-like, and so the L_1 constraint would be satisfied by time-like polarizations ζ. However, this would lead to negative-norm states because $\zeta^2 < 0$ for time-like ζ. These are the ghost states we encountered before, and, for $a > 1$, the Virasoro constraints are not sufficient to eliminate them.

b) For $a = 1$, the spacetime momentum k is null. In this case, the polarization vector ζ can be either *transversal*, ζ_T, or *longitudinal*, $\zeta_L \propto k$, to the momentum. The transverse polarizations are $(D-2)$ states $|\zeta_T; k\rangle$ with positive norm $\propto \zeta^2 > 0$. This is the correct number of degrees of freedom for a massless vector boson in D dimensions with polarization ζ_T in spacetime. The longitudinal one, on the other hand, is a single degree of freedom $|\lambda k; k\rangle$ with zero norm; in fact, it is orthogonal to *all physical states* $|\zeta; k'\rangle$:

$$\langle k; \lambda k | \zeta; k' \rangle = \lambda (k \cdot \zeta) \delta(k - k') = 0 \quad \text{since } k' \cdot \zeta = 0\,. \tag{3.26}$$

Therefore, the longitudinal state decouples from the physical states. Similarly to QED, this state is a "pure gauge" state and has no physical relevance; we will come back to such state in more detail below. In summary, the Virasoro constraints for $a = 1$ precisely restricts the level-one states to consistently describe spacetime vector bosons.

c) For $a < 1$, k is time-like. So the polarization ζ is a space-like vector with $(D-1)$ independent entries; the norm of such states is $\zeta^2 > 0$. This describes a massive vector boson in D dimensions with positive norm states, which is a priori acceptable from a spacetime perspective.

We conclude that implementing the Virasoro constraints at level one leads us to restrict the normal ordering constant to $a \leq 1$. For $a = 1$, we get massless vector bosons with a correct decoupling behavior for the pure gauge states. However, as we have seen above, the ground state would be tachyonic in this case. For $a < 1$, the level one states would describe massive vector bosons. In the following, we will focus on the $a = 1$ case, and see how this eventually leads to a consistent picture if $D = 26$. We will comment on $a < 1$ at the end of this section.

3.2.3 Spacetime dimension from constraints at level two

Consider level two states of the form

$$|\phi\rangle = (c_1 \alpha_{-1} \cdot \alpha_{-1} + c_2\, k \cdot \alpha_{-2} + c_3\, (k \cdot \alpha_{-1})^2)|0; k\rangle\,. \tag{3.27}$$

We can adjust the prefactors c_i and the spacetime momentum k to make this state physical. That is, we can use the three parameters (overall scaling factor not physical) to solve the three linear equations

$$(L_0 - a)|\phi\rangle = (L_0 - 1)|\phi\rangle = 0\,, \quad L_1|\phi\rangle = 0\,, \quad L_2|\phi\rangle = 0\,, \tag{3.28}$$

with all the higher Virasoro constraints satisfied automatically due to “negative level” arguments. Making use of the commutator relations

$$[L_m, \alpha_n^\mu] = -n\, \alpha_{m+n}^\mu\,, \tag{3.29}$$

it is a straightforward, but somewhat lengthy computation to show that the L_0 constraint requires $k^2 = -2/\alpha'$. Then, $L_1|\phi\rangle = L_2|\phi\rangle = 0$ sets

$$c_2 = c_1 \frac{D-1}{5}\,, \qquad c_3 = c_1 \frac{D+4}{10}\,, \tag{3.30}$$

where the spacetime dimension appears as the trace of the metric, $D = \eta_{\mu\nu}\eta^{\mu\nu}$.

However, the norm of this state is

$$\langle\phi|\phi\rangle = \frac{2|c_1|^2}{25}(D-1)(26-D)\,, \tag{3.31}$$

which is negative for $D > 26$ or $D < 1$. To avoid ghosts with $a = 1$, we must therefore restrict ourselves to $1 \leq D \leq 26$ spacetime dimensions.

In principle, we can continue with higher levels. A classic result by Brower and Goddard/Thorn (1972) is the “No Ghost Theorem for OCQ”, which proves that, for $a = 1$, $D = 26$, the space of physical states defined as the solutions of the Virasoro constraints is free of ghosts. In this case, the above states all have vanishing norm, and correspond to so-called *null states*. In fact, as explained below for the interested reader, the proliferation of null states in the critical dimension is a sign of a large gauge symmetry, namely, conformal symmetry. This is the same underlying physical principle as in the modern covariant quantization via the path-integral formalism, where the critical dimension is required to cancel the "conformal anomaly", i.e., to restore conformal symmetry in the quantum theory.

In the formalism discussed here, we cannot address the cases $a < 1$ and/or $D < 26$. However, a more refined analysis that includes string interactions will uncover violations of unitarity in these cases. This leads us to *the* bosonic string theory with $a = 1$ and $D = 26$.

Null states in the critical dimension

Here we will briefly comment on the appearance of states with vanishing norm, and how these are a "desirable" feature of a theory with gauge symmetries.

Recall from the Gupta–Bleuler formulation of QED that canonically quantized gauge theories have Hilbert spaces with states that are orthogonal to all physical states. Such states are called *spurious*. It is possible for a physical state to be spurious itself; in this case, it is also called a *null state*, as it must have zero norm (orthogonal to itself).

Null states are physical states that decouple entirely from the dynamics. Roughly speaking, they are excitations of the gauge field that can be generated by a residual (i.e., respecting the gauge fixing) gauge transformation. Therefore, two physical states are considered equivalent if they differ by a null state, $|\psi\rangle_{\text{phys}} \sim |\psi\rangle_{\text{phys}} + |\phi\rangle_{\text{null}}$. It is then customary to define the *reduced* Hilbert state,

$$\mathcal{H}_{\text{red}} := \frac{\mathcal{H}_{\text{phys}}}{\mathcal{H}_{\text{null}}}\,, \tag{3.32}$$

which describes physically distinct states.

On the string worldsheet, we know that the Virasoro generators L_m generate conformal gauge transformations. Classically, we set $L_m = 0$ for all m to ensure conformal invariance. Quantum mechanically, we can only require $L_{m>0}$ to act trivially on physical states. If conformal symmetry should remain a valid gauge symmetry, then any physical state of the form $L_{-m}|\phi\rangle$ for $m > 0$ must be null. Since there are infinitely many L_{-m}, we expect to find "many" null vectors in the physical Hilbert space.

Let us make this slightly more precise now. For the worldsheet theory, the null states are *defined* by

$$|s\rangle \;\text{null} \quad :\Longleftrightarrow \quad \begin{cases} \langle\phi|s\rangle = 0 \;\forall\; \text{physical} \;\; |\phi\rangle\,, \quad \text{and} \\ (L_0 - a)|s\rangle = 0\,, \quad L_m|s\rangle = 0 \;\; (m > 0)\,. \end{cases} \tag{3.33}$$

The orthogonality condition, i.e., $|s\rangle$ being spurious, is satisfied by the ansatz

$$|s\rangle = \sum_{m>0} L_{-m}|\chi_m\rangle\,. \tag{3.34}$$

In fact, it can be argued rigorously that any spurious state must be of this form; hence, it makes the intuition about spurious states being pure conformal gauge transformations very concrete.

Therefore, null states are the subset of these that satisfy the Virasoro constraints. It is easy to verify that the L_0 constraint requires

$$L_0|\chi_m\rangle = (a - m)|\chi_m\rangle\,. \tag{3.35}$$

To see the higher constraints in action, let us look at a simple example. We take any state $|\chi_2\rangle$ with $L_{m>0}|\chi_2\rangle = 0$, and build the spurious state

$$|s\rangle = (L_{-2} + \gamma L_{-1}^2)|\chi_2\rangle = L_{-2}|\chi_2\rangle + L_{-1}(\gamma L_{-1}|\chi_2\rangle)\,. \tag{3.36}$$

It is easy to verify that, if $L_0|\chi_2\rangle = (a-2)|\chi_2\rangle$, then $L_0(L_{-1}|\chi_2\rangle) = (a-1)L_{-1}|\chi_2\rangle$, so $|s\rangle$ satisfies $(L_0 - a)|s\rangle = 0$.

Because $L_{m>0}|\chi_2\rangle = 0$, the physical state conditions $L_m|s\rangle = 0$ for $m > 2$ are automatically satisfied. For $m = 1, 2$ we use the Virasoro commutation relations (3.16) to compute

$$
\begin{aligned}
L_1|s\rangle &= ([L_1, L_{-2}] + \overbrace{L_{-2}L_1}^{=0 \text{ on } |\chi_2\rangle} + \gamma(\overbrace{[L_1, L_{-1}]}^{=2L_0} + L_{-1}L_1)L_{-1})|\chi_2\rangle \\
&= (3L_{-1} + 2\gamma L_0 L_{-1} + \gamma L_{-1}([L_1, L_{-1}] + \underbrace{L_{-1}L_1}_{=0 \text{ on } |\chi_2\rangle}))|\chi_2\rangle \\
&= (3L_{-1} + 2\gamma([L_0, L_{-1}] + L_{-1}L_0) + 2\gamma L_{-1}L_0)|\chi_2\rangle \\
&= (3L_{-1} + 2\gamma L_{-1} + 4\gamma L_{-1}L_0)|\chi_2\rangle \\
&= (3 + 2\gamma + 4\gamma(a-2))L_{-1}|\chi_2\rangle \stackrel{!}{=} 0 \\
\Leftrightarrow \quad \gamma &= \frac{3}{6-4a}\,,
\end{aligned}
\tag{3.37}
$$

and

$$
\begin{aligned}
L_2|s\rangle &= ([L_2, L_{-2}] + \overbrace{L_{-2}L_2}^{=0 \text{ on } |\chi_2\rangle} + \gamma(\overbrace{[L_2, L_{-1}]}^{=3L_1} + L_{-1}L_2)L_{-1})|\chi_2\rangle \\
&= (4L_0 + \tfrac{D}{2} + 3\gamma L_1 L_{-1} + \gamma L_{-1}(\underbrace{[L_2, L_{-1}]}_{=3L_1=0 \text{ on } |\chi_2\rangle} + \underbrace{L_{-1}L_2}_{=0 \text{ on } |\chi_2\rangle}))|\chi_2\rangle \\
&= (4L_0 + \tfrac{D}{2} + 3\gamma(2L_0 + L_{-1}L_1))|\chi_2\rangle \\
&= ((4 + 6\gamma)(a-2) + \tfrac{D}{2})|\chi_2\rangle \\
\Leftrightarrow \quad D &= (8 + 12\gamma)(2-a)\,.
\end{aligned}
\tag{3.38}
$$

From these, we deduce that when $a = 1$ (to impose physicality at level-one), then the above states are only consistently null if $\gamma = 3/2$, which then implies the critical dimension $D = 26$.

This construction generalizes to arbitrary level, provided $D = 26$. So the critical dimension is to ensure that all states associated residual conformal transformations generated by L_{-m} indeed decouple.

Unfortunately, this argument is not quite enough to rule out consistency in $D < 26$: In these cases, the above states would be orthogonal to all physical states, but they cannot be made physical themselves. So a priori, there is no contradiction to finding them in the Hilbert space. As mentioned before, the inconsistencies for $D < 26$ only arise in string loops; and so far we have only discussed strings at "tree level".

3.3 Closed strings and spacetime gravity

Let us now inspect the low-level physical states of the closed string in the critical setting, $a = 1$ and $D = 26$. The physicality conditions are now imposed by both "left-" and "right-moving" Virasoro generators, $\tilde{L}_{m\geq 0}$ and $L_{m\geq 0}$. Recall that the $m = 0$ constraints can be recast as level matching, $N = \tilde{N}$, and mass shell condition, $M^2 = -k^2 = (2(N+\tilde{N}) - 4a)/\alpha'$. Level matching allows us to organize the *physical* states by a single level number, either the left- or the right-moving one.

At level zero, we just have the oscillator vacua $|0, \tilde{0}; k\rangle$,[4] which is still tachyonic, since $M^2 = -k^2 = -4a/\alpha' = 4/\alpha'$ for $a = 1$. Note that this is four times the mass-squared of the open string ground state; or, closed string tachyon is twice as "massive" as the open string tachyon.

At level one, the level matching condition requires one excitation in the left- and right-moving sector each. So the general state takes the form

$$|\Omega; k\rangle := \Omega_{\mu\nu}\alpha^{\mu}_{-1}\tilde{\alpha}^{\nu}_{-1}|0, \tilde{0}; k\rangle\,. \tag{3.39}$$

The mass shell condition then becomes $M^2 = -k^2 = (2(1+1) - 4a)/\alpha' = 0$ in the critical setting. So all these states are massless in spacetime. The $L_1/\tilde{L}_1$ constraints are then

$$\begin{aligned} L_1|\Omega; k\rangle &= L_1\Omega_{\mu\nu}\alpha^{\mu}_{-1}\tilde{\alpha}^{\nu}_{-1}|0, \tilde{0}; k\rangle = \Omega_{\mu\nu}[L_1, \alpha^{\mu}_{-1}]\tilde{\alpha}^{\nu}_{-1}|0, \tilde{0}; k\rangle \\ &= \Omega_{\mu\nu}\alpha^{\mu}_{0}\tilde{\alpha}^{\nu}_{-1}|0, \tilde{0}; k\rangle = \sqrt{\tfrac{\alpha'}{2}}\Omega_{\mu\nu}k^{\mu}\tilde{\alpha}^{\nu}_{-1}|0, \tilde{0}; k\rangle \stackrel{!}{=} 0\,, \\ \tilde{L}_1|\Omega; k\rangle &= \sqrt{\tfrac{\alpha'}{2}}\Omega_{\mu\nu}k^{\nu}\alpha^{\mu}_{-1}|0, \tilde{0}; k\rangle \stackrel{!}{=} 0\,, \\ \Longrightarrow \qquad \Omega_{\mu\nu}k^{\mu} &= 0 \quad \text{and} \quad \Omega_{\mu\nu}k^{\nu} = 0\,. \end{aligned} \tag{3.40}$$

This is the condition for a 2-tensor to be transverse to the spacetime momentum k. In the spacetime interpretation, we can think of $|\Omega; k\rangle$ as the Fourier modes (labelled by the momentum k) of a tensor field $\Omega_{\mu\nu}(x)$.

[4] Previously, when we wrote $\mathcal{H}^{\text{Fock}}_{\text{closed}} \cong \mathcal{H}^{\text{Fock}}_{\text{open}} \otimes \mathcal{H}^{\text{Fock}}_{\text{open}}$, we were slightly careless about the ground state. In fact, we should have identified the two oscillator vacua on the right-hand side, which amounts to modding out the tensor product by an equivalence relation. This gives a type of ground states for the closed string, which we denote by $|0, \tilde{0}; k\rangle$ here, to distinguish from the open string ground states.

There is more to be said about the properties of this field, due to the existence of null states at level one. Consider

$$
\begin{aligned}
L_{-1}(\zeta\cdot\tilde{\alpha}_{-1}|0,\tilde{0};k\rangle) &= (\zeta\cdot\tilde{\alpha}_{-1})L_{-1}|0,\tilde{0};k\rangle = (\zeta\cdot\tilde{\alpha}_{-1})(\alpha_{-1}\cdot\alpha_0)|0,\tilde{0};k\rangle \\
&= \sqrt{\tfrac{\alpha'}{2}}(\zeta\cdot\tilde{\alpha}_{-1})(k\cdot\alpha_{-1})|0,\tilde{0};k\rangle = \sqrt{\tfrac{\alpha'}{2}}\,\underbrace{k_\mu\zeta_\nu}_{=\omega_{\mu\nu}}\,\alpha^\mu_{-1}\tilde{\alpha}^\nu_{-1}|0,\tilde{0};k\rangle\,.
\end{aligned}
\tag{3.41}
$$

By construction, this state is orthogonal to all physical states, and, analogously to above, is physical if $k^2 = 0$ (which implies $\omega_{\mu\nu}k^\mu = 0$) and

$$
\omega_{\mu\nu}k^\nu = k_\mu\zeta_\nu k^\nu \overset{!}{=} 0 \quad\Leftrightarrow\quad \zeta\cdot k = 0\,. \tag{3.42}
$$

Hence, it is a null state. Likewise,

$$
\tilde{L}_{-1}(\zeta'\cdot\alpha_{-1}|0,\tilde{0};k\rangle) = \sqrt{\tfrac{\alpha'}{2}}\,\underbrace{\zeta'_\mu k_\nu}_{\omega'_{\mu\nu}}\,\alpha^\mu_{-1}\tilde{\alpha}^\nu_{-1}|0,\tilde{0};k\rangle \tag{3.43}
$$

is a null state if $\zeta'\cdot k = 0$. Therefore, the level one states

$$
|\Omega_{\mu\nu};k\rangle \sim |\Omega_{\mu\nu} + k_\mu\zeta_\nu + k_\nu\zeta'_\mu;k\rangle = |\Omega_{\mu\nu};k\rangle + \sqrt{\tfrac{2}{\alpha'}}(|k_\mu\zeta_\nu;k\rangle + |k_\nu\zeta'_\mu;k\rangle) \tag{3.44}
$$

are equivalent as physical states if $k\cdot\zeta = 0 = k\cdot\zeta'$. In other words, the tensor $\Omega_{\mu\nu}$ enjoys a "spacetime gauge symmetry" $\Omega_{\mu\nu} \to \Omega_{\mu\nu} + k_\mu\zeta_\nu + k_\nu\zeta'_\mu$.

To make the spacetime interpretation more transparent, it is useful to recall the decomposition of a general 2-tensor Ω into irreducible representations of the spacetime Lorentz symmetry:

$$
\Omega_{\mu\nu} = \underbrace{\gamma_{(\mu\nu)}}_{\substack{\text{symmetric,}\\ \text{traceless}}} + \underbrace{b_{[\mu\nu]}}_{\text{anti-sym.}} + \underbrace{\varphi}_{\text{trace}}\,\eta_{\mu\nu}\,. \tag{3.45}
$$

Then, the physical states associated with $\gamma_{(\mu\nu)}$ are transverse (to k) symmetric traceless tensors that enjoy the gauge invariance

$$
\gamma_{\mu\nu} \to \gamma_{\mu\nu} + k_\mu\zeta_\nu + k_\nu\zeta_\mu \tag{3.46}
$$

with $\zeta \cdot k = 0$ (so $\zeta' = \zeta$ in the general parametrization above). In the spacetime interpretation, this is exactly the same degrees of freedom as a *graviton* in traceless harmonic gauge:

$$g_{\mu\nu}(x) = \eta_{\mu\nu} + \gamma_{\mu\nu}(x) , \quad \text{with} \quad \gamma_{\mu\nu}(x) \sim \gamma_{\mu\nu}(x) + \partial_\mu \zeta_\nu(x) + \partial_\nu \zeta_\mu(x) , \quad (3.47)$$

where ζ parametrizes the gauge transformations coming from diffeomorphisms. In momentum space, this transformation evidently becomes $\gamma_{\mu\nu}(k) \sim \gamma_{\mu\nu}(k) + k_\mu \zeta_\nu + k_\nu \zeta_\mu$, with $k \cdot \zeta = 0$ being the gauge fixing condition. This result suggests that a quantum theory of strings automatically includes gravity.

Analogously, the anti-symmetric part (transverse to k) also enjoys a gauge symmetry (with $\zeta' = -\zeta$),

$$b_{\mu\nu} \longrightarrow b_{\mu\nu} + k_\mu \zeta_\nu - k_\nu \zeta_\mu \, . \quad (3.48)$$

Moreover, notice that the gauge parameter ζ itself, which is constrained by $\zeta \cdot k = 0$, has a redundancy in $\zeta \longrightarrow \zeta + \lambda k$. Including both, we recognize that $b_{\mu\nu}$ has the spacetime interpretation of a *2-form gauge field*,

$$b_{\mu\nu}(x) dx^\mu \wedge dx^\nu \equiv b^{(2)} \sim b^{(2)} + d\zeta^{(1)} \quad (\zeta^{(1)} \sim \zeta^{(1)} + d\lambda) \, . \quad (3.49)$$

Lastly, we turn to the trace part, which ought to be a spacetime scalar field. The Virasoro constraints seem to eliminate all non-zero momentum modes,

$$\begin{aligned} L_1(\varphi\, \eta_{\mu\nu} \alpha^\mu_{-1} \tilde{\alpha}^\nu_{-1} |0, \tilde{0}; k\rangle) &= \varphi\, \eta_{\mu\nu} \tilde{\alpha}^\nu_{-1} ([L_1, \alpha^\mu_{-1}] + \alpha^\mu_{-1} L_1) |0, \tilde{0}; k\rangle \\ &= \varphi\, \eta_{\mu\nu} \tilde{\alpha}^\nu_{-1} \alpha^\mu_0 |0, \tilde{0}; k\rangle = \varphi \sqrt{\tfrac{\alpha'}{2}} k_\nu \tilde{\alpha}^\nu_{-1} |0, \tilde{0}; k\rangle \stackrel{!}{=} 0 \, , \\ \tilde{L}_1(\varphi\, \eta_{\mu\nu} \alpha^\mu_{-1} \tilde{\alpha}^\nu_{-1} |0, \tilde{0}; k\rangle) &= \varphi \sqrt{\tfrac{\alpha'}{2}} k_\mu \alpha^\mu_{-1} |0, \tilde{0}; k\rangle \stackrel{!}{=} 0 \\ \Longrightarrow \ k^\mu &= 0 \, , \end{aligned} \quad (3.50)$$

which would force the field to be constant. So there would be no degree of freedom.

To see that the level-one states truly contain a spacetime scalar field, we define the state

$$\begin{aligned} |\varphi_{\rho,\tilde{\rho}}; k\rangle &:= [(\rho \cdot \alpha_{-1})(\tilde{\alpha}_0 \cdot \tilde{\alpha}_{-1}) + (\alpha_0 \cdot \alpha_{-1})(\tilde{\rho} \cdot \tilde{\alpha}_{-1}) + \varphi\, \alpha_{-1} \cdot \tilde{\alpha}_{-1}] \, |0, \tilde{0}; k\rangle \\ &= \left[(\rho \cdot \alpha_{-1}) \tilde{L}_{-1} + (\tilde{\rho} \cdot \tilde{\alpha}_{-1}) L_{-1} + \varphi\, \alpha_{-1} \cdot \tilde{\alpha}_{-1}\right] |0, \tilde{0}; k\rangle \, . \end{aligned} \quad (3.51)$$

To make it physical, we need $k^2 = 0$ to satisfy the $L_0/\tilde{L}_0$ constraints. The evaluate the $L_1/\tilde{L}_1$ constraints, it is useful to compute their actions on each of the three terms separately. For the first, we find

$$\begin{aligned}
\tilde{L}_1\left[(\rho\cdot\alpha_{-1})\tilde{L}_{-1}|0,\tilde{0};k\rangle\right] &\propto \tilde{L}_1\tilde{L}_{-1}|0,\tilde{0};k\rangle = 2L_0|0,\tilde{0};k\rangle \propto k^2 = 0\\
L_1\left[(\rho\cdot\alpha_{-1})\tilde{L}_{-1}|0,\tilde{0};k\rangle\right] &= \tilde{L}_{-1}\left[\rho_\mu([L_1,\alpha^\mu_{-1}] + \alpha^\mu_{-1}L_1)|0,\tilde{0};k\rangle\right]\\
&= \tilde{L}_{-1}\rho_\mu\alpha^\mu_0|0,\tilde{0};k\rangle = \sqrt{\tfrac{\alpha'}{2}}(\rho\cdot k)(\tilde{\alpha}_{-1}\cdot\tilde{\alpha}_0)|0,\tilde{0};k\rangle\\
&= \sqrt{\tfrac{\alpha'}{2}}(\rho\cdot k)\left(\sqrt{\tfrac{\alpha'}{2}}k_\mu\tilde{\alpha}^\mu_{-1}\right)|0,\tilde{0};k\rangle\,.
\end{aligned} \tag{3.52}$$

For the second, the analogous computation shows

$$\begin{aligned}
L_1\left[(\tilde{\rho}\cdot\tilde{\alpha}_{-1})L_{-1}|0,\tilde{0};k\rangle\right] &= 0\\
\tilde{L}_1\left[(\tilde{\rho}\cdot\tilde{\alpha}_{-1})L_{-1}|0,\tilde{0};k\rangle\right] &= \sqrt{\tfrac{\alpha'}{2}}(\tilde{\rho}\cdot k)\left(\sqrt{\tfrac{\alpha'}{2}}k_\mu\alpha^\mu_{-1}\right)|0,\tilde{0};k\rangle\,.
\end{aligned} \tag{3.53}$$

The third term was already computed above in (3.50). Therefore, the Virasoro constraint for the state (3.51) leads to

$$\begin{aligned}
L_1|\varphi_{\rho,\tilde{\rho}};k\rangle &= \left[\sqrt{\tfrac{\alpha'}{2}}(\rho\cdot k)\left(\sqrt{\tfrac{\alpha'}{2}}k_\mu\tilde{\alpha}^\mu_{-1}\right) + \varphi\left(\sqrt{\tfrac{\alpha'}{2}}k_\mu\tilde{\alpha}^\mu_{-1}\right)\right]|0,\tilde{0};k\rangle \stackrel{!}{=} 0\,,\\
\tilde{L}_1|\varphi_{\rho,\tilde{\rho}};k\rangle &= \left[\sqrt{\tfrac{\alpha'}{2}}(\tilde{\rho}\cdot k)\left(\sqrt{\tfrac{\alpha'}{2}}k_\mu\alpha^\mu_{-1}\right) + \varphi\left(\sqrt{\tfrac{\alpha'}{2}}k_\mu\alpha^\mu_{-1}\right)\right]|0,\tilde{0};k\rangle \stackrel{!}{=} 0\,,
\end{aligned} \tag{3.54}$$

from which we conclude the physicality condition $\rho\cdot k = \tilde{\rho}\cdot k = -\varphi\sqrt{\frac{2}{\alpha'}}$.

Naively, these are too many degrees of freedom for a spacetime scalar: we have two vectors and one scalar, related by with two constraints, so $2D-1$ free parameters for $|\varphi_{\rho,\tilde{\rho}};k\rangle$. However, we can still utilize the existence of null states, to identify most of these degrees of freedom as redundant! Indeed, looking at the ansatz (3.51), we immediately see that the first two terms are of the form $\tilde{L}_{-1}|\rho\rangle$ and $L_{-1}|\tilde{\rho}\rangle$, which are orthogonal to all physical states. So it is tempting to remove these by subtracting an appropriate null state.

The null states in question are schematically of the same form,

$$|\rho_n;k\rangle := \tilde{L}_{-1}\left[\rho_n\cdot\alpha_{-1}|0,\tilde{0};k\rangle\right]\,,\quad |\tilde{\rho}_n;k\rangle := L_{-1}\left[\tilde{\rho}_n\cdot\tilde{\alpha}_{-1}|0,\tilde{0};k\rangle\right]\,, \tag{3.55}$$

but in order to be null (hence the subscript n), they need $k^2 = 0$ for the $L_0/\tilde{L}_0$ constraints, and $\rho_n \cdot k = \tilde{\rho}_n \cdot k = 0$ for the $L_1/\tilde{L}_1$ constraints (which follow in an obvious manner from (3.52) and (3.53)). These are $2D-2$ parameters that are unphysical, or "pure gauge", which can be used to always eliminate all but one degree of freedom in $|\varphi_{\rho,\tilde{\rho}}; k\rangle$.[5]

To summarize, we have seen that the critical closed bosonic string gives rise to massless spacetime tensor fields at level one. Importantly, the spectrum of null states behave just in the right way to describe, in the spacetime interpretation, the following objects:

- A symmetric traceless tensor $\gamma_{\mu\nu}$ with gauge symmetry $\gamma_{\mu\nu} \to \gamma_{\mu\nu} + \partial_\mu \zeta_\nu + \partial_\nu \zeta_\mu$; there are classic arguments that such a field *must be* metric perturbations, i.e., *dynamical gravity*.
- An anti-symmetric tensor $b_{\mu\nu}$ with gauge symmetries $b_{\mu\nu} \to b_{\mu\nu} + \partial_\mu \zeta_\nu - \partial_\nu \zeta_\mu$ and $\zeta_\mu \to \zeta_\mu + \partial_\mu \lambda$; this means that b is a differential form of degree 2, or a *2-form gauge field*. In the string theory context, this field is known as the "Kalb–Ramond" field.
- A scalar φ; this field is called the *dilaton*.

In String Theory II, you will learn that these spacetime fields persist for the superstring, which has the additional benefit of having no tachyonic ground state.

The emergence of the symmetric tensor from quantized strings is one of the reasons why string theory is generally said to be a quantum theory of gravity. However, the other fields are just as important to furnish a *consistent* quantum theory of gravity. We will come back to some basic aspects of these fields later on.

4 Scattering of Strings

So far, all discussion has been limited to the quantization of the Polyakov action (2.30), which describes a free field theory in $(1+1)$ dimensions. To describe

[5] To see this directly, first note that $|\varphi_{\rho,\tilde{\rho}}; k\rangle - |\rho_n; k\rangle = |\varphi_{\rho-\rho_n,\tilde{\rho}}; k\rangle$. Now imagine picking spacetime coordinates such that the null momentum k is $(t, t, 0, 0, \ldots)$. Physicality of $|\varphi_{\rho,\tilde{\rho}}; k\rangle$ means $\rho = \left(\sqrt{\frac{\alpha'}{2}}\frac{\varphi}{t} + \rho^1, \rho^1, \rho^2, \ldots\right)$. For $|\rho_n; k\rangle$ we can then pick $\rho_n = (\rho^1, \rho^1, \rho^2, \rho^3, \ldots)$ which obviously satisfies $\rho_n \cdot k = 0$. This leaves $\rho - \rho_n = \left(\sqrt{\frac{\alpha'}{2}}\frac{\varphi}{t}, 0, 0, \ldots\right)$, which depends only on φ (and the momentum). Analogously, we can remove any free parameters in $\tilde{\rho}$.

interactions, the standard procedure would be to add interaction terms to this action, then utilize the LSZ formalism to compute scattering amplitudes via correlation functions. However, in the case of the worldsheet theory, it is impossible to add any (standard) interaction terms without breaking conformal invariance. As we have seen previously (though in somewhat hidden way), having conformal invariance as a gauge symmetry is crucial for the free theory to make sense; if broken, any hope of doing perturbation theory around the free limit in a controllable manner is lost. Miraculously, even though conformal symmetry seems to be forbidding any QFT-esque interactions, it is also what makes *stringy interactions* possible.

The idea behind this goes back to the worldline formulation of relativistic quantum mechanics and its interpretation inside a path-integral. There, interactions are encoded in the splitting and joining of worldlines at "interaction vertices", and the resulting diagrams are (the position space version of) Feynman diagrams. Fixing the rules of joining worldlines at vertices, which is equivalent to specifying the details of the interactions, it is well-known that by summing all possible Feynman diagrams with fixed external legs (and integrating over the positions of all vertices), one recovers an asymptotic series expression for the scattering amplitude.

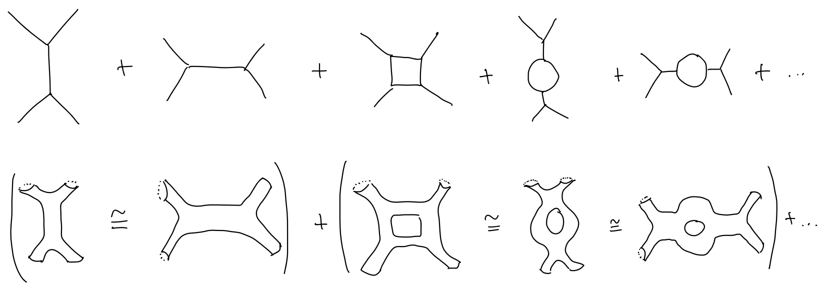


Fig. 3: Feynman diagrams in QFT and their stringy version. The diagrams can be thought of representing open and closed strings; for the open string, the solid lines are all boundaries, while for the closed string, only the (obvious) ends are boundaries of the worldsheet. Notice how different particle-diagrams become topologically equivalent string diagrams.

For the worldsheet theory, this gives an obvious generalization: we can draw all possible diagrams of strings splitting and joining, with some chosen physical "in and out" states on the external legs, and formally sum over all of them, see Figure 3. By doing so, we immediately recognize a major difference to point particles and worldlines. Whereas the vertices joining worldlines are "localized singularities" of the Feynman diagram, any string diagram locally looks like a smooth 2d manifold. So locally, there is neither any chance, nor any *need*, to "add an interaction term".

But then, how do we capture the global structure of the string diagram that clearly exhibits some features of "interaction"? Here is the place where the magic of 2d conformal symmetry comes into play: any boundary component of a 2d manifold, or surface, can be mapped by conformal transformations to a *marked point*, or *puncture*, with suitable data that "keeps track" of the states on external leg.

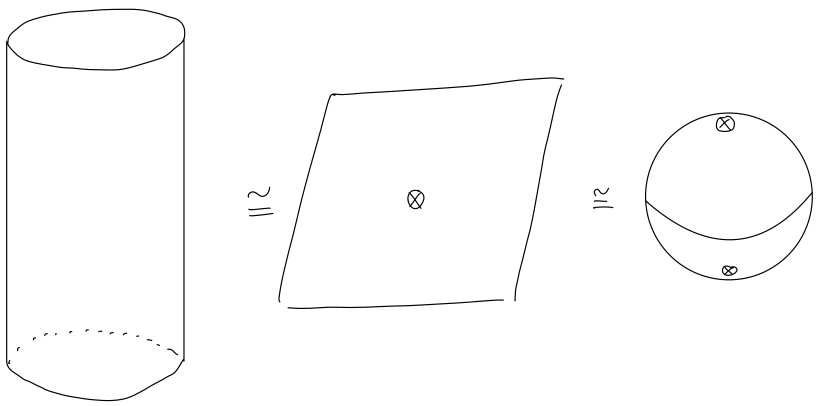


Fig. 4: A cylinder can be conformally mapped to the plane, with a puncture at origin representing one end. It can also be mapped to a sphere with two punctures at poles, which can be seen as the one-point compactification of the plane (that additional point being the second puncture for the other end of the cylinder).

As we will discuss in more detail below, the punctures (plus the physical data attached to them) correspond to the insertion of so-called **vertex operators**, which are certain local operators in the free worldsheet theory. Morally, they describe the absorption/emission of one physical string state, see Figure 5.

The "value" of each diagram is then a correlation function of these operators, evaluated according to the topology of the surface, see Figure 6. One important difference to point-particles is, as already highlighted in Figure 3, is that there are less distinct topologies, and hence string diagrams at given loop order than for Feynman diagrams. This is a very crude explanation for why string diagrams are "less" divergent than (unrenormalized) Feynman diagrams. In many ways, the finite width of the string "resolves" some of the singularities, and provides a natural cutoff for momenta running in loops. In the early days of string theory as an attempt to describe the strong interaction, these properties correctly predicted some aspects of the physics of mesons and baryons, and has led to the development of the "dual resonance model".

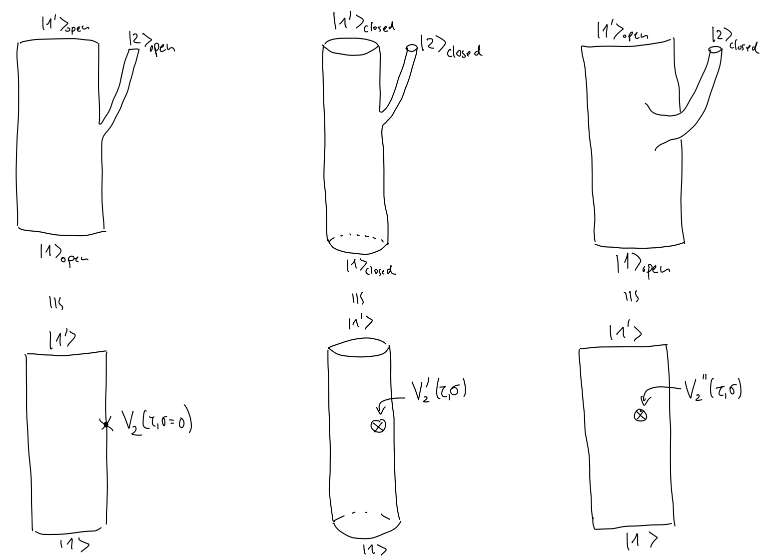


Fig. 5: Vertex operators describe the emission and absorption of physical states. Open string states are associated to vertex operator insertions on the boundary of the worldsheet, while closed string states correspond to insertions on the interior.

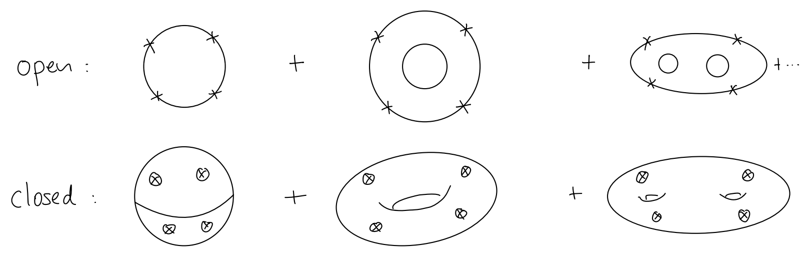


Fig. 6: Using conformal symmetry, scattering amplitudes of strings are expressed as an asymptotic series in topologies of 2d surfaces (genus expansion) with fixed punctures labelled by vertex operators that encode the external states.

4.1 Vertex operators

To make the diagrammatic notion of "emitting/absorbing" a state on the string worldsheet precise, note that any of the processes shown in figure 5 can be thought of as a linear transformation $|1\rangle \longrightarrow |1'\rangle$. Such a transformation should be implemented by a local operator acting on Hilbert space. How do we describe such an operator? Just as in standard QFT, a generic local operator $O(\xi)$ on the worldsheet can be build from the basic fields $X^\mu(\xi)$ and their conjugate momenta $\Pi^\mu(\xi)$. But in the worldsheet theory, associating an operator to any specific point on Σ would break diffeomorphism invariance. To remedy this, we must integrate over the insertion locus of the local operator. More precisely, we have

$$\begin{aligned} &\text{closed string:} && \int_\Sigma d\tau d\sigma\, O(\tau,\sigma)\,, \\ &\text{(NN) open string:} && \int_{-\infty}^{\infty} d\tau\, O(\tau,\sigma=0 \text{ or } l)\,, \end{aligned} \tag{4.1}$$

where on the open string the operator is inserted on the boundary.

For such an operator insertion to describe a physical absorption/emission process, it is obvious that we need to map physical states to physical states, and, moreover, null states to null states. As in the previous section, this is implemented by the Virasoro generators. Focusing on the (NN) open string with operators inserted on the boundary, we have, for $m > 0$,

$$\begin{aligned} L_m\left[\int d\tau\, O(\tau,0)|\text{phys}\rangle\right] &= \int d\tau [L_m, O(\tau,0)]|\text{phys}\rangle \stackrel{!}{=} 0 \\ \Longleftrightarrow \quad [L_m, O(\tau,0)] &\propto \partial_\tau(\ldots)\,, \\ \int d\tau\, O(\tau,0)\,(L_{-m}|\phi\rangle) &= \int d\tau\,([O(\tau,0),L_{-m}]|\phi\rangle + L_{-m}O(\tau,0)|\phi\rangle) \\ \Longleftrightarrow \quad [O(\tau,0),L_{-m}] &\propto \partial_\tau(\ldots) \quad (\text{if } L_{-m}|\phi\rangle \text{ is physical})\,. \end{aligned} \tag{4.2}$$

These conditions specify the transformation behavior of $O(\tau,\sigma)$ under infinitesimal conformal transformations.

More generally, in a conformally invariant theory, one defines so-called *primary operators* $A(\tau)$ *of weight* h if under a *conformal* transformation $\tau \to \tilde{\tau}(\tau)$, one has

$$A(\tau) \to \tilde{A}(\tilde{\tau}) = A(\tau)\left(\frac{\partial\tau}{\partial\tilde{\tau}}\right)^h\,. \tag{4.3}$$

Clearly, the integrated operator $\int d\tilde{\tau}\tilde{A}(\tilde{\tau}) = \int d\tilde{\tau}\left(\frac{\partial\tau}{\partial\tilde{\tau}}\right)^h A(\tau)$ is invariant if $h = 1$.[6]

Infinitesimally, $\tau \longrightarrow \tilde{\tau} = \tau + \epsilon(\tau)$, the transformation of a primary operator implies

$$\begin{aligned}\tilde{A}(\tilde{\tau}) &= A(\tau)\left(\frac{\partial\tau}{\partial\tilde{\tau}}\right)^h = A(\tau)\left(\frac{\partial\tilde{\tau}}{\partial\tau}\right)^{-h} = A(\tau)(1 + \partial_\tau\epsilon(\tau) + O(\epsilon^2))^{-h} \\ &= A(\tau)(1 - h\partial_\tau\epsilon(\tau)) + O(\epsilon^2)\,.\end{aligned} \tag{4.4}$$

On the other hand, we also have

$$\begin{aligned}\tilde{A}(\tilde{\tau}) &= \tilde{A}(\tau + \epsilon(\tau)) = \tilde{A}(\tau) + \epsilon(\tau)\partial_\tau\tilde{A}(\tau) + O(\epsilon^2) \\ &= \tilde{A}(\tau) + \epsilon(\tau)\partial_\tau\tilde{A}(\tilde{\tau} - \epsilon(\tau)) + O(\epsilon^2) \\ &= \tilde{A}(\tau) + \epsilon(\tau)\partial_\tau\left[\tilde{A}(\tilde{\tau}) - \epsilon(\tau)\partial_{\tilde{\tau}}\tilde{A}(\tilde{\tau}) + O(\epsilon^2)\right] + O(\epsilon)^2 \\ &= \tilde{A}(\tau) + \epsilon(\tau)\partial_\tau A(\tau) + O(\epsilon^2)\,,\end{aligned} \tag{4.5}$$

where we also make use of the transformation properties of A as a primary operator. This means that

$$\begin{aligned}\delta_\epsilon A(\tau) = \tilde{A}(\tau) - A(\tau) &= -\epsilon(\tau)\partial_\tau A(\tau) - hA(\tau)\partial_\tau\epsilon(\tau) \\ &= -\partial_\tau(A(\tau)\epsilon(\tau)) - (h-1)A(\tau)\partial_\tau\epsilon(\tau)\,,\end{aligned} \tag{4.6}$$

which is a total derivative for $h = 1$.

Since this is the variation of $A(\tau)$ with respect to the infinitesimal conformal transformation $\tau \longrightarrow \tilde{\tau} = \tau + \epsilon(\tau)$, it also encodes the *classical* Poisson bracket action, $\{L_m, A(\tau)\}_{\mathrm{PB}} = \delta_{\epsilon_m}A(\tau)$, of the Virasoro generators L_m, which are parametrized by the conformal Killing vectors $\epsilon_m = -e^{im\tau}$.[7] The quantum commutator expression is then obtained via "$[\cdot,\cdot] \cong -i\{\cdot,\cdot\}_{\mathrm{PB}}$", which is

$$[L_m, A(\tau)] = e^{im\tau}(-i\partial_\tau + mh)A(\tau)\,, \tag{4.7}$$

[6] This is essentially just a definition of "primary" in the classical setting. There are subtleties arising in the quantization, and will be discussed in detail in the CFT course.

[7] We are at $\sigma = 0$, so $\xi^\pm(\sigma = 0) = \tau$. Moreover, we will from now on drop the length scale l on the worldsheet, which means that σ is a dimensionless quantity that runs from 0 to 2π on the closed string, and to π on the open string. This simplifies the expressions in Section 2, by setting $2\pi/l$ (in expressions for the closed string) and π/l (in expressions for the open string) to 1. However, having the explicit worldsheet scale l is useful when discussing, e.g., the Casimir energy and its relationship to the normal ordering constant.

which can be seen as an equivalent definition of weight h primaries in the open string quantum theory. Sometimes h is also called the conformal dimension. A **vertex operator** is then a primary operator of weight $h = 1$.

The story for the closed string is analogous when we are in lightcone coordinates. Here, a primary operator is labelled by two weights, $(\bar{h}, h)$, and is characterized by

$$A(\xi^+, \xi^-) \to \tilde{A}(\tilde{\xi}^+, \tilde{\xi}^-) = \left(\frac{\partial \xi^+}{\partial \tilde{\xi}^+}\right)^{\bar{h}} \left(\frac{\partial \xi^-}{\partial \tilde{\xi}^-}\right)^{h} A(\xi^+, \xi^-)\,, \tag{4.8}$$

whose infinitesimal transformation under $\tilde{\xi}^\pm = \xi^\pm + \epsilon^\pm$ is again a total derivative if $h = \bar{h} = 1$:

$$\delta A = -\partial_+(\epsilon^+ A) - (\bar{h} - 1)A\partial_+\epsilon^+ - \partial_-(\epsilon^- A) - (h - 1)A(\partial_-\epsilon)\,. \tag{4.9}$$

For $\epsilon^\pm = \frac{i}{2} e^{2im\xi^\pm}$ (we include the factor of $(-i)$ so that the expression directly matches the quantum commutators), this gives the action of the Virasoro generators:

$$\begin{aligned} [L_m, A(\xi^\pm)] &= \tfrac{1}{2} e^{2im\xi^-}(-i\partial_- + 2mh)A(\xi^\pm)\,, \\ [\tilde{L}_m, A(\xi^\pm)] &= \tfrac{1}{2} e^{2im\xi^+}(-i\partial_+ + 2m\bar{h})A(\xi^\pm)\,. \end{aligned} \tag{4.10}$$

For now, let us restrict ourselves to the (NN) open string, and get some intuition for (boundary) vertex operators. We start with the basic scalar field $X^\mu(\tau, 0) \equiv X^\mu(\tau)$. To check its conformal dimension, we simply compute the action of the Virasoro generators:

$$\begin{aligned} [L_m, X^\mu(\tau)] &= \tfrac{1}{2} \sum_n \left[\alpha_{m-n} \cdot \alpha_n\,,\; x^\mu + 2\alpha' p^\mu \tau + i\sqrt{2\alpha'} \sum_{\ell \neq 0} \frac{\alpha^\mu_\ell}{\ell} e^{-i\ell\tau} \right] \\ &= -i\sqrt{2\alpha'} \sum_n \alpha^\mu_n e^{-i(n-m)\tau} = -i e^{im\tau} \frac{\partial}{\partial \tau}(X^\mu(\tau))\,, \end{aligned} \tag{4.11}$$

implying $h = 0$, and so it is not a vertex operator.

It turns out that the simplest vertex operator built from X^μ is of the form $\exp(ik_\mu X^\mu(\tau)) = \exp(ik \cdot X(\tau)$. Notice that, from a spacetime perspective, the commutator $[P^\mu, \exp(ik \cdot X(\tau))] = k^\mu \exp(ik \cdot X(\tau))$ suggests the interpretation that this operator increases the spacetime momentum by k, just as colliding an incoming state with another one that has momentum k, see Figure 7.

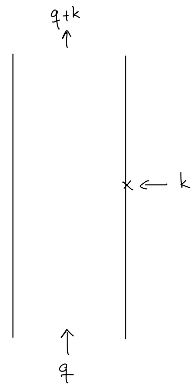


Fig. 7: Inserting a vertex operator of the schematic form $\exp(ik \cdot X)$ describes an incoming state with momentum k.

One important subtlety is operator ordering in taking products of X^μ. As usual, we work with the *normal ordered* product

$$V_k(\tau) := :e^{ik\cdot X(\tau)}: \\ = \exp\left(\sqrt{2\alpha'}\sum_{n>0}\frac{k\cdot\alpha_{-n}}{n}e^{in\tau}\right)e^{ik\cdot(x+2\alpha' p\tau)}\exp\left(-\sqrt{2\alpha'}\sum_{n>0}\frac{k\cdot\alpha_n}{n}e^{-in\tau}\right). \tag{4.12}$$

There could in general be some normal ordering constants appearing here, stemming from the fact that

$$\exp(ik\cdot\alpha_m)\exp(ik\cdot\alpha_n) = \exp(ik\cdot(\alpha_m+\alpha_n) - \tfrac{1}{2}(k\cdot k)\delta_{m+n,0})\,; \tag{4.13}$$

notice that any such ambiguity vanishes for $k^2 = 0$. On the Problem Sheet, you will show that normal ordering in the definition of $V_k(\tau)$ gives the conformal transformation

$$[L_m, :e^{ik\cdot X(\tau)}:] = e^{im\tau}\left(-i\partial_\tau + m\,(\alpha' k\cdot k)\right) :e^{ik\cdot X(\tau)}: \,. \tag{4.14}$$

Therefore, $V_k(\tau)$ is a vertex operator if $\alpha' k^2 = 1$, which is precisely the mass-shell condition for the tachyon! In fact, this is the unique vertex operator that describes the absorption/emission of a tachyon; more on that later.

For level-one states, the mass-shell condition requires $k^2 = 0$. So the associated operator $V_k(\tau)$ has weight $h = 0$. To construct a vertex operator, i.e., a primary of weight $h = 1$, we make use of

$$\begin{aligned}
[L_m, \partial_\tau V_k(\tau)] = \partial_\tau [L_m, V_k(\tau)] &= \partial_\tau \left(e^{im\tau}(-i\partial_\tau + mh) V_k(\tau) \right) \\
&\overset{h=0}{=} ime^{im\tau}(-i\partial_\tau) V_k(\tau) + e^{im\tau}(-i\partial_\tau^2) V_k(\tau) \\
&= e^{im\tau}(-i\partial_\tau + m)\partial_\tau V_k(\tau) ,
\end{aligned} \tag{4.15}$$

to identify $\partial_\tau V_k(\tau) = ik \cdot \partial_\tau X :e^{ik\cdot X}:$ as a candidate. Since $k \cdot \partial_\tau X = k_\mu \, \partial_\tau X^\mu \equiv k_\mu \dot{X}^\mu$ is really a sum of individual terms, it is not hard to see (and also not hard to verify) that, in fact, the primary operator $\dot{X}^\mu : e^{ik\cdot X}:$ also has weight $h = 1$. This allows us to construct the vertex operator

$$W_{\zeta,k}(\tau) := \sqrt{\tfrac{1}{2\alpha'}} (\zeta \cdot \dot{X}(\tau)) :e^{ik\cdot X(\tau)}: , \tag{4.16}$$

where ζ can be interpreted as the polarization of the vector boson at level one, whose absorption/emission is controlled by $W_{\zeta,k}(\tau)$. One may wonder about the normal ordering between $\zeta \cdot \dot{X}$ and the exponential factor. Recall from (2.45) the mode expansion (with $\pi/l \cong 1$)

$$\dot{X}^\mu(\tau) \equiv \partial_\tau X^\mu|_{\sigma=0} = (\partial_+ + \partial_-) X^\mu|_{\sigma=0} = \sqrt{2\alpha'} \sum_n \alpha_n^\mu e^{-in\tau} , \tag{4.17}$$

from which it is not hard to see that all potential ambiguities come from commutators of the form

$$[\alpha_m \cdot \zeta, \alpha_n \cdot k] \propto \delta_{m+n,0} \, (\zeta \cdot k) . \tag{4.18}$$

This precisely vanishes for physical states $|\zeta; k\rangle$ at level one, where we have $\zeta \cdot k = 0$. Since these are all physical states at level one, we have constructed a vertex operator $W_{\zeta,k}(\tau)$ for every one of them. Notice that for $\zeta \propto k$, i.e., the longitudinal modes, we have

$$k \cdot \dot{X} e^{ik\cdot X} = -i\partial_\tau (e^{ik\cdot X}) , \tag{4.19}$$

which formally vanishes after integrating over τ.

We could continue with at level two, and show with more elaborate arguments that every physical state has a corresponding vertex operator. Of course, this continues to any level. Moreover, it turns out that also the converse is true: every vertex operator corresponds to a physical state.

4.2 The State-Operator Correspondence

The identification between physical states and vertex operators is a special feature of CFTs known as the state-operator correspondence: In a general 2d CFT, each primary operator A_j (of any weight) is in one-to-one correspondence to so-called highest weight states $|\phi_j\rangle \in \mathcal{H}$ of the Hilbert space; by acting on these with $L_{n<0}$, one obtains the entire Hilbert space. Focusing on the applications to the string worldsheet, we can further sharpen the statement:

There is a *one-to-one correspondence* between physical states and primary operators of weight $h = 1(= \bar{h})$, i.e., vertex operators.

To see how this comes about, let us consider, for concreteness, the action of the open string tachyon vertex operator (4.12) on the string vacuum. We can further use the Hamiltonian $H = \frac{\pi}{l} L_0 \equiv L_0 = \alpha' p^2 + N$ as the time evolution operator, to write

$$\begin{aligned} V_k(\tau)|0;0\rangle &= e^{i\tau L_0} :e^{ik\cdot X(0)}: \underbrace{e^{-i\tau L_0}|0;0\rangle}_{=|0;0\rangle} \\ &= e^{i\tau L_0} \exp\left(\sqrt{2\alpha'}\sum_{n>0}\frac{k\cdot\alpha_{-n}}{n}\right) e^{ik\cdot x}|0;0\rangle \\ &= e^{i\tau L_0} \exp\left(\sqrt{2\alpha'}\sum_{n>0}\frac{k\cdot\alpha_{-n}}{n}\right) |0;k\rangle \,. \end{aligned} \tag{4.20}$$

Now we perform a common method in field theory, namely Wick rotation, to pass to imaginary or Euclidean time $t = i\tau$ on the worldsheet. Then, using $\alpha' k^2 = 1$ for the tachyon, we have

$$\begin{aligned} &:e^{ik\cdot X(\tau)}: |0;0\rangle = e^{tL_0} \exp\left(\sqrt{2\alpha'}\sum_{n>0}\frac{k\cdot\alpha_{-n}}{n}\right)|0;k\rangle \\ =&e^{tL_0}\left[1 + \sqrt{2\alpha'}\left(k\cdot\alpha_{-1} + \frac{k\cdot\alpha_{-2}}{2} + \ldots\right) + \frac{2\alpha'}{2}\left((k\cdot\alpha_{-1})^2 + \ldots\right) + \ldots\right]|0;k\rangle \\ =&\left[e^t + e^{t(1+1)}\sqrt{2\alpha'}k\cdot\alpha_{-1} + e^{t(1+2)}\left(\sqrt{\tfrac{\alpha'}{2}}k\cdot\alpha_{-2} + \alpha'(k\cdot\alpha_{-1})^2\right)\right]|0;k\rangle + \ldots \end{aligned} \tag{4.21}$$

From this, we recover the tachyon $|0;k\rangle$ in the limit

$$\begin{aligned} |0;k\rangle &= \lim_{t\to-\infty} e^{-t} V_k(-it)|0;0\rangle \\ &\equiv \lim_{z\to 0} \frac{1}{z} V_k(-i\log z)|0;0\rangle\,, \end{aligned} \tag{4.22}$$

where, in the second line, we have defined $z = e^{i\tau} = e^t$.

The same procedure works for the "photon", i.e., with level-one states ($k^2 = 0$):

$$\begin{aligned} W_{\zeta,k}(\tau)|0;0\rangle &= \tfrac{1}{\sqrt{2\alpha'}}\zeta\cdot\dot{X}(\tau) :e^{ikX(\tau)}: |0;0\rangle \\ &= e^{tL_0}\sum_{n>0}(\zeta\cdot\alpha_{-n})\exp\left(\sqrt{2\alpha'}\sum_{m>0}\frac{k\cdot\alpha_{-m}}{m}\right)|0;k\rangle \\ &= \ldots = \left[z(\zeta\cdot\alpha_{-1}) + z^2(\zeta\cdot\alpha_{-2} + (\zeta\cdot\alpha_{-1})(k\cdot\alpha_{-1})) + \ldots\right]|0;k\rangle\,, \end{aligned} \tag{4.23}$$

from which we then again have

$$|\zeta;k\rangle = \lim_{z\to 0}\frac{1}{z}W_{\zeta,k}(-i\log z)|0;0\rangle\,. \tag{4.24}$$

It is not hard to guess the pattern for any physical state $|\psi\rangle$ with corresponding vertex operator $V_\psi(\tau)$:

$$|\psi\rangle = \lim_{z\to 0}\frac{1}{z}V_\psi(-i\log z)|0;0\rangle\,. \tag{4.25}$$

Analogously, each vertex operators also generates an "out" state. For the tachyon ($\alpha' k^2 = 1$), this would be

$$\begin{aligned} \langle 0;0|V_k(\tau) &= \langle 0;k|\exp\left(\sqrt{2\alpha'}\sum_{n>0}\frac{k\cdot\alpha_n}{n}\right)e^{-i\tau L_0} \\ &= \langle 0;k|\left[e^{-t} + e^{-2t}\sqrt{2\alpha'}k\cdot\alpha_1 + \ldots\right] \\ \implies\quad \langle 0;k| &= \lim_{t\to\infty} e^t\langle 0;0|V_k(-it) = \lim_{z\to\infty} z\langle 0;0|V_k(-i\log z)\,. \end{aligned} \tag{4.26}$$

Of course, there is an almost identical version for the closed string. Here, we just have to be slightly more careful about the spatial coordinate since we are inserting vertex operators on the interior of the worldsheet. Having Wick-rotated to Euclidean time, $\tau = -it$, the lightcone coordinates become $\xi^{\pm} = -i(t \pm i\sigma)$.[8] The map from vertex operators V_ψ (i.e., primaries of weight $h = 1 = \bar{h}$) to states $|\psi\rangle$ is then

$$\begin{aligned}|\psi\rangle &= \lim_{t\to-\infty} e^{-4t} V_\psi(-it,\sigma)|0,\tilde{0};0\rangle \\ &= \lim_{|z|\to 0} \frac{1}{(z\bar{z})^2} V_\psi(-\tfrac{i}{2}\log(z\bar{z}), -\tfrac{i}{2}\log\tfrac{z}{\bar{z}})|0,\tilde{0};0\rangle\,,\end{aligned} \tag{4.27}$$

with $z = e^{i\xi^+}$ and $\bar{z} = e^{i\xi^-}$.

Physically, what we have done is to time-evolve backwards the state generated by acting with the vertex operator $V_\psi(\tau)$ on the vacuum $|0;k\rangle$ to past infinity. In Euclidean time, the evolution to past infinity, $t \to -\infty$ by e^{tH}, suppresses higher energy-eigenstates, so the lowest-energy state dominates.

The coordinates $(z,\bar{z})$ further give a geometric interpretation. Take a closed string worldsheet that topologically it is a cylinder, so (τ,σ) are global coordinates on it. After Wick rotation, the Euclidean metric is $dt^2 + d\sigma^2 = d(i\xi^+)d(i\xi^-)$. Then the new coordinate $z = e^{i\xi^+} = e^t e^{i\sigma} \equiv |z|e^{i\sigma}$ has a natural interpretation as polar coordinates on the plane $\mathbb{R}^2 \cong \mathbb{C}$. Note that the standard metric on $\mathbb{C}$, $dzd\bar{z} = d(i\xi^+)d(i\xi^-)z\bar{z} = (dt^2 + d\sigma^2)e^{2t}$ is clearly conformally equivalent to the cylinder metric. This is precisely the first conformal transformation depicted in Figure 4. Then, the state/operator map makes the introductory statement about asymptotic states (initial state $|\psi\rangle$ prepared at $\tau = -it = -\infty$) being "generated" by vertex operator insertions/punctures precise.

4.3 Tree-level string amplitudes

3-point scattering

We now have the technical tools to describe 3-point tree-level string amplitudes. The simplest example is the scattering of three open string tachyons, $|0;k_i\rangle$, with $\alpha' k_i^2 = 1$. We have already seen its corresponding diagram, on the left of Figure 5. The amplitude of this scattering process is just the transition amplitude from, say, $|0;k_1\rangle$ to $|0;k_3\rangle$ via the absorption of $|0;k_3\rangle$, the latter of which is equivalently

[8] Implicitly we have also worked with these coordinates on the open string worldsheet above. However, we only cared about vertex operators inserted on the boundary $\sigma = 0$.

described by a vertex operator insertion. The naive guess of the amplitude is then (with k_3 incoming momentum)

$$\begin{aligned}
&\langle 0; -k_3| \int d\tau V_{k_2}(\tau)|0; k_1\rangle \\
&= \int d\tau \langle 0; -k_3|e^{i\tau L_0} V_{k_2}(0) e^{-i\tau L_0}|0; k_1\rangle \quad (L_0|0; k_i\rangle = \alpha' k_i^2 |0; k_i\rangle = |0; k_i\rangle) \\
&= \int d\tau \langle 0; -k_3|e^{i\tau} V_{k_2}(0) e^{-i\tau}|0; k_3\rangle \quad (\text{use normal ordering of } V_{k_2}) \\
&= \int d\tau \langle 0; -k_3|e^{ik_2 \cdot x}|0; k_1\rangle = \int d\tau \langle 0; -k_3|0; k_1 + k_2\rangle \\
&= \delta(k_1 + k_2 + k_3) \int d\tau \,.
\end{aligned} \tag{4.28}$$

The fact that the seemingly infinite number $\int_{-\infty}^{\infty} d\tau$ appears is because the theory still has unfixed gauge freedoms in terms of conformal transformations. To account for this correctly, we need to "divide out" by their "volume Vol(conf$_{\text{res}}$)", which is equivalent to have fixed the gauge (e.g., following the Faddeev–Popov procedure, see below), in which case the "empty integral" would not have appeared in the first place. We should, however, include an a priori undetermined factor g_o that corresponds to a "coupling constant" (for the open string). This results in the (correct) 3-point amplitude with open string tachyons:

$$\mathcal{A}_{3,\,\text{open}}(k_1, k_2, k_3) = g_o\, \delta(k_1 + k_2 + k_3)\,. \tag{4.29}$$

The computation with closed string tachyons proceeds analogously. Though we have not shown it explicitly, the closed-string tachyon vertex operator takes the same form, $V_k^{\text{cl}}(\xi^\pm) =: e^{ik\cdot X(\xi^\pm)} :$ with $\alpha' k^2 = 4$. The insertion into the bulk, as in the middle of Figure 5, is described by an insertion at the "origin", $\xi^\pm = 0$, followed by a translation in $\xi^\pm$ using $\tilde{L}_0$ and L_0, respectively. Note that $L_0|0;k_i\rangle = \tilde{L}_0|0;k_i\rangle = |0;k_i\rangle$. The amplitude, with g_c the closed string coupling constant, is then

$$\begin{aligned}
&\mathcal{A}_{3,\,\text{closed}}(k_1,k_2,k_3)\\
&= g_c \int d^2\xi \langle 0;-k_3|\, e^{2i(\xi^- L_0+\xi^+\tilde{L}_0)}\, V^{\text{cl}}_{k_2}(0)\, e^{-2i(\xi^- L_0+\xi^+\tilde{L}_0)}|0;k_1\rangle/\text{Vol}(\text{conf}_{\text{res}})\\
&= g_c \int d^2\xi \langle 0;-k_3|e^{ik_2\cdot x}|0;k_1\rangle/\text{Vol}(\text{conf}_{\text{res}})\\
&= g_c\, \delta(k_1+k_2+k_3) \int d^2\xi/\text{Vol}(\text{conf}_{\text{res}}) = g_c\, \delta(k_1+k_2+k_3)\,.
\end{aligned} \tag{4.30}$$

On the problem sheet, you will also compute the 3-point amplitudes with tachyons and level-one states.

The Veneziano amplitude

In the tree-level scattering of three states, we only have one relevant diagram. It is most apparent for the open string, where, after the conformal transformation, we are led to inserting the vertex operators on the boundary of a disc. After fixing two insertions corresponding to $|0;k_1\rangle$ and $|0;k_3\rangle$ at the antipodal points, it makes physically no difference where the third one between them sits. Notice that, a prior, we could also have treated any other pair as in- and out-states, with the third being inserted in between.

In fact, the state/operator correspondence gives another rather miraculous feature of string amplitudes: it is given by a vacuum expectation value, or correlation function, of vertex operators corresponding to all in- and out-states in a free 2d theory. Where in QFT, one has to use the LSZ-machinery to reduce amplitudes to correlation functions, the worldsheet description somehow provides this for free.

The modern way of computing correlation functions is via the path-integral formalism. This is particularly powerful in dealing with residual gauge symmetries, by means of the Faddeev–Popov method. If we perform this analysis on worldsheets that are topologically a disk or a sphere, i.e., for tree-level diagrams of open or closed string states, respectively, we reduce the full conformal group to

$$\text{conf}_{\text{res}} = \begin{cases} \text{PSL}(2,\mathbb{R}) \text{ with } \dim_{\mathbb{R}} = 3 & \text{(disc)},\\ \text{PSL}(2,\mathbb{C}) \text{ with } \dim_{\mathbb{R}} = 6 & \text{(sphere)}. \end{cases} \tag{4.31}$$

These residual transformations made it possible for the 3-point amplitudes to fix the positions of all three vertex operators completely.[9] For four or more states, we can therefore completely fix the gauge by specifying the position of three vertex operators, and then integrate over the remaining ones to maintain diffeomorphism invariance.

With this intermezzo, let us now turn to the scattering of four open-string tachyons $|0; k_i\rangle \equiv |i\rangle$. After conformally mapping to the disc and fixing, say, the operator insertions for states $i = 1, 2, 4$, the remaining operator insertion can be split into three contributions, see top row of Figure 8.

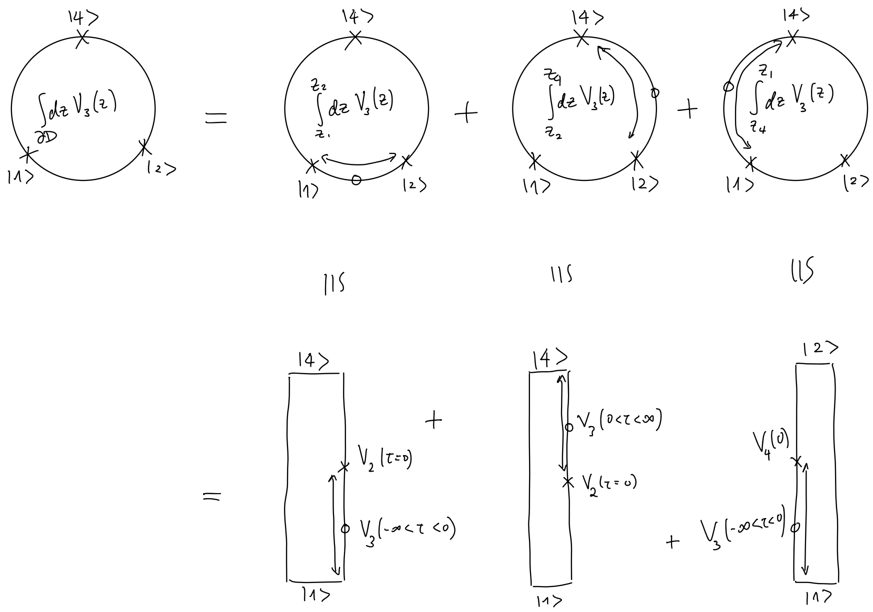


Fig. 8: After fixing the residual conformal gauge symmetries, the scattering of four open-string tachyons can be divided into three contributions, depending on the insertion location of the vertex operator $V_{k_3} \equiv V_3$.

Let us focus on the first contribution. To evaluate it with the techniques available to us, we have to map the disc back into the "strip" geometry; the second state is associated to the vertex operator with fixed position at finite τ, say, at $\tau = 0$,

[9] Note that, since the open string vertex operators are inserted on the boundary, there are only three real degrees of freedoms that precisely soak up the three generators of PSL(2, $\mathbb{R}$). For closed string vertices that are inserted in the interior, we can again soak up all real degrees of freedom for precisely three states.

see bottom row in Figure 8. Then the position of the second vertex operator is at $\tau \in [-\infty, 0]$. So, interpreting the diagram as a transition amplitude from $|1\rangle$ to $|4\rangle$, we have

$$\mathcal{A}^{(1)}_{4,\,\text{open}} = g_o^2 \int_{-\infty}^{0} d\tau \langle 0; -k_4 | V_{k_2}(0) V_{k_3}(\tau) | 0; k_1 \rangle \,, \tag{4.32}$$

where the coefficient g_o^2 can be intuitively explained by imagining the 4-point diagram as joining two 3-point diagrams (we will give a better explanation later). To obtain a convergent integral, it is again standard to Wick-rotate the integration contour into the imaginary / Euclidean time. As usual in QFT, one must make sure that the contour rotation does not cross any poles in the interior of the complex τ-plane; this is indeed possible for the integral above. In the end, we end up with the expression $\int_{-\infty}^{0} dt \langle 0; -k_4 | V_{k_2}(0) V_{k_3}(-it) | 0; k_1 \rangle$, which you are ask to compute on the Problem Sheet. The final result is

$$\mathcal{A}^{(1)}_{4,\,\text{open}} = g_o^2\, \delta\left(\sum_{i=1}^{4} k_i\right) \frac{\Gamma(-\alpha(s))\Gamma(-\alpha(u))}{\Gamma(-\alpha(s)-\alpha(u))}\,, \qquad \alpha(x) := 1 + \alpha' x \,, \tag{4.33}$$

where $s = -(k_1 + k_2)^2$ and $u = -(k_1 + k_4)^2$, and $\Gamma(x) = \int_0^\infty dt\, t^{x-1} e^{-t}$ is Euler's Gamma function. The two other diagrams can be evaluated in a similar fashion, and one finds

$$\begin{aligned} \mathcal{A}^{(2)}_{4,\,\text{open}} &= g_o^2\, \delta\left(\sum_i k_i\right) \frac{\Gamma(-\alpha(u))\Gamma(-\alpha(t))}{\Gamma(-\alpha(t)-\alpha(u))}\,, \\ \mathcal{A}^{(3)}_{4,\,\text{open}} &= g_o^2\, \delta\left(\sum_i k_i\right) \frac{\Gamma(-\alpha(s))\Gamma(-\alpha(t))}{\Gamma(-\alpha(t)-\alpha(s))}\,, \end{aligned} \tag{4.34}$$

where we also introduced the third Mandelstam variable $t = -(k_1 + k_3)^2$.

Using the relationship $B(a, b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$ between Euler's Gamma and Beta (B) functions, the full amplitude can be expressed as

$$\mathcal{A}_{4,\,\text{open}} = g_o^2\, \delta(\sum_i k_i)\ \left[B(-\alpha(s), -\alpha(u)) + B(-\alpha(u), -\alpha(t)) + B(-\alpha(s), -\alpha(t))\right] \,. \tag{4.35}$$

This is the famous *Veneziano amplitude*. It displays a variety of remarkable features that highlights the difference between particle and string scatterings.

The first one that is easy to see is the presence of poles at $s = \frac{n-1}{\alpha'}$, or $u = \frac{n-1}{\alpha'}$, or $t = \frac{n-1}{\alpha'}$ for $n \in \mathbb{N}_0$ a non-negative integer. These come from the Gamma functions in the numerators, which have poles at $-\alpha(s,u,t) = 0, -1, -2, \ldots$ Notice that they correspond precisely to the masses of level-n open string states! Of course, these are just the on-shell resonances of tree-level 2-to-2 scatterings that should also be familiar from point-particles. What is different is that there are *infinitely many* of them, which would seem bizarre from a particle perspective, but can now be naturally interpreted as all the excitations of a single string.

Another property worth mentioning is the "high-energy, fixed-angle" limit: we let $|1\rangle$ and $|2\rangle$ collide head-on with $s = -(k_1+k_2)^2 \longrightarrow \infty$ and "measure" the outgoing states at a fixed angle θ, see Figure 9. This means that the ratios $t/s \approx -\sin^2\theta/2$, $u/s \approx -\cos^2\theta/2$, and $t/u \approx \tan^2\theta/2$ are fixed. Using Stirling's approximation, $\Gamma(n+1) = n! \approx \sqrt{2\pi n}(n/e)^n$, and the relationship $s + t + u = \sum_i M_i^2 = -4/\alpha'$ for Mandelstam variables, one can show that

$$\mathcal{A}_{4,\,\text{open}} \approx F(\theta)^{-\alpha' s}\,, \tag{4.36}$$

where $F(\theta)$ is a positive function in the angle θ.

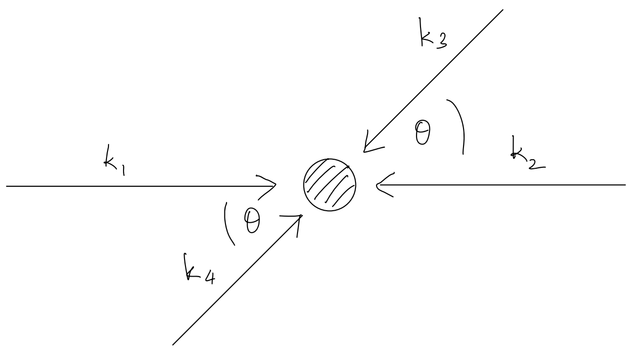


Fig. 9: Head-on collision of two particles / string states with momenta k_1 and k_3, producing two new particles / states with momenta k_3 and k_4 that are deflected from the original trajectories by an angle θ.

Contrast to this exponential fall-off behavior for strings, the fixed-angle scattering of point-particles typically exhibits a power-law. E.g., in scatterings that exchange spin-J, mass M particles, the amplitudes behaves as

$$\mathcal{A} \sim \frac{t^J}{s - M^2}\,. \tag{4.37}$$

Naively, having different contributions from higher-spin particles would make the total amplitude diverge — which is the case for point-particle theories. However, the string has an *infinite* tower of higher-spin states, and they conspire in a specific way to cancel out the high-energy divergences. There is another interpretation: Since the $s \to \infty$ limit probes smaller and smaller distances, the point-particle will inevitably "detect" the singular nature of the interaction vertex. On the other hand, the string length $\ell_s \sim \sqrt{\alpha'}$ naturally resolves this singularity, and thus regularizes the amplitude by itself.

4.4 String coupling, loop amplitudes, and other remarks

In the following, we will discuss a few other distinct aspects of string amplitudes, without going into the details too much.

More on the genus expansion

First, let us give some more motivation to the origin of the coupling coefficients g_o and g_c. They are tied to the "uniqueness" of worldsheet theory, in that we cannot add any interaction terms without breaking conformal and Weyl invariance. There is one exception that was discussed in Section 2 and on the first Problem Sheet: the Euler *density*, whose integral gives the Euler number,

$$\chi = \frac{1}{4\pi}\int_\Sigma d^2\xi\sqrt{-\det\gamma}\mathcal{R}(\gamma) + \frac{1}{2\pi}\int_{\partial\Sigma} ds\,\mathcal{K}(\gamma)\,. \tag{4.38}$$

As a result of the Riemann–Roch-theorem, this is a topological invariant, and can be expressed as

$$\chi = 2 - 2g - b\,, \tag{4.39}$$

where g is the genus (= number of handles) and b is the number of boundary components. As a few example, we have

- sphere: $(g, b) = (0, 0)$, torus: $(g, b) = (1, 0)$, any closed compact surface: $(g, b) = (n, 0)$;
- disc: $(g, b) = (0, 1)$, cylinder: $(g, b) = (0, 2)$, pants: $(g, b) = (0, 3)$.

If we include this term, the full action becomes $S[X,\gamma] = S_{\mathrm{P}}[X,\gamma] + \lambda\chi$, with $\lambda \in \mathbb{R}$, which has the same local worldsheet dynamics. Now, in the path-integral formalism, we have to sum over all possible background configurations; in the present case, this now has to include different topologies of the worldsheet. Then the correlation functions are computed schematically as

$$\begin{aligned} \mathcal{A}(|1\rangle, \dots, |n\rangle) &= \sum_{\text{topologies}} \int \frac{\mathcal{D}[X,\gamma]}{\mathrm{Vol}(\mathrm{conf}_{\mathrm{res}})}\, e^{-S[X,\gamma]} \prod_{i=1}^{n} \mathcal{V}_{|i\rangle} \\ &= \sum_{\text{topologies}} (e^{\lambda})^{-\chi} \frac{\mathcal{D}[X,\gamma]}{\mathrm{Vol}(\mathrm{conf}_{\mathrm{res}})}\, e^{-S_{\mathrm{P}}[X,\gamma]} \prod_{i=1}^{n} \mathcal{V}_{|i\rangle}\,, \end{aligned} \tag{4.40}$$

where $\mathcal{V}_{|i\rangle}$ are the integrated vertex operator insertions associated to the states $|i\rangle$. Note that the sum over topologies is precisely the asymptotic series shown in Figure 6, with $g_s = e^{\lambda}$ being the "expansion parameter".

Moreover, consider any worldsheet and attaching a "handle" to it, see Figure 10, top row. Since this increases g by one, the Euler characteristic decreases, $\chi \longrightarrow \chi - 2$, so the amplitude picks up a factor g_s^2. Physically, this new diagram has a closed string loop, which in the Feynman diagram limit would have two vertices, each describing the emission and absorption of a closed string state. Hence, it is natural to identify the closed string coupling constant with $g_c = g_s = e^{\lambda}$.

Likewise, we can also add to any worldsheet an "interior" boundary (see Figure 10, bottom row), which increases b by one, so decreases $\chi \to \chi - 1$. The two corresponding amplitudes are related to each other by adding an open string loop, which leads to an additional factor g_s. Since this would also have two identical vertices in the Feynman diagram limit describing open-string emission/absorption, we conclude that $g_o = \sqrt{g_s} = \sqrt{g_c}$.

To keep track of the powers of g_s that gives to a good particle-limit interpretation, one defines the integrated vertex operators $\mathcal{V}_{|i\rangle}$ as carrying a factor of g_o if $|i\rangle$ is an open-string state, and g_c if $|i\rangle$ is a closed-string state. E.g., the open/closed string tachyons would have the integrated vertex operators

$$\mathcal{V}_{|0;k\rangle_o} = \sqrt{g_s} \int d\tau\, V_k(\tau)\,, \quad \mathcal{V}_{|0,0;k\rangle_c} = g_s \int d^2\xi\, V_k^{\mathrm{cl}}(\xi)\,. \tag{4.41}$$

Inserting these for the 3- and 4-point tachyon amplitudes, with the Euler characteristics of the disc / sphere, gives the coupling coefficients shown in the results above.

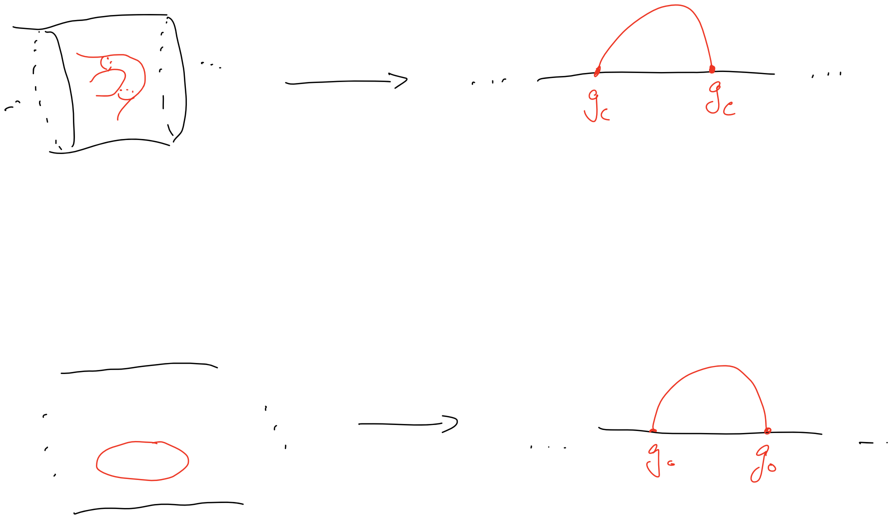


Fig. 10: Top: adding a handle to a worldsheet, which changes $g_s^{-\chi} \to g_s^{-\chi} g_s^2$ corresponds to adding a closed-string state in a loop, which comes with a factor g_c^2. Bottom: adding a boundary to a worldsheet, which changes $g_s^{-\chi} \to g_s^{-\chi} g_s$ corresponds to adding an open-string state in a loop, which comes with a factor g_o^2.

Moduli space of Riemann surfaces and 1-loop amplitudes

After gauging fixing, the path-integral $\frac{\mathcal{D}[\gamma]}{\text{Vol}(\text{conf}_{\text{res}})}$ over the metric reduces to an integral over the so-called *moduli space* for topologies with $\chi \leq 0$. This is a space of parameters that characterize the string worldsheet as a *complex* 2d compact manifold, also known as a Riemann surface.

The easiest example is the torus with $\chi = 0$. As a complex manifold, the torus can be represented by a parallelogram in $\mathbb{C}$, whose parallel sides are identified. The metric on the torus is given by $ds^2 = |dx + \tau dy|^2$, where $\tau \in \mathbb{C}$ (this is the standard notation, not to be confused with time coordinate on Σ!) defines the "shape" of parallelogram, see Figure 11. This is the single (complex) modulus of the torus, also known as the *complex structure*.

Naively, one might expect that any $\tau \in \mathbb{C}$ defines a distinct torus. However, the gauge fixing procedure eliminates most of the complex plane, because different

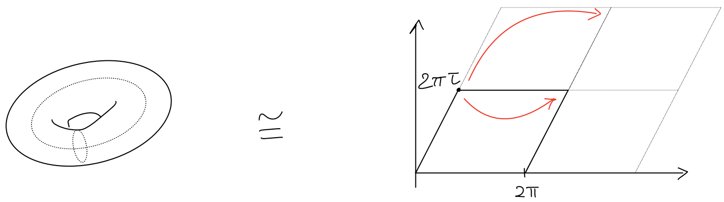


Fig. 11: The torus can be thought of a parallelogram with opposite sides glued together. Equivalently, it can be thought of as the quotient $\mathbb{C}/\Lambda$ by a lattice $\Lambda \subset \mathbb{C}$ spanned by 2π and $2\pi\tau$. Möbius transformations $\tau \mapsto \tau'$ leave the lattice invariant, hence also the associated torus.

complex structure define the *same* torus if they are related by an $\mathrm{PSL}(2,\mathbb{Z})$, or "Möbius transformation",

$$\tau \longrightarrow \tfrac{a\tau+b}{c\tau+d}\,, \qquad \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \cong -\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{SL}(2,\mathbb{Z})\,. \tag{4.42}$$

Hence, inequivalent tori are labelled by points in the *fundamental domain*,

$$F_0 = \frac{\mathbb{C}}{\mathrm{PSL}(2,\mathbb{Z})} = \{\tau \mid |\tau| \geq 1,\ -\tfrac{1}{2} \leq \mathrm{Re}(\tau) \leq \tfrac{1}{2}\,,\ \mathrm{Im}(\tau) > 0\}\,, \tag{4.43}$$

see Figure 12. For an n-point closed string scattering, the 1-loop amplitude is the computed as follows. We first evaluate the path-integral over X, with n vertex operator insertions, over a torus with fixed τ and the corresponding flat metric, and then integrate the result as a function of τ over F_0.

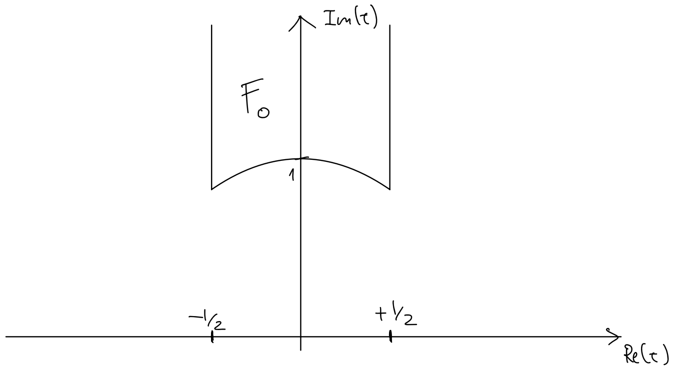


Fig. 12: The fundamental domain F_0 of $\mathrm{PSL}(2,\mathbb{Z})$; every $\tau \in F_0$ defines a unique torus.

To see the physical relevance of this prescription, let us turn to the simplest amplitude with *no* operator insertions: the 1-loop vacuum amplitude. The computation would require the full machinery of the path-integral, which we will not present here; it can be found in varying level of details in any of the references to this course. The result is the following expression

$$\mathcal{A}_{T^2} \sim \int_{F_0} \frac{d\tau d\bar{\tau}}{\mathrm{Im}(\tau)^2}\, \mathrm{Im}(\tau)^{-12}\, |\eta(\tau)|^{-48}\,; \tag{4.44}$$

here, $\eta(\tau)$ is the so-called "Dedekind η function". A generally crucial property which we are just mentioning in passing here is that the integrand is *modular invariant*, i.e., it is invariant under any $\mathrm{PSL}(2,\mathbb{Z})$ transformation on τ. Here it is merely a consistency check on the result, but it is an extremely restricting property for torus amplitudes of general 2d CFTs.

To get some more intuition about this expression, let us compare with the 1-loop vacuum amplitude of an ordinary d-dimensional QFT of a free particle of mass m. This is computed by summing over all particle paths that are topologically a circle in spacetime, and yields, up to some irrelevant overall factors, the expression

$$\begin{aligned}\mathcal{A}_{S^1}(m^2) &\sim \int \frac{d^d k}{(2\pi)^d} \int_0^\infty \frac{dl}{l} \exp(-l(k^2+m^2)) \\ &\sim \int_0^\infty dl \frac{e^{-m^2 l}}{l^{1+d/2}}\,.\end{aligned} \tag{4.45}$$

The exponent $l(k^2+m^2)$ comes from the Hamiltonian for the particle, and the l-integral takes into account all values of the radius/circumference of the circle (including accounting for the unfixed symmetry $\mathrm{Diff}(S^1)$) that the particles is moving on; carrying out the momentum integral, we obtain (up to constant factors) the second line. Now, if we would just approximate the string amplitude in the QFT-limit as having infinitely many particle-like states, then the standard QFT prescription would be to sum over all such states (more precisely, over all mass eigenstates). In the closed string Hilbert space, the mass-squared of a mass eigenstate $|\psi\rangle$ is the eigenvalue of $\frac{2}{\alpha'}(L_0+\tilde{L}_0-2)$. Importantly, they are subject to the level matching constraint, $L_0=\tilde{L}_0$, which can be implemented via a Kronecker δ in its integral representation,

$$\delta_{L_0,\tilde{L}_0} = \int_{-\frac{1}{2}}^{\frac{1}{2}} \frac{ds}{2\pi} e^{2\pi i s(L_0-\tilde{L}_0)}\,. \tag{4.46}$$

So, the particle-motivated result for the 1-loop vacuum amplitude is

$$\begin{aligned}\mathcal{A}_{\text{1-loop}} &\sim \sum_{|\psi\rangle}\int_0^\infty dl\frac{1}{l^{14}}\int_{-\frac{1}{2}}^{\frac{1}{2}} ds\,\langle\psi|\exp(2\pi i s(L_0-\tilde{L}_0)-\tfrac{2l}{\alpha'}(L_0+\tilde{L}_0-2))|\psi\rangle\\ &= \int_0^\infty \frac{dl}{l^2}\int_{-\frac{1}{2}}^{\frac{1}{2}} ds\sum_{|\psi\rangle} l^{-12}\langle\psi|\exp(2\pi i\tau(L_0-1)-2\pi i\bar{\tau}(\tilde{L}_0-1))|\psi\rangle\,,\end{aligned} \tag{4.47}$$

where we have defined $\tau = s + i\frac{l}{\pi\alpha'}$.

It is a non-trivial result that the sum over all states $|\psi\rangle$ of the closed string Hilbert space of the complicated looking exponential is precisely $\eta(\tau)^{-24}\bar{\eta}(\bar{\tau})^{-24} = |\eta(\tau)|^{-48}$. Since up to constants we have $l \propto \mathrm{Im}(\tau)$, the final result is almost identical to the string amplitude (4.44). The key difference is in the integration domain: while in (4.47) we integrate $\mathrm{Im}(\tau)$ from 0 onwards, the integral in (4.44) is over the fundamental domain F_0 of $\mathrm{PSL}(2,\mathbb{Z})$, which has $\mathrm{Im}(\tau) \geq \sqrt{\frac{3}{4}}$.

This is significant because the $l \approx 0$ region in the particle setting is responsible for the UV-divergence of the amplitude: these correspond to arbitrarily high-energy particles that run in the loop. In contrast, the actual string-theoretic amplitude avoids this region completely. This is possible because, by the $\mathrm{PSL}(2,\mathbb{Z})$ transformation $\tau \to -1/\tau$, the $\mathrm{Im}(\tau) \approx 0$ region is "already covered" by the integral over the $\mathrm{Im}(\tau) \to \infty$ region.

Therefore, the 1-loop vacuum amplitude is *UV-finite* in string theory. The same principle — the integration domain being F_0 — also applies to amplitudes with vertex operator insertions (including gravitons!) on the torus, so all 1-loop amplitudes are UV-finite. For higher loops, the computation quickly becomes unfeasible. It is still an ongoing topic of research to show UV-finiteness at all loop orders. However, with the state-of-the-art result extending to 5-loop amplitudes of the superstring, it is widely believed that (super-)string theory is UV-finite.

What about the IR behavior? This is the limit $\mathrm{Im}(\tau) \to \infty$, so the part where the would-be UV-divergent contributions are mapped onto by $\mathrm{PSL}(2,\mathbb{Z})$. In this regime we can use the expansion $\eta(\tau) = q^{\frac{1}{24}} \prod_{n=1}^{\infty}(1-q^n)$, with $q = e^{2\pi i \tau}$, to find

$$\begin{aligned} \int_{F_0} \frac{d^2\tau}{\mathrm{Im}(\tau)^{14}} |\eta(\tau)|^{-48} &= \int_{F_0} \frac{d^2\tau}{\mathrm{Im}(\tau)^{14}} (q\bar{q})^{-1}(1+24q+\ldots)(1+24\bar{q}+\ldots) \\ = \int_{-\frac{1}{2}}^{\frac{1}{2}} d\mathrm{Re}(\tau) \int_{\sqrt{1-\mathrm{Re}(\tau)^2}}^{\infty} \frac{d\mathrm{Im}(\tau)}{\mathrm{Im}(\tau)^{14}} &\left(e^{4\pi \mathrm{Im}(\tau)} + 24^2 + \ldots\right) , \end{aligned} \tag{4.48}$$

where we have dropped terms that vanish upon integrating over $\mathrm{Re}(\tau)$. From this, we see that this integral does exhibit IR-divergence, but it comes from the term $e^{4\pi\mathrm{Im}(\tau)}$. In a more careful analysis of the origin of the η function, one would find that this term arises from having the tachyon; the next term, 24^2, comes from the massless states, and is finite. All higher level states also give finite contributions. For the superstring, the absence of a tachyon then implies that the 1-loop amplitude is also IR-finite.

Open-close duality and tadpole cancellation

The 1-loop amplitude of the open string is defined by the cylinder-topology, see left of Figure 13. There is only one modulus t that measures the circumference / radius of the cylinder, which does not enjoy any modular symmetries like $\mathrm{PSL}(2,\mathbb{Z})$. Hence, it can take any value, $0 \le t \le \infty$. Quoting just the result of the computation, the amplitude is

$$\mathcal{A} \sim \int_0^{\infty} \frac{dt}{t^{14}} \eta(it)^{-24} \,. \tag{4.49}$$

The IR limit is very similar to the closed string: there is a contribution coming from the tachyon, which diverges for $t \to \infty$. Once we go to the superstring, the amplitude becomes IR-finite.

However, unlike for the closed string, the UV-region, $t \approx 0$, is part of the integration domain, and there is no way to avoid the resulting divergence. Thus, it seems that a theory with open strings contradicts our previous claim about string theory being UV-finite.

Of course, there is a resolution of this contradiction, and it simply comes by reinterpreting the cylinder-topology of the worldsheet as a *tree-level* amplitude of a closed string! (See Figure 13.) This is a general feature of string worldsheets called the open-closed (string channel) duality, and can also be applied to other amplitudes.

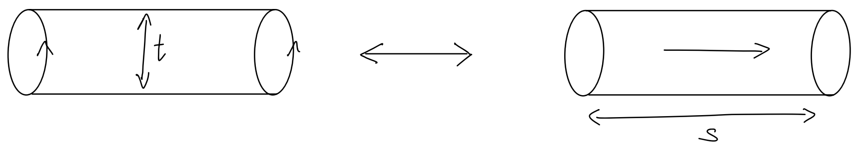


Fig. 13: The open string 1-loop vacuum diagram has the worldsheet topology of a cylinder with modulus t. This can be equivalently viewed as a tree-level propagation of closed string modes with momenta $|k| \approx 1/s$.

For the 1-loop open string vacuum amplitude, the corresponding tree-level closed string amplitude can be technically obtained by exchanging the roles of the (Euclidean) time and the spatial coordinates on the worldsheet. Geometrically, this actually affects the modulus of the cylinder, via $t \to s = 1/t$. Under this transformation, the η function becomes $\eta(it) = \sqrt{s}\,\eta(is)$, so the total amplitude transforms as

$$\begin{aligned}\mathcal{A} \sim \int_0^\infty \frac{dt}{t^{14}}\,\eta(it)^{-24} &= -\int_\infty^0 \frac{ds}{s^2}\,s^{14}\,s^{-12}\eta(is)^{-24} = \int_0^\infty ds\,\eta(is)^{-24} \\ &= \int_0^\infty ds(e^{2s} + 24 + O(e^{-2s}))\,.\end{aligned} \tag{4.50}$$

The closed string interpretation is that we have to sum over all possible closed string states (which gives the η function) that propagate along the cylinder with momenta $|k| \sim s$ (which is integrated over). Now, the divergence again comes from the $s \to \infty$ region, where, as before, the tachyon gives the most divergent piece. In superstring theories, the leading term is then due to the second contribution coming from the massless states.

This kind of IR divergence is completely analogous to IR divergences in standard QFTs. They arise from the propagator $1/k^2$ of massless particles of vanishing momentum. There are may ways of dealing with such divergences in QFT, e.g., introducing IR cutoffs, adding appropriate counterterms etc. For string theory, there is another nice interpretation, namely the existence of so-called tadpole diagrams, i.e., where a state is created from the vacuum. Indeed, the divergent contributions to the above amplitude can be thought of as connecting two tadpoles by the propagator of a massless string state, see Figure 14. In QFT, a tadpole is another kind of instability to the vacuum, stemming from a perturbative expansion around a field configuration ϕ_0 that is not a critical point of the potential, $V'(\phi_0) \neq 0$.

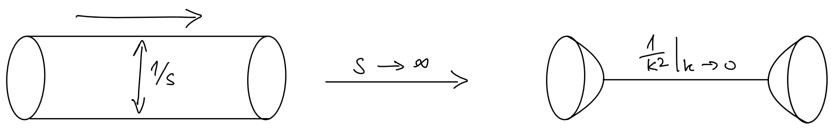


Fig. 14: In the IR-region $s \to \infty$, the closed string amplitude can be loosely interpreted as coming from two closed string tadpoles that are connected by a propagator of a massless particle with momentum $|k| \sim 1/s \to 0$.

To guarantee the stability of the vacuum, we must therefore “cancel” the tadpoles. It turns out that the way to do it correctly is to introduce *spacetime filling* branes, to which the endpoints of the open strings that caused the problem in the first place are attached. The nature and number of such branes can be determined by more involved computations of string amplitudes on *non-oriented* worldsheets (such as the Möbius strip or the Klein bottle). The result of this analysis, however, actually implies the existence of *non-abelian gauge fields* in spacetime!

This comes about because the massless states of the open string, which were spacetime vector bosons, now carry “labels” (so-called Chan–Paton factors) of the brane that they end on. As we will elaborate on in Section 5.3.2, these labels can be interpreted as vector indices of elements in a Lie algebra $\mathfrak{g}$. Moreover, the tree-level scattering amplitudes of these vector bosons turn out to be consistent with those of a $\mathfrak{g}$-gauge field at low energies, thus justifying the interpretation above.

In the bosonic string, the brane system that facilitates the tadpole cancellation gives rise to the gauge algebra $\mathfrak{g} = \mathfrak{so}(8192) = \mathfrak{so}(2^{13})$. In (one of the) superstring theories, similar arguments fix the gauge symmetry in 10d to be $\mathfrak{so}(32)$.

These numbers may look arbitrary and just a result of some complicated computations. But one may also view them as a profound *feature* of string theory that further distinguishes it from ordinary QFTs: to facilitate a *consistent* UV-finite theory that includes both gauge bosons and gravitons in the (low-energy) particle-limit, the allowed gauge dynamics is heavily constrained. There are many more such examples of consistency conditions on effective field theories that arise in the particle-limit of string theory, and they culminate in the development of the “Swampland Program”, which is an active field of research.

String theory, non-perturbatively

By now, it should have become clear that the genus expansion is essentially the stringy version of Feyman diagram expansion in QFT. While powerful for small couplings, there are many field theory phenomena that cannot be handled in this perturbative ansatz. These not only include strong-coupling effects such as confinement, but also non-perturbative effects like instantons and solitons at weak coupling.

In field theory, the path-integral in principle provides a fully non-perturbative definition. In the modern perspective, the fields over which we path-integrate have quantized fluctuations that give rise to the particles whose worldlines build up the Feynman diagrams. If we want to extend this picture to string theory, then we would need "string fields", whose excitations are strings.

This is the approach of *string field theory*, and it has had some remarkable successes, particularly with open strings, including finding and understanding the true, non-perturbative vacuum of bosonic string theory. However, it struggles with closed strings, and in some sense, this is related to the fact that the closed string sector includes a theory of gravity, which does not have off-shell quantities. As we have seen above, an open string theory must also incorporate closed strings, and just how to describe this from string field theory is still to be understood.

Another approach is based on the background field description which we will introduce momentarily. In this approach, non-perturbative g_s-effects can be often understood via (spacetime) geometry, branes, and (string) dualities. An important fine print here is that these approaches usually rely on some (spacetime) supersymmetry to be computationally feasible. Nevertheless, frameworks such as M-/F-theory or the AdS/CFT correspondence gave rise to tremendous insights about non-perturbative dynamics in field theory limits with and without gravity.

5 Strings in Background Fields

Having seen the various intriguing features of string amplitudes *from the spacetime perspective*, we will now further study the implications of having propagating strings in spacetime.

5.1 Einstein's equations from the worldsheet

After having introduced the Polyakov action S_{P} (2.19) with an arbitrary spacetime metric $g_{\mu\nu}$ in the classical setting, we have quickly resorted to setting it to the flat Minkowski background to simplify things. However, having seen that there are massless gravitons in the closed string spectrum, we can get some intuition for the general setting, by adding a small perturbation to the flat metric,

$$g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu} \,. \tag{5.1}$$

If we insert this into the path-integral, then the exponential of the action (gauge fixed to have flat worldsheet metric) becomes

$$\exp(-S[h]) = \exp(-S_{\mathrm{P}}) \left(1 - \frac{1}{4\pi\alpha'} \int_\Sigma d^2\xi \partial_a X^\mu \partial^a X^\nu h_{\mu\nu}(X(\xi)) + \ldots \right) \,. \tag{5.2}$$

This can be interpreted as the insertion of an operator

$$\mathcal{V}_h = \frac{1}{2\pi\alpha'} \int d^2\xi \partial_a X^\mu \partial^a X^\nu h_{\mu\nu}(X(\xi)) \,. \tag{5.3}$$

But we know about consistency conditions of such operator insertions: they must be vertex operators!

Though we have not discussed this explicitly for the closed string states, it is apparent that, if we choose $h_{\mu\nu}(X) = \gamma_{\mu\nu} :e^{ik\cdot X}:$, with $\gamma_{\mu\nu}$ a symmetric traceless tensor, this precisely corresponds to a vertex operator that generates a *plane gravity wave* with polarization $\gamma_{\mu\nu}$. A general vertex operator is a linear superposition $h_{\mu\nu}$ of such plane waves — this is the physical consistency condition on the perturbation $h_{\mu\nu}$ that comes from having a quantum string in spacetime.

Extending this to a general spacetime metric $g_{\mu\nu}$ is hard, because the worldsheet theory defined by the Polyakov action,

$$S = \frac{1}{4\pi\alpha'} \int d^2\xi \, g_{\mu\nu}(X) \partial_a X^\mu \partial^a X^\nu \,, \tag{5.4}$$

will no longer be free. To see the type of (worldsheet) interactions such a modification would introduce, we now make a "perturbation ansatz" for the coordinate fields,

$$X^\mu(\xi) = X_0^\mu(\xi) + \sqrt{\alpha'} Y^\mu(\xi) \,. \tag{5.5}$$

For simplicity, we will assume that the "background value" $X_0^\mu(\xi)$ is constant on the worldsheet, though this need not be the case. Then, a Taylor expansion of $g_{\mu\nu}$ in the Lagrangian yields

$$\begin{aligned} & g_{\mu\nu}(X)\partial_a X^\mu \partial^a X^\nu \\ & = \alpha' \left[g_{\mu\nu}(X_0) + \sqrt{\alpha'} g_{\mu\nu,\rho}(X_0) Y^\rho(\xi) + \frac{\alpha'}{2} g_{\mu\nu,\rho\sigma}(X_0) Y^\rho(\xi) Y^\sigma(\xi) + \ldots \right] \partial_a Y^\mu \partial^a Y^\nu , \end{aligned} \tag{5.6}$$

where we have use the standard GR-notation $g_{\mu\nu,\rho} \equiv \partial_\rho g_{\mu\nu}$. In this expansion, each term now represents an interaction for the fluctuations Y^μ, with interaction coefficient given by the metric derivatives $g_{\mu\nu,\ldots}$.

Notice that this expansion is controlled by powers of $\sqrt{\alpha'}$. In a spacetime with "typical" length scales r_c (e.g., radius of sphere, Hubble constant, ...), the metric derivatives behave like $\partial g/\partial X \sim 1/r_c$, so the effective dimensionless coupling constant is of order $\sqrt{\alpha'}/r_c$. Hence, the above couplings can be treated perturbative if $\sqrt{\alpha'} \sim \ell_s \ll r_c$, i.e., when the size of the string is much smaller than typical length scales. At each α'-order, we still have the full genus expansion for amplitude computations on the worldsheet. This means that perturbative string theory enjoys a *double expansion* in g_s and α'.

For $\ell_s/r_c \ll 1$, the results may be interpreted as an infinite set of particles (with special properties coming from the genus expansion) that probe the spacetime geometry. However, for $\ell_s/r_c \approx 1$, this interpretation breaks down badly; now, the spatial extension of the string becomes important, and gives rise to the notion of "stringy geometry". E.g., while particle interactions would break down in geometries with certain singularities, a string can effectively "resolve" such singularities and be well-behaved.

Of course, all of these statements only make sense if we actually *can* rely on the genus expansion. More explicitly, the results of the previous sections relied heavily on the conformal gauge invariance of the worldsheet theory. Classically, the action built from (5.6) is conformally invariant, but this is not necessarily true at the quantum level, because, unlike before, we now have an interacting field theory on the worldsheet. These interactions can lead to unphysical divergences of (worldsheet) correlation functions. In a typical field theory, the go-to method to deal with this issue is *renormalization*. However, as you will learn in the "Advanced QFT" course, renormalization generally introduces an explicit scale dependence of correlation functions, in which case the theory is no longer conformally invariant. A prime example of a classically conformally invariant theory that develops a scale dependence quantum mechanically is Yang–Mills theory.

In slightly more technical terms, the scale dependence comes in the form of the “running of the coupling”. That is, a coupling parameter, in this case the spacetime metric $g_{\mu\nu}$, in the properly renormalized theory depends on the energy scale M, or momenta of the involved states; it is measure by the so-called β-function,

$$\beta(g_{\mu\nu}) = M\frac{\partial}{\partial M}g_{\mu\nu}\,. \tag{5.7}$$

So, if we insist on preserving conformal invariance of the quantum worldsheet theory in the presence of the interactions (5.6), then we must ensure that the β-functions of the couplings all vanish. This in turn constrains the spacetime metric $g_{\mu\nu}(X)$. Note that this is indeed the generalization of the constraints on the infinitesimal metric perturbations (5.1) to define *vertex* operators.

To actually compute the β-function, let us first make a simplification familiar from GR: locally around X_0, we choose *Riemann normal coordinates*; this has the effect that the metric locally takes the form

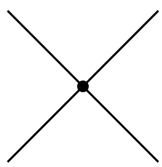
$$g_{\mu\nu}(X_0 + \sqrt{\alpha'}Y) = \eta_{\mu\nu} - \frac{\alpha'}{3}R_{\mu\lambda\nu\kappa}(X_0)Y^\lambda Y^\kappa + O(Y^3)\,, \tag{5.8}$$

with $R_{\mu\lambda\nu\kappa}(X_0)$ the spacetime Riemann tensor at X_0. Then, the action with the Lagrangian (5.6) becomes

$$S = \frac{1}{4\pi}\int d^2\xi\left((\partial_a Y)\cdot(\partial^a Y) - \frac{\alpha'}{3}R_{\mu\lambda\nu\kappa}Y^\lambda Y^\kappa \partial_a Y^\mu \partial^a Y^\nu + \ldots\right), \tag{5.9}$$

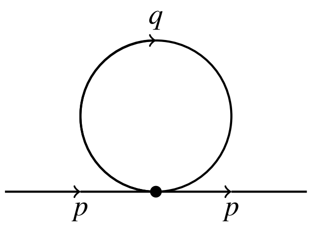
with the leading interaction being a quartic coupling in Y.

Such an interaction can be treated in the standard fashion: we use (momentum space) Feynman diagrams (now for the 2d worldsheet theory!) with the Feynman rule



$$\simeq \quad R_{\mu\lambda\nu\kappa}p_a^\mu (p^a)^\nu\,, \tag{5.10}$$

where p_a^μ ($a = 0, 1$) is the 2d momentum of the scalar field Y^μ. This gives the divergent 1-loop contribution



$$= \int \frac{d^2q}{(2\pi)^2} R_{\mu\lambda\nu\kappa} p_a^\mu (p^a)^\nu \frac{\eta^{\lambda\kappa}}{q^2}\,, \tag{5.11}$$

where $\eta^{\lambda\kappa}/q^2$ is the propagator in the loop.

The leading order divergence can be determined using "dimensional regularization", which yields

$$\int \frac{d^dq}{(2\pi)^d} R_{\mu\lambda\nu\kappa} p_a^\mu (p^a)^\nu \frac{\eta^{\lambda\kappa}}{q^2} \overset{d=2+\epsilon}{=} \frac{1}{2\pi\epsilon} R_{\mu\nu} p_a^\mu (p^a)^\nu + \dots\,, \tag{5.12}$$

with $R_{\mu\nu}$ the Ricci tensor associated to the metric $g_{\mu\nu}$. To cancel this divergence, one needs to add to the interaction the counter-term

$$R_{\mu\lambda\nu\kappa} Y^\lambda Y^\kappa \partial_a Y^\mu \partial^a Y^\nu \longrightarrow R_{\mu\lambda\nu\kappa} Y^\lambda Y^\kappa \partial_a Y^\mu \partial^a Y^\nu - \frac{1}{\epsilon} R_{\mu\nu} \partial_a Y^\mu \partial^a Y^\nu\,. \tag{5.13}$$

One can show that this can be absorbed by a "wave-function renormalization" and a "coupling renormalization",

$$Y^\mu \longrightarrow Y^\mu + \frac{\alpha'}{\epsilon} R^\mu_\nu Y^\nu\,, \quad g_{\mu\nu} \longrightarrow g_{\mu\nu} + \frac{\alpha'}{\epsilon} R_{\mu\nu}\,, \tag{5.14}$$

which then implies the 1-loop β-function

$$\beta(g_{\mu\nu}) \propto \alpha' R_{\mu\nu}\,. \tag{5.15}$$

Thus, worldsheet conformal invariance restricts the spacetime metric $g_{\mu\nu}$ to be Ricci-flat, $R_{\mu\nu} = 0$. This in particular implies that it solves the vacuum Einstein equation,

$$R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R = 0\,. \tag{5.16}$$

So a (probe) quantum string enforces, on scales $r_c \gg \ell_s \propto \sqrt{\alpha'}$, the dynamics of General Relativity.

If we continue to higher loops, i.e., to higher orders in α', then we start to find deviations from standard GR. For example, including the 1- and 2-loop contributions, the beta function is

$$\beta^{(2)}(g_{\mu\nu}) = \alpha' R_{\mu\nu} + \frac{1}{2}(\alpha')^2 R_{\mu\lambda\rho\sigma} R_\nu{}^{\lambda\rho\sigma} \stackrel{!}{=} 0\,. \tag{5.17}$$

5.2 The non-linear σ-model

In analogy to identifying the (spacetime) metric perturbations as insertions of the graviton vertex operator, we can try to extend the Polyakov action such that the net effect in the path-integral is to generate operator insertions for the Kalb–Ramond ($b_{\mu\nu}$) and dilaton (φ) fields. It turns out that the corresponding action, called a *non-linear σ-model*, takes the form

$$S_\sigma = \frac{1}{4\pi\alpha'}\int_\Sigma d^2\xi\sqrt{-\det\gamma}\left[\left(\gamma^{ab}g_{\mu\nu}(X) + i\epsilon^{ab}b_{\mu\nu}(X)\right)\partial_a X^\mu \partial_b X^\nu + \alpha'\mathcal{R}\varphi(X)\right], \tag{5.18}$$

where we have re-introduced the worldsheet metric γ_{ab} (not to be confused with the graviton tensor). The coupling to the b-field requires the 2d Levi-Civita symbol ϵ^{ab} (the factor i is due to the Euclidean signature), and $\mathcal{R} \equiv \mathcal{R}(\gamma)$ the Ricci scalar on the worldsheet.

Let us briefly explain the name "non-linear σ-model" for this kind of action. Roughly, a σ-model is a field theory for a field $\phi : \Sigma \to \mathcal{M}$ that takes values in a *manifold* $\mathcal{M}$. Traditionally, Σ is the spacetime on which the field theory lives, and $\mathcal{M}$ is called the **target space**. If the target space carries some kind of "linear structure", e.g., a vector space, then the whole physical system is called a *linear* σ-model. For general manifolds, such as generic Riemannian/Lorentzian manifolds, it is then called a "non-linear" σ-model.[10]

Returning to the specific σ-model for the string, we recognize the last term in (5.18) as a generalization to the topological term $\lambda\chi \propto \int d^2\xi\lambda\mathcal{R}$ which gave rise to the string coupling constant. However, in view of the σ-model, we learn that the string coupling is actually not a constant, but rather a spacetime field

[10] Studying the (σ-model) spacetime renormalization effects on the target space $\mathcal{M}$, such as we did to find Einstein's equations, is an instance of so-called "Ricci flow". This mathematical framework has many applications in the study of geometric and topological questions. In particular, it was instrumental in Grigori Perelman's proof of the Poincaré Conjecture.

$e^{\varphi(X)}$! In particular, this means that there are regions in spacetime where the genus expansion for string amplitudes breaks down, namely, when $|\varphi| \to \infty$. In certain superstring theories, such regions can be understood as the vicinity of branes, which are "sources" for φ or g_s. These can be treated fully non-perturbatively in the framework of F-theory. For the upcoming discussion, it is useful to *define* the string coupling by the *asymptotic* value of the dilaton,

$$g_s := e^{\varphi_0}\,, \quad \text{with } \varphi_0 = \lim_{|X|\to\infty} \varphi(X)\,. \tag{5.19}$$

To understand the second term in (5.18) from the spacetime perspective, let us briefly recall that, in classical electrodynamics, the coupling of a charged particle along a worldline $\Gamma : \tau \mapsto X^\mu(\tau)$ to a background electromagnetic field with gauge potential A_μ is described by the term

$$\int_\Gamma A_\mu(X(\tau))\,\dot{X}^\mu(\tau)\,d\tau\,. \tag{5.20}$$

In the language of differential forms, $A = A_\mu dX^\mu$ is a 1-form, and this integral is simply the integral of the pull-back of this 1-form to the worldline. In this form, gauge invariance is simply a consequence of Stokes' theorem,

$$\int_\Gamma d\tau\, A_\mu(X(\tau))\,\dot{X}^\mu(\tau) = \int_\Gamma A \longrightarrow \int_\Gamma (A + d\lambda) = \int_\Gamma A + \lambda|_{\partial\Gamma} = \int_\Gamma A\,, \tag{5.21}$$

with the usual assumption that the gauge parameter λ vanishes at infinity, i.e., the boundaries of Γ.

The generalization from a particle to a string proceeds straightforwardly, by now having a 2-form $b = \frac{1}{2} b_{\mu\nu} dX^\mu \wedge dX^\nu$ instead of a 1-form gauge potential which we pull-back to the worldsheet. This pull-back is precisely

$$b_{\mu\nu}\partial_a X^\mu \partial_b X^\nu d\xi^a \wedge d\xi^b = b_{\mu\nu}\partial_a X^\mu \partial_b X^\nu \epsilon^{ab}\, d^2\xi\,. \tag{5.22}$$

Therefore, the σ-model action describes the string as the charged object under the 2-form gauge field b. The gauge invariance under $b \to b + dc$, i.e., $b_{\mu\nu} \to b_{\mu\nu} + \partial_\mu c_\nu - \partial_\nu c_\mu$ can be checked straightforwardly. The physically invariant *field strength* is then a 3-form,

$$H = db\,, \text{ with } H_{\mu\nu\rho} = \partial_\mu b_{\nu\rho} + \partial_\nu b_{\rho\mu} + \partial_\rho b_{\mu\nu}\,. \tag{5.23}$$

As for the spacetime metric, we can interpret the additional terms to the Polyakov action as interaction terms of the 2d worldsheet theory, and compute their β-functions from having properly renormalized the divergences they introduce. The 1-loop results are the following:

$$\begin{aligned}\beta(g_{\mu\nu}) &= \alpha'\Big(R_{\mu\nu} - \frac{1}{4}H_{\mu\rho\kappa}H_{\nu}{}^{\rho\kappa} + 2\nabla_\mu\nabla_\nu\varphi\Big) + O(\alpha'^2)\,,\\ \beta(b_{\mu\nu}) &= \alpha'\Big(\frac{1}{2}\nabla^\rho H_{\rho\mu\nu} + \nabla^\rho\varphi H_{\rho\mu\nu}\Big)\,,\\ \beta(\varphi) &= \frac{D-26}{6} + \alpha'\Big(\frac{1}{2}\nabla_\mu\varphi\nabla^\mu\varphi - \frac{1}{2}\nabla^2\varphi - \frac{1}{24}H_{\mu\nu\rho}H^{\mu\nu\rho}\Big) + O(\alpha'^2)\,,\end{aligned} \tag{5.24}$$

which have to all vanish as a consistency condition (worldsheet conformal invariance).

The $(D-26)/6$ term in the last line clearly vanishes for the critical setting, but it is worth mentioning its origins. Basically, this is due to the quantum breaking of Weyl invariance in the *free* worldsheet theory. Each coordinate field X^μ contributes $1/6$ to this "Weyl anomaly". The proper path-integral treatment will, however, dictate the addition of "ghost"-fields to implement the gauge-fixing conditions à la Faddeev–Popov. The ghost system in turn contributes $-26/6$ to this anomaly. In the $\beta(\varphi)$-function computation, this anomaly shows up again at leading order.

5.2.1 The low-energy effective action in spacetime

We have seen that the spacetime background fields (g, b, φ) are constrained by the conformal invariance of the string. From a spacetime perspective, these fields should be dynamical, so the constraint equations $\beta = 0$ should be their equations of motions. As you will see on the Problem Sheet, one can indeed derive them from the Euler–Lagrange equations for the *low-energy effective action*

$$S_D^{(S)} = \frac{1}{2\kappa_0^2}\int d^D x\sqrt{-\det g}\,e^{-2\varphi}\left(\frac{2(26-D)}{3\alpha'} + \mathcal{R}(g) - \frac{1}{12}H_{\mu\nu\rho}H^{\mu\nu\rho} + 4\nabla_\mu\varphi\nabla^\mu\varphi\right). \tag{5.25}$$

The superscript (S) stands for "string frame", because the spacetime fields (g, φ) appear as in the string σ-model action.

For spacetime calculations, it is often more convenient to go to the "Einstein frame", by defining $\kappa = \kappa_0 e^{\varphi_0}$, $\tilde{\varphi} = \varphi - \varphi_0$, and raise/lower indices with the metric $\tilde{g} = \exp(\varphi/6)g$:

$$S_D^{(E)} = \frac{1}{2\kappa^2}\int d^D x\sqrt{-\det\tilde{g}}\left(\frac{2(26-D)}{3\alpha'}e^{2(\varphi_0-\tilde{\varphi})} + \mathcal{R}(\tilde{g}) - \frac{e^{-\tilde{\varphi}/3}}{12}|H|^2 - \frac{1}{6}|\nabla_\mu\tilde{\varphi}|^2\right). \tag{5.26}$$

Now the Einstein–Hilbert term takes the canonical form, and so we can identify $\kappa^2 = 8\pi G_N$ with the gravitational coupling, and G_N as Newton's constant in this D dimensional theory. Hence, we can see explicitly in this "frame" (really, its just a convenient choice of field variables) that the spacetime theory is General Relativity, coupled to additional matter fields.

Note that these actions reproduce only the 1-loop β-functions as equations of motions. To capture the additional terms at higher-loops, e.g., the Riemann^2-term in the 2-loop β-function (5.17), the spacetime action also needs α' corrections:

$$S_D = \underbrace{S_D^{(0)}}_{S_D^{(S)}/S_D^{(E)}} + \alpha' \underbrace{S_D^{(1)}}_{\text{4-deriv. terms}} + (\alpha')^2 \underbrace{S_D^{(2)}}_{\text{6-deriv. terms}} + \ldots \tag{5.27}$$

This is an "effective field theory" expansion with cutoff scale $M_S = (\alpha')^{-1/2}$, and can be interpreted as the effective action obtained upon integrating out massive string states.

To summarize, in the spacetime perspective there are several different scales, or expansion parameters:

- The gravitational coupling, $\kappa = \kappa_0 e^{\varphi_0} \sim G_N^{1/2} \sim (M_{\text{Planck}})^{(2-D)/2}$, is related to the Planck scale, and determines the scale above which genuine quantum gravity effects become important.
- The "string scale", $M_s = (\alpha')^{-1/2}$, is the scale where the finite size of the string will start to become important. The net effect are deviations from GR, in form of *higher-derivative* terms in the *classical* effective action.
- The dimensionless ration $M_s/M_{\text{Planck}} \sim e^{2\varphi_0/(D-2)}$ is proportional to (some power of) the string coupling, and it controls the higher contributions in the genus expansion, which are higher-loop orders, i.e., quantum effects from the spacetime perspective.

This (hopefully) clarifies the notion of the *low-energy* effective action: it is an action for dynamics at energy scales $E \ll M_s$ in the limit $M_s/M_{\text{Planck}} \to 0$, so we also suppress spacetime quantum effects. In turn, the worldsheet quantum corrections, controlled by α', are corrections to the classical action, and can be incorporated as deviations from GR.

5.3 Open strings in background fields and D-branes

So far, we have looked at the background fields associated to massless states from the closed string sector. We can repeat the analysis for the open string, and this will lead us to the *Dirac–Born–Infeld (DBI) action*, which is the worldvolume action for **D-branes**.

Recall that for the open string, the massless states at level one assemble into the degrees of freedom of a vector boson. The corresponding vertex operator was

$$\mathcal{W}_{\zeta,k} \sim \int_{\partial\Sigma} d\tau\, \zeta \cdot \dot{X} :e^{ik\cdot X}:, \tag{5.28}$$

with $k^2 = k \cdot \zeta = 0$. Following the same logic around (5.3) as we argued for the general metric perturbations $h_{\mu\nu}(X)$ to be a superposition of plane waves $\gamma_{\mu\nu} :e^{ik\cdot X}:$, we can also identify a spacetime vector field $A_\mu(X)$ whose Fourier modes are the vertex operators $\zeta_\mu :e^{ik\cdot X}:$. Because of the freedom of adding null states, $A_\mu(X) \to A_\mu(X) + \partial_\mu \lambda(X)$ is a gauge symmetry. The corresponding term that one needs to add to the Polyakov action to produce this operator insertion in the path-integral is then

$$S_{\text{endpoints}} = \int_{\partial\Sigma} A_\mu(X)\partial_\tau X^\mu d\tau\,. \tag{5.29}$$

But this is just like coupling the *endpoints* of the open string to an electromagnetic gauge potential A_μ!

More precisely, we had established in Section 2 that the boundary conditions for the coordinate fields X^μ can be separated into Neumann-type and Dirichlet-type. The endpoint of the string can move in the Neumann-type directions $\mu = 0, 1, \ldots, p$ (which we take to include a time-like direction in spacetime), and are fixed in the

Dirichlet-type directions $I = p+1, \ldots, D-1$. We can split the total worldsheet action as $S = S_{\text{Neumann}} + S_{\text{Dirichlet}}$, with

$$\begin{aligned} S_{\text{Neumann}} &= \frac{1}{4\pi\alpha'} \int_\Sigma d^2\xi \, \partial_a X^\mu \partial^a X^\nu \eta_{\mu\nu} + i \int_{\partial\Sigma} A_\mu(X) \dot{X}^\mu d\tau \quad (\mu, \nu = 0, \ldots, p) \,, \\ S_{\text{Dirichlet}} &= \frac{1}{4\pi\alpha'} \int_\sigma d^2\xi \, \partial_a X^I \partial^a X^J \delta_{IJ} \quad (I, J = p+1, \ldots, D-1) \,, \end{aligned} \tag{5.30}$$

where the imaginary factor in the coupling to A_μ is due to being in Euclidean signature on the worldsheet. Furthermore, we have again restricted ourselves to flat spacetime metric for simplicity.

While the Dirichlet part just goes along for a ride, the Neumann part introduces worldsheet interactions with couplings associated to the spacetime vector field A_μ, which can introduce a scale dependence in the renormalized quantum theory. So we can again impose conformal invariance as a consistency condition to constrain the spacetime dynamics of A_μ and its associated field strength $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$. Going through the 1-loop computations and properly renormalize the divergences, the vanishing of the β-function turns out to imply

$$\partial_\mu \left(F_{\nu\rho} \frac{\eta^{\mu\nu}}{1 - 4\pi\alpha'^2 F^2} \right) = 0 \quad \text{(all indices from 0 to } p) \,. \tag{5.31}$$

Note that, to zeroth order in α', this is just the familiar equation of motion $\partial^\mu F_{\mu\rho} = 0$ for electromagnetism.

This equation of motion has been known for a long time, and arises in the *Born–Infeld (BI) model* which is a non-linear alternative to Maxwell's theory.[11] The corresponding spacetime action is

$$S = -T_p \int d^{p+1}x \sqrt{-\det(\eta_{\mu\nu} + 2\pi\alpha' F_{\mu\nu})} \,, \tag{5.32}$$

which is an integral only over the Neumann-type spacetime directions. For field strengths that are small compared to the string tension, $|F_{\mu\nu}| \ll 1/\alpha'$, this action can be expanded as

$$S = -T_p \int d^{p+1}x \left(1 + \frac{(2\pi\alpha')^2}{4} F_{\mu\nu} F^{\mu\nu} + O(\alpha'^2) \right) \,, \tag{5.33}$$

[11] The Born–Infeld model resolves the (classical) divergence of the electric field at the location of a point-charge, by modifying the equations of motions for large field values $|F_{\mu\nu}| \approx 1/\alpha'$; in the original model, $1/\alpha'$ is just a large parameter of mass dimension 2 that controls the deviation from Maxwell's equations, which in string theory gets identified with the (inverse) string tension.

which gives the $(p+1)$-dimensional volume, and then the Maxwell action, $\int d^{p+1}x\,\frac{1}{4g}F_{\mu\nu}F^{\mu\nu}$.

In the full D-dimensional spacetime, this describes the propagation of $U(1)$ gauge field along a $(p+1)$-dimensional submanifold. This submanifold is the *worldvolume* of a Dp-brane, which is an extended object with tension T_p. Notice that the combination of brane and string tension, $((2\pi\alpha')^2T_p)^{-1} = T^2/T_p = g$, gives the $U(1)$ gauge coupling.

5.3.1 The DBI action

One can generalize the analysis now to non-flat spacetimes, including non-trivial backgrounds for the dilaton φ and the Kalb–Ramond 2-form field b. Of course, there is a corresponding worldsheet computation that can be carried out that will justify the resulting spacetime, or rather worldvolume description. In the following, we will give some more intuitive explanation for the form of the effective field theory action.

First, we should introduce again "intrinsic" coordinates ξ^a, $a = 0, \ldots, p$, on the worldvolume, with spacetime coordinate fields $X^\mu(\xi)$ that describe the embedding of the brane into spacetime. Then, the $\eta_{\mu\nu}$-term in the BI action (5.32) should be replaced with the pull-back metric

$$h_{ab} = \partial_a X^\mu \partial_b X^\nu g_{\mu\nu}\,. \tag{5.34}$$

In fact, switching off A, φ and b, the worldvolume action $S = -T_p\int d^{p+1}\xi\sqrt{-\det(h_{ab})}$ is the obvious generalization of the Nambu–Goto action (2.13) for the string.

The gauge field $A = A_a d\xi^a$ is now a 1-form on the worldvolume, with corresponding field strength $F_{ab} = \partial_a A_b - \partial_b A_a$. Combining with the pull-back metric, we obtain the generalization that historically is known as the *Dirac–Born–Infeld (DBI) action*,

$$S = -T_p \int d^{p+1}\xi\,\sqrt{-\det(h_{ab} + 2\pi\alpha' F_{ab})}\,. \tag{5.35}$$

Turning to the Kalb–Ramond field, it is helpful to look at the classical string worldsheet action including both the b field and the gauge field A. The relevant terms are

$$\frac{1}{4\pi\alpha'}\int_\Sigma d^2\xi\, \epsilon^{ab}\,\underbrace{\partial_a X^\mu \partial_b X^b b_{\mu\nu}}_{=B_{ab}} + \int_{\partial\Sigma} A = \frac{1}{2\pi\alpha'}\int_\Sigma B + \int_{\partial\Sigma} A\,, \tag{5.36}$$

written in differential-form notation, with $B = \frac{1}{2}B_{ab}d\xi^a \wedge d\xi^b$ the pull-back of the spacetime 2-form b onto Σ (strictly speaking, A here is also a pulled-back 1-form). Now consider the effect of a spacetime gauge transformation,

$$b \to b + dc \quad \overset{\text{pull-back}}{\Longrightarrow} \quad B \to B + dC \tag{5.37}$$

for c a 1-form, which pulls-back to C on Σ. This shifts the first term by a total derivative, which vanishes automatically for the closed string, so we did not bother discussing it. However, for the open string, this results in a boundary contribution; using Stokes' theorem in differential-form notation, this is easily shown to be

$$\int_\Sigma B \to \int_\Sigma (B + dC) = \int_\Sigma B + \int_\Sigma dC = \int_\Sigma B + \int_{\partial\Sigma} C \tag{5.38}$$

The worldsheet action is then only invariant if the $U(1)$ gauge fields also shifts with the gauge parameter C, namely as

$$A \to A - \frac{1}{2\pi\alpha'}C\,. \tag{5.39}$$

A symmetry structure where a lower-degree gauge field, in this case A transform under a gauge transformation of a higher-degree gauge field, in this case b, is nowadays known as a *higher-group* symmetry. It is quite remarkable that this relatively recently developed concept has been in plain sight within string theory for decades!

Returning to our spacetime action, this means that, in the presence of a non-trivial b-field background, it is not the $U(1)$ field strength F, but rather the combination

$$B_{ab} + 2\pi\alpha' F_{ab} \tag{5.40}$$

which is the gauge invariant anti-symmetric tensor that replaces F_{ab} under the square-root in (5.35). Physically, it is basically the statement that the charged objects under b and A in spacetime are not independent of each other, but rather just the same one: it's the string and its endpoints!

For the dilaton, we do not have any slick spacetime argument for how to include it correctly, so we will now just state the full DBI-action:

$$S = -T_p \int d^{p+1}\xi \, e^{\varphi_0 - \varphi} \sqrt{-\det(h_{ab} + 2\pi\alpha' F_{ab} + B_{ab})} \,. \tag{5.41}$$

This means that the value of the dilaton actually modifies the *effective* brane tension $T_p e^{\tilde{\varphi}}$. Implicitly, this relates the "asymptotic" brane tension T_p to the asymptotic string coupling, $T_p \sim 1/g_s = e^{-\varphi_0}$.

5.3.2 Non-Abelian gauge fields from brane stacks

As we have seen, open strings ending on a Dp-brane give rise to a $U(1)$ gauge field in the $(p+1)$-dimensional worldvolume. For completeness, let us also mention that string theory provides a straightforward way to obtain non-Abelian gauge fields, namely by *stacking* branes.

More precisely, this procedure relies on the ingenious realization by Polchinski that branes are not just booking keeping devices for open string endpoints, but are dynamical objects themselves (that are non-perturbative in g_s, as their tension behave as $T_p \sim 1/g_s$). This means that we can talk about moving several Dp-branes relative to each other. In particular, we can put N of them on top of each other, so that they fill the same worldvolume.

Though it may seem a bit odd a first, it turns out that the correct way to describe open string on such a stack is to assign labels $n = 1, \ldots, N$ to *each* endpoint. Schematically, in the coupling (5.29), we could divide up the boundary $\partial\Sigma$ into two parts (for the two endpoints) and assign to each one a label. The result is the same as if we just declare that the gauge field A now carries two additional labels:

$$(A_a)^m{}_n \,. \tag{5.42}$$

Note that these still should be thought of as coming from level-one states $|\zeta; k\rangle^m{}_n$ of the open string, now "starting" on the m-th and "ending" on the n-th brane, but which are still massless vector bosons.

The labels (m, n) are called Chan–Paton labels, and can be shown to endow the set of vector fields $(A_a)^m{}_n$ with the structure of the adjoint representation of

$U(N)$. This gives the hint that the spacetime (or worldvolume) interpretation of these open string modes should be that of a $U(N)$ gauge field. Indeed, the set of gauge transformations determined by the adding null states in the open string Hilbert space with Chan–Paton labels turns out to be the adjoint action of $U(N)$ on the states $|\zeta; k\rangle^m{}_n$. The corresponding low-energy effective action on the brane worldvolume then includes the standard Yang–Mills form, which we won't write out here explicitly.

Note that with oriented string, we are limited to $U(N)$ gauge groups. However, as already mentioned earlier, this limitation is lifted once we include non-oriented strings, which can be equivalently thought of as having brane-like objects which "reverse" the orientation of the strings that end on them. These so-called orientifold-planes, or Op-planes for short, are more exotic branes, in that they can have *negative* tension. But we can still stack them on Dp-branes, in which case the worldvolume can exhibit $SO(N)$ or $Sp(N)$ gauge groups. In non-perturbative regions (where g_s is not small, and thus are most naturally described in F-theory), one can have exceptional groups E_n and more exotic non-simply laced Lie groups.

6 Circle Compactification and T-Duality

The fact that string theory is only well-defined in more than four spacetime dimensions is clearly troublesome to anyone who is interested in understanding *our* universe. The most straightforward process method to bridge this conceptual gap is "compactification": we take our higher-dimensional spacetime ($D = 26$ for the bosonic string) $\mathcal{M}_D$ to be a product $\mathcal{M}_D = \mathbb{R}^{1,D-1-d} \times X_d$, and postulate that the size of X_d is "small" enough, so that we can hide any dynamics happening "inside" X_d in an effective field theory description on $\mathbb{R}^{1,D-1-d}$, whose cutoff is the (inverse length) scale of X_d.

The idea of compactification predates string theory, and has already been studied in the 1920s by Kaluza and Klein. In these early works, it was shown that, classically, the *circle* or S^1 *reduction* of 5d general relativity, i.e., taking

$$\mathcal{M}_5 = \mathbb{R}^{1,3} \times S^1 , \tag{6.1}$$

leads to 4d GR plus an electromagnetic field satisfying Maxwell's equations, and an electrically charged scalar field in the limit of small circle radius. Moreover, the circle radius determines the charge-to-mass ratio of the quantum excitations of the scalar field.

The methods of Kaluza and Klein apply also to higher-dimensional internal manifolds X_d, and more general (quantum) field theories in $\mathcal{M}_D$. In such *Kaluza–Klein (KK) reductions*, the geometric and topological properties of X_d determine the parameters of the low-energy effective theory in $\mathbb{R}^{1,D-1-d}$. In string compactifications, this interplay becomes even richer and more intricate, and has led to cross-fertilization between research in geometry/topology and physics.

The key difference between a field-theoretic KK-reduction and the compactification of string theory is again due to the finite size of the string, which sees the internal geometry in a different way than particles. Below, we will study explicitly an S^1-compactification, and see how the genuinely stringy phenomenon of "T-duality" arises from the σ-model description.

6.1 KK-reduction on a circle

Let us first understand the circle reduction in the field theory context. To illustrate the role of the S^1, we keep the remaining directions to be flat, i.e., $\mathcal{M}_D = \mathbb{R}^{1,D-2} \times S^1$, with coordinates $X^\mu = (X^i, X^{D-1})$ where $i = 0, ..., D-2$. For further simplicity, we consider for now only field configurations that are independent of the circle direction of radius r, $X^{D-1} \sim X^{D-1} + 2\pi r$. Then, the field components of the metric, a 2-form field $b_{\mu\nu}$, and a scalar φ in D dimensions decompose into

$$\begin{aligned} g_{\mu\nu}(X) &\longrightarrow \{g_{ij}(X^i),\, g_{i,D-1}(X^i),\, g_{D-1,D-1}(X^i)\}\,, \\ b_{\mu\nu}(X) &\longrightarrow \{b_{ij}(X^i),\, b_{i,D-1}(X^i)\}\,, \\ \varphi(X) &\longrightarrow \varphi(X^i)\,. \end{aligned} \tag{6.2}$$

This gives a set of fields $(g_{ij}(X^i),\, b_{ij}(X^i),\, \varphi(X^i))$ that are the $(D-1)$-dimensional versions of the graviton, 2-form, and the dilaton fields. In addition, we also have two vector fields, $g_{i,D-1}$ and $b_{i,D-1}$ in $\mathbb{R}^{1,D-1}$, which are called the *graviphoton* and the *Kalb–Ramond photon*, and an additional scalar field $g_{D-1,D-1}$, the *radion*.

The two vector fields are actually 1-form fields, i.e., $U(1)$ gauge fields, in $\mathbb{R}^{1,D-2}$. The graviphoton inherits this gauge symmetry from the D-dimensional diffeomorphisms, $\delta_\epsilon g_{\mu\nu} = \nabla_\mu \epsilon_\nu + \nabla_\nu \epsilon_\mu$: by taking $\epsilon_{D-1} = \lambda(X^i)$ for any (scalar) function λ, and other ϵ-directions to be 0, we have

$$A_i := g_{i,D-1} \to g_{i,D-1} + \delta g_{i,D-1} = A_i + \nabla_i \lambda = A_i + \partial_i \lambda\,. \tag{6.3}$$

Similarly, the Kalb–Ramond photon (or Kalb–Ramond $U(1)$) also inherits the gauge symmetry

$$\tilde{A}_i := b_{i,D-1} \longrightarrow \tilde{A}_i + \partial_i \tilde{\lambda}\,, \tag{6.4}$$

by picking the gauge parameter as $c_{D-1} = \tilde{\lambda}(X^i)$ for the D-dimensional symmetry $\delta_c b_{\mu\nu} = \partial_\mu c_\nu - \partial_\nu c_\mu$. These gauge symmetries are correctly reflected in the $(D-1)$-dimensional action. To obtain that (see Problem Sheet), one follows, more carefully, the Kaluza–Klein ansatz,

$$\begin{aligned} g_{\mu\nu} dX^\mu dX^\nu &= g_{ij} dX^i dX^j + e^{2\sigma}(dX^{D-1} + A_i dX^i)^2\,, \quad \text{with } g_{D-1,D-1} = e^{2\sigma}\,, \\ b_{\mu\nu} dX^\mu \wedge dX^\nu &= b_{ij} dX^i \wedge dX^j + \tilde{A}_i dX^i \wedge dX^{D-1}\,, \\ \phi \equiv \varphi^{(D-1)} &= \varphi - \frac{\sigma}{2}\,, \end{aligned} \tag{6.5}$$

and plugs this into the D-dimensional effective action.

Now we can start adding non-trivial profiles for the D-dimensional fields on the circle. For that, one can simply use a Fourier decomposition in the periodic coordinate X^{D-1}, in which case the $(D-1)$-dimensional fields we have considered so far are the zero-modes of this decomposition. For illustration, consider the dilaton field:

$$\varphi(X^\mu) = \sum_{n\in\mathbb{Z}} e^{inX^{D-1}/r}\, \varphi_n(X^i) \qquad (\text{with } \varphi_n^* = \varphi_{-n}). \tag{6.6}$$

From the kinetic term in the D-dimensional action, we obtain, after integrating over the circle,

$$\begin{aligned} \int d^D X\, \partial_\mu \varphi \partial^\mu \varphi &= \int d^D X\, ((\partial_i \varphi)^2 + (\partial_{D-1}\varphi)^2) \\ &= 2\pi r \int d^{D-1}X \sum_{n\in\mathbb{Z}} \left(\partial_i \varphi_n \partial^i \varphi_{-n} + \frac{n^2}{r^2} |\varphi_n|^2 \right)\,. \end{aligned} \tag{6.7}$$

That is, a single scalar field in $\mathbb{R}^{1,D-2} \times S^1$ looks like an infinite tower of scalar fields on $\mathbb{R}^{1,D-2}$ labelled by the *Kaluza–Klein level* n, with mass

$$M_n^2 = \frac{n^2}{r^2}\,. \tag{6.8}$$

The level n field in $(D-1)$ dimensions arise from D-dimensional field configurations that have a definite momentum $p^{(D)} = n/r$ along the circle. In the limit $r \to 0$, or equivalently, at energy scales $E \ll 1/r \sim M_{\mathrm{KK}}$, all KK modes with $n \neq 0$ decouple, and the low-energy effective theory in $\mathbb{R}^{1,D-2}$ just contains the massless KK zero-modes.

Finally, re-consider the effect of diffeomorphisms $\epsilon_{D-1} = \lambda(X^i)$ that gave rise to the gauge transformation (6.3) for the graviphoton A_i. At the level of the spacetime coordinates, this diffeomorphism is simply a "rotation" $X^{D-1} \to X^{D-1} + \lambda(X^i)$ on the S^1. From the Fourier decomposition (6.6), we can see that this also has an effect on the KK-modes:

$$\varphi_n \to e^{in\lambda/r}\varphi_n \,. \tag{6.9}$$

This tells us that the n-th KK-mode, in addition of having mass $M_n \propto n$, also has *electric charge* $Q_n \propto n$ under the graviphoton A_i in $(D-1)$-dimensions.

On the other hand, there are no charged objects under the Kalb–Ramond photon $\tilde{A}_i$ coming from the (field-theoretic) KK-reduction of the D-dimensional effective action. We will see momentarily how the perspective from the string worldsheet changes this.

6.2 Strings on a circle

Intuitively, strings differ from particles in S^1-compatifications by being able to *wind* around the circle. In the remaining spacetime directions, this gives another type of particle-like objects that are not present in the field-theoretic KK-reduction; these will turn out to now be charged under the Kalb–Ramond $U(1)$.

To see this in detail, we turn to the non-linear σ-model (5.18) of the worldsheet, and specify the target space to be the $D = 26$ dimensional manifold $\mathbb{R}^{1,24} \times S^1$. In the classical theory, the periodic target space direction has two immediate effects. First, it leads to a quantized spacetime momentum in the mode expansions,

$$p^{25} = \frac{n}{r}\,, \quad n \in \mathbb{Z}\,. \tag{6.10}$$

Moreover, for the closed string, we can relax the boundary condition $X(\xi^1 + 2\pi) = X(\xi^1)$ to

$$X^{25}(\xi^1 + 2\pi) = X^{25} + 2\pi w r\,, \tag{6.11}$$

where $w \in \mathbb{Z}$ encodes how many times the string winds around S^1. For this reason, it is also called the *winding number*.

Classically, the momentum and winding numbers label disconnected components of phase space. That is, the mode expansion differs by discrete parameters (n, w),

$$X^{25} = x^{25} + \alpha' p^{25} \xi^0 + wr\xi^1 + \text{ oscillators} = X_L^{25} + X_R^{25}\,, \tag{6.12}$$

with the left- and right-movers (in lightcone coordinates $\xi^\pm = \xi^0 \pm \xi^1$) given by

$$\begin{aligned} X_L^{25}(\xi^+) &= \frac{1}{2}x^{25} + \frac{1}{2}\alpha' \underbrace{\left(\frac{n}{r} + \frac{wr}{\alpha'}\right)}_{=:p_L} \xi^+ + i\sqrt{\frac{\alpha'}{2}} \sum_{m\neq 0} \frac{1}{m}\tilde{\alpha}_m^{25} \exp(-im\xi^+)\,, \\ X_R^{25}(\xi^+) &= \frac{1}{2}x^{25} + \frac{1}{2}\alpha' \underbrace{\left(\frac{n}{r} - \frac{wr}{\alpha'}\right)}_{=:p_R} \xi^- + i\sqrt{\frac{\alpha'}{2}} \sum_{m\neq 0} \frac{1}{m}\alpha_m^{25} \exp(-im\xi^-)\,. \end{aligned} \tag{6.13}$$

These are mostly identical to the original mode expansion (2.39), but differ crucially in the zero-mode oscillator part,

$$\begin{aligned} \alpha_0^{25} &:= \sqrt{\tfrac{\alpha'}{2}}p_R \neq \sqrt{\tfrac{\alpha'}{2}}p_L =: \tilde{\alpha}_0^{25}\,, \\ \implies \tilde{\alpha}_0^{25} + \alpha_0^{25} &= \sqrt{2\alpha'}p^{25} = \sqrt{2\alpha'}\tfrac{n}{r}\,, \quad \tilde{\alpha}_0^{25} - \alpha_0^{25} = \sqrt{\tfrac{2}{\alpha'}}wr\,, \end{aligned} \tag{6.14}$$

The quantization proceeds with these modified zero-mode operators. More precisely, while we previously had $\tilde{\alpha}_0^{25} = \alpha_0^{25}$ as quantum operators, now they differ by a "c-number". This does not affect commutators of the oscillators, so the oscillator Fock space is identical to (3.4), and organizes the states into levels labelled by the eigenvalues under N and $\tilde{N}$. However, whereas the oscillator vacuum $|0,\tilde{0};k^{(26)}\rangle$ was previously labelled by a 26d momentum vector k, it is now labelled by a 25d momentum vector and the integers (n, w) that describes the basis in momentum space,

$$\begin{aligned} & |0,\tilde{0};k,n,w\rangle\,, \quad k \in \mathbb{R}^{1,24}\,, \\ \text{with} \quad & p^i|0,\tilde{0};k,n,w\rangle = k^i|0,\tilde{0};k,n,w\rangle\,, \\ & p_{L/R}^{25}|0,\tilde{0};k,n,w\rangle = \left(\tfrac{n}{r} \pm \tfrac{wr}{\alpha'}\right)|0,\tilde{0};k,n,w\rangle\,. \end{aligned} \tag{6.15}$$

Likewise, the expressions for the Virasoro generators remain the same, but it is more appropriate to reorganize them using $v \cdot v' := v_i\, v^i$ as the 25d scalar product,

$$
\begin{aligned}
L_0 &= \frac{1}{2}\big(\alpha_0 \cdot \alpha_0 + (\alpha_0^{25})^2\big) + \overbrace{\sum_{\ell>0} \alpha_{-\ell} \cdot \alpha_\ell + \sum_{\ell>0} \alpha_{-\ell}^{25}\alpha_\ell^{25}}^{N}\,, \\
L_m &= \frac{1}{2}\sum_\ell \alpha_{m-\ell} \cdot \alpha_m + \frac{1}{2}\sum_\ell \alpha_{m-\ell}^{25}\alpha_\ell^{25}\,,
\end{aligned}
\tag{6.16}
$$

and similarly for the left-movers $\tilde{L}_m$. Because the structures are identical, the normal ordering and criticality discussions go through as before, but now the mass-shell and level-matching conditions are modified:

$$
\begin{aligned}
L_0 - 1 &= \frac{\alpha'}{4}(p \cdot p + p_R^2) + N - 1 \overset{!}{=} 0 \;\Rightarrow M_{(25)}^2 = -p \cdot p = p_R^2 + \frac{4}{\alpha'}(N-1)\,, \\
\tilde{L}_0 - 1 &= \frac{\alpha'}{4}(p \cdot p + p_L^2) + \tilde{N} - 1 \overset{!}{=} 0 \;\Rightarrow M_{(25)}^2 = -p \cdot p = p_L^2 + \frac{4}{\alpha'}(\tilde{N}-1)\,,
\end{aligned}
\tag{6.17}
$$

which after reorganizing becomes

$$
\begin{aligned}
M_{(25)}^2 &= \frac{1}{2}(p_R^2 + p_L^2) + \frac{2}{\alpha'}(N + \tilde{N} - 2) = \frac{n^2}{r^2} + \frac{w^2 r^2}{\alpha'^2} + \frac{2}{\alpha'}(N + \tilde{N} - 2)\,, \\
N - \tilde{N} &= \frac{\alpha'}{4}(p_L^2 - p_R^2) = n\, w\,.
\end{aligned}
\tag{6.18}
$$

Note that in the mass-shell condition (first line), the mass in 25 dimensions receives both contributions from the momentum along the circle (same as in the KK-reduction), as well as the winding modes (these are now genuine stringy contributions).

String spectrum in 25d

First, we can recover the familiar massless states in 25d, which we had already found in the Kaluza–Klein reduction, from level one string states with zero momentum and winding along the circle, i.e., $(n, w) = (0, 0)$:

- The graviton, Kalb–Ramond field, and the dilaton are of the form $\alpha_{-1}^i \tilde{\alpha}_{-1}^j |0, \tilde{0}; k, 0, 0\rangle$.

- The graviphoton A_i and the Kalb–Ramond photon $\tilde{A}_i$ are associated to states of the form $(\alpha^i_{-1}\tilde{\alpha}^{25}_{-1}+\alpha^{25}_{-1}\tilde{\alpha}^i_{-1})|0,\tilde{0};k,0,0\rangle$ and $(\alpha^i_{-1}\tilde{\alpha}^{25}_{-1}-\alpha^{25}_{-1}\tilde{\alpha}^i_{-1})|0,\tilde{0};k,0,0\rangle$, respectively.
- The radion is associated to $\alpha^{25}_{-1}\tilde{\alpha}^{25}_{-1}|0,\tilde{0};k,0,0\rangle$.

As before, their gauge symmetries and correct number of degrees of freedom is guaranteed by the physicality constraints imposed by the Virasoro generators. In the correspondence between spacetime fields and worldsheet vertex operators, the two photons are associated to vertex operators

$$V_{\pm}[\zeta,k] = \frac{1}{\sqrt{2\alpha'}}\int d^2\xi\,\zeta_i(\partial_+X^i\partial_-X^{25}\pm\partial_+X^{25}\partial_-X^i):e^{ik\cdot X}: \,. \tag{6.19}$$

The states with non-trivial circle momentum and winding are in general not massless, due to the modified mass-shell formula. The simplest example would be a "winding tachyon" $|0,\tilde{0};k,n=0,w\rangle$ with $-k^2=(wr/\alpha')^2-4/\alpha'$. We can now compute a tree-level 3-point scattering for two such winding tachyons with the Kalb–Ramond photon,

$$\begin{aligned}\mathcal{A} &\sim \langle 0,\tilde{0};-k_3,0,w|\,V_-[\zeta,k_2]\,|0,\tilde{0};k_1,0,w\rangle\\ &= \tfrac{1}{\sqrt{2\alpha'}}\langle 0,\tilde{0};-k_3,0,w|\,(\zeta\cdot\tilde{\alpha}_0)\,\alpha_0^{25}-(\zeta\cdot\alpha_0)\,\tilde{\alpha}_0^{25}\,|0,\tilde{0};k_1+k_2,0,w\rangle\\ &= \tfrac{1}{\sqrt{2\alpha'}}\zeta\cdot(k_1+k_2)\langle 0,\tilde{0};-k_3,0,w|\,\alpha_0^{25}-\tilde{\alpha}_0^{25}\,|0,\tilde{0};k_1+k_2,0,w\rangle\\ &= -\frac{wr}{\alpha'}\,\zeta\cdot(k_1+k_2)\,\delta(k_1+k_2+k_3) = \frac{wr}{\alpha'}\,(\zeta\cdot k_3)\,\delta(k_1+k_2+k_3)\,.\end{aligned} \tag{6.20}$$

In the 25d spacetime, this is the tree-level contribution to an emission/absorption of a Kalb–Ramond photon from a winding tachyon with winding number w. As usual in spacetime gauge theories, this means that the winding tachyon is charged under the Kalb–Ramond $U(1)$-field $\tilde{A}_i$ with charge proportional to wr/α'.

More generally, one can show that, for states with general circle momentum and winding numbers (n,w), their graviphoton-charge is $p_L+p_R\sim\frac{n}{r}$, and their Kalb–Ramond $U(1)$-charge is $p_L-p_R\sim\frac{wr}{\alpha'}$. The latter is quite obvious following our discussion in the previous section, where we saw that the entire string is electrically charged under $b_{\mu\nu}$ in 26 dimensions. This gives a full charge *lattice* of in general massive states under the $U(1)\times U(1)$ gauge symmetry in 25d.

On the Problem Sheet, you will further show an interesting phenomenon that only occurs in string compactifications, and not in "standard" field-theory reductions.

Namely, for special values of the circle radius r, the momentum and winding contributions can cancel against the oscillator contributions in the mass-shell formula, to give additional massless states. These massless states can *enhance* the $U(1) \times U(1)$ gauge symmetry, depending on r up to $SU(2) \times SU(2)$ at $r = \sqrt{\alpha'}$. Such a non-Abelian gauge enhancement is again a genuinely stringy effect that is absent in field theoretic compactifications.

6.3 T-duality

The circle compactification of the bosonic string exhibits another remarkable feature which is the prototype of a string duality. Such dualities generally relate seemingly different descriptions of the same underlying physical system. In the present case, the differently looking descriptions are compactifications on circles with radii r_1 and $r_2 = \alpha'/r_1$.

As Kaluza–Klein reductions, different circle radii would lead to genuinely different physical models in lower dimensions, since the KK-tower would have different masses. However, if we look at the mass-shell and level matching condition (6.18) for string states on an S^1, we see that they are invariant under the exchange, known as a **T-duality** transformation,

$$r \rightarrow r' = \alpha'/r\,, \quad (n,w) \rightarrow (n',w') = (w,n)\,, \tag{6.21}$$

i.e., the roles of momentum and winding modes are exchanged as go from a “small” circle to a “large” circle and vice versa.

Consider first the limit $r \rightarrow \infty$. This clearly “decompactifies” the S^1-direction, thus giving back the 26d theory. At the level of the mass spectrum, this is signaled by the momentum modes labelled by n developing a continuum of massless states. Naively, the opposite limit $r \rightarrow 0$ should correspond to the limit of obtaining a genuine 25d theory, by decoupling the massive momentum modes. However, the string also has winding modes labelled by w that become light and form a continuum in this limit, indicating another “dual” space dimensions that opens up.

To make this more precise, let us inspect more closely the worldsheet CFT. Here, the relevant data are the Hilbert space $\mathcal{H}$ of states $\prod_\ell \alpha_{-n_\ell}|0;k,n,w\rangle$, and the separation into (un-)physical and null states by the Virasoro generators. These are in turn derived from the stress tensor $T_{\pm\pm} = -\partial_\pm X_\mu \partial_\pm X^\mu$. Any “transformation” that results in a re-organization of this data that do not modify the physicality structure on $\mathcal{H}$, i.e., leave the Virasoro algebra invariant, will therefore not change the actual physical system, but instead just provide a change in the *description*.

We can easily verify that the T-duality transformation (6.21) is precisely of this type. To see this, note that in the mode expansion (6.13), the transformation acts on the momenta,

$$\begin{aligned} &p_L = \frac{n}{r} + \frac{wr}{\alpha'} \longrightarrow \frac{wr}{\alpha'} + \frac{n}{r} = p_L\,, \quad p_R = \frac{n}{r} - \frac{wr}{\alpha'} \longrightarrow \frac{wr}{\alpha'} - \frac{n}{r} = -p_R\,, \\ &\Rightarrow \quad \alpha_0^{25} \longrightarrow -\alpha_0^{25}\,, \quad \tilde{\alpha}_0^{25} \longrightarrow \tilde{\alpha}_0^{25}\,. \end{aligned} \tag{6.22}$$

Motivated by this, we *define* a new "coordinate" fields,

$$Y^i = X^i \text{ for } i = 0, \ldots, 24\,, \quad Y^{25}(\xi^+, \xi^-) = X_L^{25}(\xi^+) - X_R^{25}(\xi^-)\,. \tag{6.23}$$

If we move around the spatial coordinate on the worldsheet, i.e., $\xi^1 \longrightarrow \xi^1 + 2\pi$, we see from (6.13) that

$$Y^{25}(\xi^0, \xi^1 + 2\pi) = Y^{25}(\xi^+ + 2\pi, \xi^- - 2\pi) = Y^{25}(\xi^\pm) + \frac{2\pi\alpha' n}{r}\,, \tag{6.24}$$

so this is a $2\pi r'$-periodic target space direction. From these fields, we can equivalently extract from the mode expansions the Heisenberg pairs (x, p) and the oscillators $(\alpha_m, \tilde{\alpha}_m)$.[12] Moreover, because

$$\partial_+ Y^{25} = \partial_+ X_L^{25} = \partial_+ X^{25}\,, \quad \partial_- Y^{25} = -\partial_- X_R^{25} = -\partial_- X^{25}\,, \tag{6.25}$$

the stress tensor is invariant,

$$\begin{aligned} T'_{++} &:= \partial_+ Y \cdot \partial_+ Y = \partial_+ X \cdot \partial_+ X = T_{++}\,, \\ T'_{--} &:= \partial_- Y \cdot \partial_- Y = \partial_- X \cdot \partial_- X = T_{--}\,. \end{aligned} \tag{6.26}$$

Therefore, the Fourier modes which give rise to the Virasoro generators L_n and $\tilde{L}_m$ are also the same.

This construction explicitly shows that in the limit $r \longrightarrow 0$ for the circle with coordinate X^{25}, the "T-dual" circle with coordinate Y^{25} becomes infinitely large, $r' = \alpha'/r \longrightarrow \infty$, as the continuum of massless states have indicated above. This also means that there is, in a very concrete sense, a "smallest" circle radius that an S^1-compactification of the bosonic string can have. This is the self-dual radius $r_c = \sqrt{\alpha'}$, which is precisely the case of maximal gauge enhancement in the non-compact 25 dimensions.

[12] There is a subtlety regarding the zero-mode in Y^{25}, in that the center-of-mass coordinate x^{25} cancels out classically. However, this ultimately does not matter, because the CFT defined by Y^μ with radius $r' = \alpha'/r$ is the same as that defined by X^μ with radius r.

Open strings under T-duality

For T-duality to work in the closed string sector, it was crucial that the string can wind around the circle. For open strings, there is no invariant notion of a winding number (see Figure 15), and we are only left with the circle momentum $p^{25} = n/r$ as a discrete parameter in the mode expansions. So what is T-duality "exchanging" for the open string?

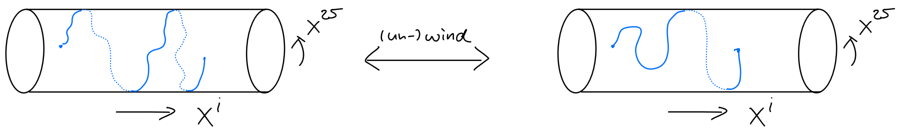


Fig. 15: An open string on a circle can unwind if its endpoints can move freely in the circle direction X^{25}.

To answer this question, we revisit the splitting into left- and right-movers, which we did not bother to write down around (2.44). The general ansatz that solves the equations of motion $\partial_+\partial_- X^\mu = 0$ is still

$$X^\mu_L(\xi^+) = \tfrac{1}{2}x^\mu + \alpha' p^\mu \xi^+ + i\sqrt{\tfrac{\alpha'}{2}} \sum_{m\neq 0} \tfrac{1}{m}\tilde{\alpha}^\mu_m e^{-im\xi^+}\,, \tag{6.27}$$

$$X^\mu_R(\xi^-) = \tfrac{1}{2}x^\mu + \alpha' p^\mu \xi^- + i\sqrt{\tfrac{\alpha'}{2}} \sum_{m\neq 0} \tfrac{1}{m}\alpha^\mu_m e^{-im\xi^-}\,, \tag{6.28}$$

with the relations between the coefficients now fixed by the boundary conditions:

- Neumann boundary conditions (NN), $\partial_{\xi^1} X^\mu|_{\xi^1=0,\pi} = 0$, require $\alpha^\mu_m = \tilde{\alpha}^\mu_m$;
- Dirichlet boundary conditions (DD), $X^\nu|_{\xi^1=0,\pi} = c^\nu$, require $x^\nu = c^\nu$, $p^\nu = 0$, and $\alpha^\nu_m = -\tilde{\alpha}^\nu_m$.

Now let us assume that we have Neumann boundary conditions (NN) in the $2\pi r$-periodic spacetime coordinate $X^{25} = X^{25}_L + X^{25}_R$ is $2\pi r$ (so $p^{25} = n/r$), meaning that both string endpoints can move freely on the circle. Then, we follow the same procedure as for the close string and *define*

$$\begin{aligned} Y^{25} &= X^{25}_L - X^{25}_R = \alpha' p^{25}(\xi^+ - \xi^-) + i\sqrt{\tfrac{\alpha'}{2}} \sum_{m\neq 0} \tfrac{1}{m}\alpha^{25}_m (e^{-im\xi^+} - e^{-im\xi^-}) \\ &= 2\alpha' p^{25}\xi^1 + i\sqrt{2\alpha'} \sum_{m\neq 0} \tfrac{1}{m}\alpha^{25}_m e^{-im\xi^0} \sin(m\xi^1)\,. \end{aligned} \tag{6.29}$$

As before, this is a $2\pi r' = 2\pi\alpha'/r$-periodic direction, which should be thought of as the "T-dual circle". Because we have

$$Y^{25}(\xi^1 = 0, \pi) = 0 \mod \underbrace{\frac{2\pi n\alpha'}{r}}_{=2\pi\alpha' p^{25}}, \quad n \in \mathbb{Z}, \tag{6.30}$$

the open string now has (DD) boundary conditions in the T-dual circle direction, with no possibility to move its center of mass along that direction. Hence, it has no momentum along the circle Y^{25}. Moreover, in the T-dual setting, the term $2\alpha' p^{25}\xi^1 = 2n\frac{\alpha'}{r}\xi^1$ in the mode expansion of Y^{25} now counts the winding of the string. It is straightforward to check that, starting from Dirichlet boundary conditions on a circle, the T-dual setting will have Neumann boundary conditions on the T-dual circle, and the winding number becomes the momentum number. This is summarized in Figure 16. Again, because from the worldsheet perspective this is just a rearrangement of the CFT data, the physics of both descriptions must agree.

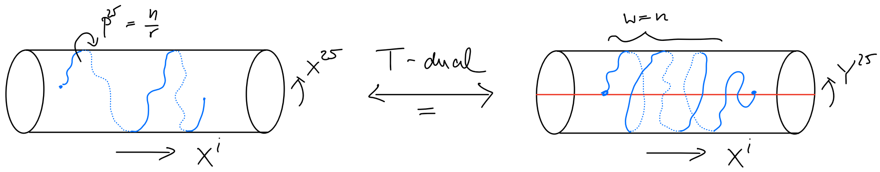


Fig. 16: Under a T-duality transformation, an open string with (NN) boundary conditions on the circle X^{25} becomes an open string with (DD) boundary conditions on the T-dual circle Y^{25}. Just as for closed strings, this duality exchanges momentum and winding numbers.

Recall from the last section that, in the spacetime perspective of open strings it is more appropriate to think of the $(p+1)$ spacetime directions in which one has Neumann boundary conditions to be filled by a Dp-brane. Then, applying T-duality on an S^1 with radius r that is contained, or "wrapped" by the worldvolume of a Dp-brane, we obtain an equivalent theory on an S^1 with radius $r' = \alpha'/r$ and a D$(p-1)$-brane that does *not* wrap the S^1, and vice versa. Said differently, T-duality exchanges the types (in this case labelled just by their dimension p) of D-branes.

6.4 Dualities beyond T

T-duality is just one example of dualities in string theory. In general, the notion of "dual" descriptions (not to be confused with proverbial dualities such as "wave-particle duality") can be an extremely powerful tool, in that it can provide complementary ways to understand the very same physical system. While dualities in concrete examples (such as electromagnetic duality) have appeared throughout the history of physics, it is fair to say that the *concept* of duality has been popularized only through string theory.

In fact, the main propellant of (super-)string theory was what is now known as the "second string revolution" in the late 90s, when it was recognized that the five different versions of the superstring where actually all related, through a *web* of dualities, among each other and an 11d theory called M-theory, see Figure 17. Thus, dualities help clarify that there is a *unique* (non-perturbative) theory underlying all superstring theories.

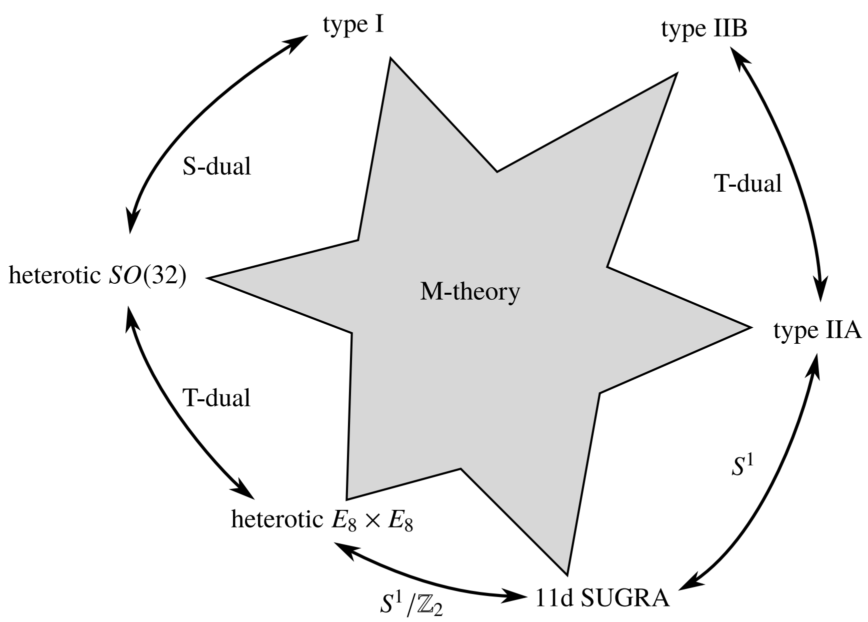


Fig. 17: A web of T-dualities and S-dualities relate the five differently looking superstring theories (type I, type IIA/IIB, heterotic $E_8 \times E_8/SO(32)$) in 10d, and 11d supergravity (via S^1 or $S^1/\mathbb{Z}_2$ compactification). These are all "perturbative" limits of a non-perturbative theory, **M-theory**, in eleven dimensions.

Besides T-duality, this web also uses "S-duality", which is a generalization of electromagnetic duality ($F_{\mu\nu} \mapsto (\star F)_{\mu\nu}$, $(\star F)_{\mu\nu} \mapsto -F_{\mu\nu}$). In the context of QFTs, S-duality maps the coupling, i.e., the coefficient $\frac{1}{4g^2}$ for $F_{\mu\nu}F^{\mu\nu}$, to $g^2/4$. That is, it exchanges a strongly-coupled QFT for a weakly-coupled QTF, both describing the same system. For this reason, S-duality is also called "strong/weak duality". This is also how it appears in the duality web: heterotic $SO(32)$ string theory at string coupling g_s is dual to type I string theory at string coupling $g_s' = 1/g_s$; and type IIB is S-*self-dual*, i.e., it is the same at string coupling g and $1/g$. This is clearly a useful kind of duality, in that it allows us to use Feynman diagrams to understand strong-coupling effects! There are many incarnations of S-duality in quantum field theory (typically supersymmetric), and oftentimes it is a string-theoretic realization of such QFT models that provides the first evidence for the existence of an S-dual description.

There is also a generalization of T-duality to higher dimensional compactification manifolds. That is, for a given string compactification on a manifold X_d, there is another compactification on $\tilde{X}_d$ that leads to the *same* physics in lower dimensions. Such a duality is known as *mirror symmetry*. The name stems from the fact that certain invariants of X_d and $\tilde{X}_d$ known as Hodge numbers, when organized into (literally) the shape of a diamond, are related to each other by *mirroring* along a diagonal line in the diamond. As can be shown (at the level of rigor acceptable to physicists) from the worldsheet perspective, i.e., in terms of equivalent σ-model CFTs, it relates compactifications of type IIA on a so-called Calabi–Yau manifold X_d to compactifications of type IIB on the *mirror-dual* Calabi–Yau $\tilde{X}_d$. This duality is highly interesting for mathematicians, because it gives a entirely novel method to learn about (hard-to-come-by) properties of X_d by performing (sometimes much easier) computations on $\tilde{X}_d$. In the context of "homological mirror symmetry", this has been formalized mathematically into a conjectured equivalence between certain (derived) categories.

Another, perhaps by now the most studied duality originating from string theory, is **gauge/gravity duality**, also known as the **holographic principle**. In this duality, the two descriptions with the same underlying physical system are so vastly different, that, prior to concrete string theoretic examples, their physical equivalence would not have been conceivable. Under this duality, a quantum field theory (typically a gauge theory) in d dimensions ought to be equivalent to a *quantum gravity* in $(d + 1)$ dimensions. Compared to others, this duality not only equates having no gravity with having gravity, but also physics in different dimensions!

In string theory, the first concrete realization of this duality is the *AdS/CFT correspondence*, which relates a conformal field theory in d-dimensional flat space with a quantum gravity theory (realized by a string theory) in $(d + 1)$-dimensional Anti-de Sitter (AdS) spacetime. The CFT can be thought of as living on the *boundary* of AdS, hence it has the flavor of being a hologram for the dynamics in the interior.[13] In a special limit, the CFT becomes a strongly-coupled field theory with large number of degrees of freedom, whereas the quantum gravity theory reduces to a weakly-coupled (supersymmetric version of) Einstein-gravity in AdS. In this case, one can learn about non-perturbative effects in CFTs using standard GR computations.

More recently, the holographic principle has also been a gateway for ideas in quantum information theory to enter the study of quantum gravity. Namely, certain insights about quantum information on the QFT-side of the duality can be translated into statements about the dual gravitational theory, which provides new insights into the quantum nature of gravity. For example, it gives a concrete way to make precise the statement "spacetime is emergent from quantum dynamics", by re-deriving geodesic equations on the gravity side (i.e., basic principles for gravitational dynamics) from quantum entanglement on the QFT-side. Furthermore, holography seems to also provide the intuition (as shown in some restrictive, but very concrete settings) how the "black hole information paradox" is resolved in a consistent theory of quantum gravity.

[13] The name "holographic principle" actually goes back further, and originates from the Bekenstein–Hawking formula for black hole entropy $S_{\text{BH}} = \frac{k}{4G_N} A$ (with k being Boltzmann's constant, and G_N the gravitational coupling constant), which scales not with the volume, but rather the surface area A of the event horizon. Since the entropy is supposed to be counting degrees of freedom of the black hole, it appears as if all the relevant data is stored on the surface, just as in a hologram.